Original Article

Proteomic Analysis of Whole Saliva in Relation to Dental Caries Resistance

(dental caries / fractions of whole human saliva / label-free quantification / mass spectrometry / proteins)

L. KULHAVÁ[1, 2], A. ECKHARDT[2], S. PATARIDIS[2], R. FOLTÁN[3], I. MIKŠÍK[2]

[1]Department of Analytical Chemistry, Faculty of Science, Charles University, Prague, Czech Republic
[2]Institute of Physiology of the Czech Academy of Sciences, Prague, Czech Republic
[3]Department of Stomatology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic

Received: November 29, 2019. Accepted: March 10, 2020.

This work was supported by Charles University, project GA UK No. 322216, within the framework of Specific University Research (SVV 260560), and by the Czech Science Foundation (No. 17-31564A).

Corresponding author: Lucie Kulhavá, Institute of Physiology of the Czech Academy of Sciences, Vídeňská 1083, 142 20 Prague 4, Czech Republic. Phone**:** (+420) 241 06 2127; e-mail**:** lucie.kulhava@fgu.cas.cz

Abbreviations: AMY1C **–** α-amylase 1 (IPI00300786), ALB – serum albumin (IPI00745872), AZGP1 – zinc-α-2-glycoprotein (IPI00166729), BPIFB1 – BPI fold-containing family B member 1 (IPI00291410), BPIFB2 – BPI fold-containing family B member 2 (IPI00296654), CA6 – carbonic anhydrase 6 (IPI00295105), CSTB – cystatin-B (IPI00021828), CST1 – cystatin-SN (IPI00305477), CST4 – cystatin-S (IPI00032294), CRNN – cornulin (IPI00297056), C6orf58 **–** protein LEG1 homologue (IPI00374315), FDR – false discovery rate, GO – gene ontology, HSPB1 – heat-shock protein β 1 (IPI00025512), H3F3A – histone H3.3 (IPI00909530), IGHA1 – immunoglobulin heavy constant α 1 (IPI00386879), Ig heavy chain V-III – immunglobulin heavy chain V-III region (IPI00854644), IGJ – immunoglobulin J chain (IPI00947235), IGK@ IGK@ – IGK@ IGK@ protein (IPI00784985), IGLV1-44 – immunoglobulin λ variable 1-44 (IPI00887169), JUP – junction plakoglobin (IPI00554711), KRT6A – keratin, type II cytoskeletal 6A (IPI00909059), KRT1 – keratin, type II cytoskeletal 1 (IPI00220327), KRT4 – keratin, type II cytoskeletal 4 (IPI01022175), KRT5 – keratin, type II cytoskeletal 5 (IPI00009867), KRT13 – keratin, type I cytoskeletal 13 (IPI00009866), KRT13 GUCA1B **–** KRT13 GUCA1B protein (IPI00930614), LCN1 – lipocalin-1 (IPI00009650), LTF – lactotransferrin (IPI00925547), LYZ – lysozyme C (IPI00019038), MS – mass spectrometry, PIGR – polymeric immunoglobulin receptor (IPI00004573), PIP – prolactin-inducible protein (IPI00022974), RAB10 – Ras-related protein Rab 10 (IPI00016513), S100A8 – protein S100-A8 (IPI00007047), S100A9 – protein S100-A9 (IPI00027462), S100A14 – protein S100-A14 (IPI00010214), TF – serotransferrin (IPI00022463), TGM3 – protein-glutamine γ-glutamyltransferase E (IPI00300376), ZG16B – zymogen granule protein (IPI00060800).

Abstract. Saliva contains possible biomarkers that are associated with dental caries. The present study aimed to analyse differences in the abundance of proteins in the saliva between caries-positive (CP; N = 15) and caries-free (CF; N = 12) males and to compare differences in the abundance of proteins between two saliva sample fractions (supernatant and pellet). We found 14 differently significantly expressed proteins in the CF group when comparing the supernatant fractions of the CP and CF groups, and three proteins in the pellet fractions had significantly higher expression in the CP group. Our results indicate very specific protein compositions of the saliva in relation to dental caries resistance (the saliva of the CP group mainly contained pellet proteins and the saliva of the CF group mainly contained supernatant proteins). This was the first time that the saliva pellet fraction was analysed in relation to the dental caries status. We detected specific calcium-binding proteins that could have decalcified enamel in the saliva pellet of the CP group. We also observed significantly up-regulated immune proteins in the saliva supernatant of the CF group that could play an important role in the caries prevention. The particular protein compositions of the saliva pellet and supernatant in the groups with different susceptibilities to tooth decay is a promising finding for future research.

Introduction

Oral fluids are very important in oral health. The physiology of human saliva and salivary secretion were reviewed in Proctor (2016). Saliva includes many markers that can indicate a risk of certain diseases (Podzimek et al., 2016). Possible biomarkers for dental caries, such as salivary electrolytes, microorganisms, proteins and peptides, were also previously reviewed along with the functional properties of saliva (Gao et al., 2016). Only a

small part of the population at the age of 30 is classified as caries-free, and the reason for the resistance is unknown. There are limited studies on the differences in the saliva protein composition between persons with carious teeth and persons without caries (Al-Tarawneh et al., 2011; Laputková et al., 2018). Differences in the abundance of salivary proteins between caries-susceptible and caries-free persons and a gender comparison were presented in our previous study (Kulhavá et al., 2018). Age-specific variations in the salivary proteome of caries-susceptible adults and elderly people have also been investigated (Wang et al., 2018a). The influence of salivary protein composition on *in vitro* dental pellicle and their correlation with dental caries was studied by two-dimensional electrophoresis (Vitorino et al., 2006). Salivary protein roles as predictors of caries risk were described in a recent review (Laputková et al., 2018). However, studies on the protein composition of saliva in relation to dental caries have varied (different approaches, methods, and use of gender-based comparisons). A study by Vitorino et al. described an assessment of pre-treatment saliva samples using in-gel and off-gel approaches, and they were the first to present the pellet fraction protein content (but not related to caries) (Vitorino et al., 2012). All comparative studies of caries-positive and caries-free individuals have analysed only the supernatant saliva fraction. The present study, based on analyses by mass spectrometry (MS) (and label-free mass spectrometry quantification), was performed to compare differences in the salivary protein abundance between caries-free and caries-positive males. We compared differences in the abundance of proteins in two saliva sample fractions: the supernatant and the pellet. Here, the pellets were analysed in this way for the first time, and a unique comparison between these saliva fractions was carried out.

Material and Methods

Saliva sample collection

Unstimulated whole saliva was collected from 27 healthy male non-smoking volunteers (without diabetes mellitus). The subjects were categorized according to the DMFT index (decayed, missing, and filled teeth index) into two groups: caries-positive (CP) (N = 15, aged 38.4 ± 5.6, DMFT ranging from 7 to 12) and caries-free (CF) (N = 12, aged 31.8 ± 7.6, DMFT ranging from 0 (N = 10) to 1(N= 2)). All procedures performed in the studies involving human participants were in accordance with the Ethical Standards and with the World Medical Association Declaration of Helsinki (version 2000). All of the volunteers were requested not to eat or drink in the morning and to brush their teeth for 1–2.5 h prior to the trial. Harvesting a 1 ml sample of saliva was performed between 8 and 10 a.m. The samples were kept on ice, and protease inhibitors (cOmplete Protease Inhibitor Cocktail Tablets (Roche Diagnostics GmbH, Mannheim, Germany)) were added to inhibit the protease activity. The unstimulated whole-saliva samples were frozen at –80 °C until further analysis.

Initial sample preparation

The saliva samples (700 µl) were centrifuged at 13,000 *g* for 30 min at 4 °C. The supernatant and pellet (P) of each sample were retained separately. The obtained supernatants and pellets were then aliquoted for use. Each sample was divided into seven parts; four parts were used for two-dimensional electrophoresis, two parts were used for two-dimensional difference gel electrophoresis and one part was used for MS analysis and label-free (LF) quantification. The pellets were lyophilised before distribution. Seven samples (supernatant/pellet) of human saliva (caries-positive, N = 7; caries-free, N = 7) were used for two-dimensional electrophoresis because high variability between each saliva sample was identified (verified by the previous one-dimensional electrophoresis), along with different protein concentrations in individual saliva samples. Four samples (supernatant/pellet) of human saliva (caries-positive, N = 4; caries-free, N = 4) were analysed by two-dimensional difference gel electrophoresis because of the cost of access method.

Preparation of samples for two-dimensional electrophoresis and two-dimensional difference gel electrophoresis

Supernatant proteins were precipitated using a final concentration of 10% (v/v) trichloroacetic acid and 0.12% (w/v) dithiothreitol (Jehmlich et al., 2013). After vortexing followed by incubation at 25 °C for 15 min, the precipitated protein was concentrated by centrifugation (13,000 *g*, 15 min, 4 °C). Then we used a protocol modified from Gonçalves et al. (2011); ice-cold acetone was used without dithiothreitol. Protein pellets from the collected supernatants (S) were washed three times with ice-cold 100% acetone, lyophilised and stored at –80 °C. Pellet proteins (P) were washed with ice-cold acetone and acetonitrile, lyophilised and stored –80 °C.

MS sample preparation – supernatant (S) and pellet (P) protein

Protein from supernatant and pellet samples was obtained similarly as in the sample preparation for the two-dimensional gel electrophoresis. Pellet proteins were digested in a solution containing 0.1 M NH_4HCO_3 and 0.2 M trypsin (1/100) (w/w; trypsin/sample) for 22 h at 37 °C. All samples were supplemented with 0.1 M NH_4HCO_3 to obtain the same volume, and the samples were lyophilised. Samples were dissolved in 100 µl of 2% formic acid.

LC-MS/MS analysis of tryptic digests

A nano liquid chromatography (n-LC) apparatus (Proxeon Easy-nLC; Proxeon, Odense, Denmark) was used for analysing the protein digests and was coupled

to a MaXis quadrupole time-of-flight (Q-TOF) mass spectrometer (Bruker Daltonics GmbH, Bremen, Germany) with a nanoelectrosprayer.

The samples were injected onto an NS-MO-10 Biosphere C18 pre-column with an NS-AC-11-C18 Biosphere C18 column, both manufactured by NanoSeparations (Nieuwkoop, Holland). The injection volume was 5 µl. The peptide separation was achieved via a linear gradient between mobile phase A (water) and B (acetonitrile), both containing 0.1% (v/v) formic acid. The separation was started by running the system with 5% mobile phase B, followed by a gradient elution to 30% B in 70 min. The next step was a gradient elution to 50% B over 10 min and then a gradient to 100% B over 10 min. Finally, the column was eluted with 100% B for 30 min. Equilibration between the runs was achieved by washing the column with 5% mobile phase B for 5 min. The flow rate was 0.25 µl/min, and the column was held at ambient temperature (25 °C). Auto MS/MS with active exclusion (after 2 spectra and release after 0.8 min) was used for MS/MS analyses. The MS settings for the MS and MS/MS experiments were as previously published (Mikšík et al., 2018), with the spectra rate set to 1 Hz and with the auto MS/MS option set to "off" for MS measurements.

The MS/MS data were processed using ProteinScape software (version 3.0, Bruker Daltonics GmbH). Proteins were identified by correlating the tandem mass spectra of saliva samples to the International Protein Index (IPI databases) and UniProt databases.

Database searches were performed with the taxonomy restricted to *Homo sapiens* to remove protein identification redundancy. Only significant hits (Mascot score ≥ 80 for proteins; ≥ 20 for peptides, http://www.matrixscience.com) were accepted.

Label-free quantification

For the label-free quantitation, an MS chromatogram was measured for each sample, and the collected data were processed with Profile Analysis software (version 2.1, Bruker Daltonics GmbH) to obtain the peptide composition values for the individual samples (caries-positive and caries-free). Each peptide must be found in at least 50 % of the samples in each group (caries-positive and caries-free groups). The peptides that fulfilled this criterion were evaluated by Student's *t*-test, and significance was defined as $P < 0.05$.

Protein fraction uniqueness by MS/MS

The criterion for identifying protein fraction (pellet (P) or supernatant (S)) uniqueness was the presence of the protein in at least 50 % of the samples of one fraction (in CP and/or CF groups). The maximum tolerance of protein impurity was determined as the presence of two of its peptides in the other fraction.

Comparison of significant and unique fraction proteins by database for annotation

We used the computational prediction of Enrich (https://amp.pharm.mssm.edu/Enrichr) to provide functional and other interpretation of the protein's predicted potential targets. We clarified the relevant functions linked to the predicted gene list, the proteins (their genes) by Gene Ontology (GO) using Enrichr. For each GO term, the P value following multiple detection corrections, such as false discovery rate (FDR) corrections, was calculated (Chen et al., 2013). We compared pellet versus supernatant proteins in both ways: 1) unique proteins in fractions, and 2) significantly changed proteins.

Results

1/ Protein composition of the fractions

We analysed all S and P samples obtained from the caries-positive and caries-free subjects. The aim was to find unique proteins that were present in only one fraction (S or P). The results are summarised in Fig. 1. A Venn diagram was created and included those protein abbreviations where at least one peptide that met the above criterion was found. In total, we identified 207 proteins across all samples (27 supernatant samples and 27 pellet samples). Approximately half of the proteins were found only in the supernatants, a quarter of the proteins were found only in the pellets, and a quarter of the proteins were found in both fractions.

We detected 1,520 peptides of which 80 met our criterion for fraction uniqueness (see Material and Methods). These peptides belonged to 31 proteins. We identified seven proteins (annexin A1; actin; cornulin; 14-3-3 protein ς; histone H2B type 1-K; lysozyme C; and histone H3.3) that exceeded the impurity criterion (unique proteins must be identified in only one fraction).

2/ Composition differences of supernatants and pellets between caries-positive and caries-free samples (label-free quantification and two-dimensional electrophoresis)

We found 162 valid proteins in supernatant samples by the label-free quantification method. The following 14 proteins had significantly higher expression levels in the supernatant samples of caries-free subjects: α-amylase 1; serum albumin; protein S100-A9; immunoglobulin heavy variable 4-31; immunoglobulin heavy constant α 1; IGK@ IGK@ protein; apolipoprotein A-I; zymogen granule protein 16 homologue B; immunoglobulin heavy variable 1-44; cystatin B; lysozyme C; polymeric immunoglobulin receptor; annexin A1; and prolactin-inducible protein) (Table 1A).

We found 73 valid proteins in the pellet samples by the label-free quantification method. We identified three

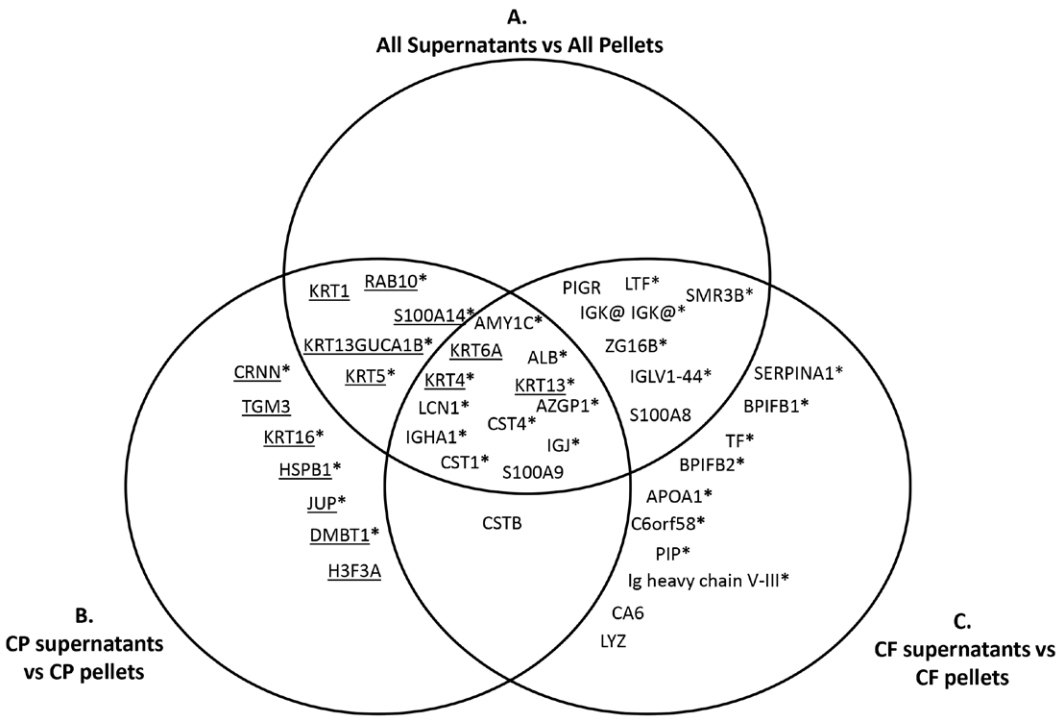

Fig. 1. Comparison of protein composition of saliva supernatants and pellets
The criterion was described in Material and Methods – Protein fraction uniqueness by MS/MS). Proteins in which at least one peptide (important peptide) was found according to the criterion in the supernatants (normal font) and in the pellets (underlined) are shown. Protein abbreviations with "*" indicate that all peptides of the protein were found only in the fraction in which the important peptide was found.

proteins (annexin A1, cornulin and 14-3-3 protein ς) with higher expression levels in the pellet samples of caries-positive subjects (Table 1B).

We compared the supernatant samples of caries-positive (N = 7) and caries-free subjects (N = 7) using two-dimensional electrophoresis (2DE). We found a significant difference in the level of polymeric immunoglobulin receptor (Student's *t*-test: P = 0.015; fold change CP/CF = 2.7). Additionally, we compared the pellet samples of caries-positive (N = 4) and caries-free subjects (N = 4) using two-dimensional difference gel electrophoresis. We found significant differences in the levels of annexin A1 (two spots; Student's *t*-test: P = 0.031 and P = 0.015; fold change CP/CF=0.08 and 0.18). The samples were applied to broad-range 7 cm, pH 3–10 NL IPG strips (Bio-Rad). Isoelectric focusing was done with a Protean IEF Cell system. After the equilibration step, samples were transferred to 12.5% SDS-polyacrylamide gel (supernatant samples), 10.5% SDS-polyacrylamide gel (pellet samples). Analysis of 2DE and DIGE gels was done using the PDQuest™ software (Bio-Rad), version 8.0.1. The local regression model was chosen as a normalization method. Spot was used to select statistically significant differential spots (Student's *t*-test; $P \leq 0.05$). Spots with differential expression were excised from the gels. Analyses of the tryptic digests were provided with nano-liquid chromatography tandem mass spectrometry.

3/ Comparison of significant and unique fraction proteins by database for annotation

We used GO analysis to identify significantly enriched GO terms (GO enrichment can directly reflect the distribution of target proteins (their genes) for each enriched GO term (GO Biological Process, and Human Gene Atlas)(https://amp.pharm.mssm.edu/Enrichr/; 15th November 2019)). We observed some significantly enriched biological processes in comparison of significantly changed proteins (and also unique proteins) in the CF group: defence response to bacterium (GO:0042742) (adjusted P value: $P < 0.010$; including these proteins: immunoglobulin heavy constant α 1, lysozyme C, immunoglobulin heavy variable 4-31, protein S100A9) and receptor-mediated endocytosis (GO:0006898) (adjusted P value: $P < 0.046$; including these proteins: serum albumin, immunoglobulin heavy variable 1-44, apolipoprotein A-I, immunoglobulin heavy constant α 1) (no significance in the CP group). We also observed a significantly enriched source of protein origin (by Human Gene Atlas) – "trachea" in the CF group (adjusted P value: $P < 0.0011$; including these proteins: polymeric immunoglobulin receptor, prolactin-inducible protein, lysozyme C) (no significance in the CP group).

Table 1A. Significantly over-expressed proteins in the saliva supernatants of the CF group in comparison with the CP group

Accession Number	Protein (peptide)	Total number of peptides[a]	Number of peptides[b]	P[c]	Average CS : Average CF[d]	MASCOT Score	n CS : n CF[e]	P (Average)[f]
IPI00300786 P04745	**α-Amylase 1**	46	10					1.0*10-5
	R.TSIVHLFEWR.W			0.0449	0.061	72.3	10 : 12	
	R.ALVFVDNHDNQR.G			0.0483	0.121	101.5	12 : 12	
	K.LGTVIR.K			0.117	0.112	46.4	11 : 10	
	R.NMVNFR.N			0.094	0.125	30.5	11 : 09	
	R.YQPVSYK.L			0.052	0.073	41.8	11 : 11	
	R.LSGLLDLALGK.D			0.101	0.044	88.7	11 : 11	
	K.NWGEGWGFMPSDR.A			0.164	0.167	96.8	10 : 08	
	K.TGSGDIENYNDATQVR.D			0.102	0.076	147.6	10 : 09	
	K.GFGGVQVSPPNENVAIHNPFRPWWER.Y			0.097	0.172	109.2	10 : 09	
	R.YFENGKDVNDWVGPPNDNGVTK.E			0.975	1.008	97.4	11 : 09	
IPI00745872 P02768	Serum albumin	31	3					4.2*10-3
	K.KVPQVSTPTLVEVSR.N			0.0381	0.092	93.2	9 : 11	
	K.LVNEVTEFAK.T			0.053	0.234	68.6	10 : 12	
	R.FKDLGEENFK.A			0.052	0.043	78.1	9 : 11	
IPI00027462 P06702	Protein S100-A9	17	8					2.4*10-3
	K.MHEGDEGPGHHHKPGLGEGTP.-			0.047	0.301	70.3	13 : 11	
	K.MSQLER.N			0.133	0.662	44.0	11 : 08	
	K.DLQNFLK.K			0.174	0.363	25,8	13 : 09	
	R.LTWASHEK.M			0.071	0.178	45.3	12 : 09	
	K.DLQNFLKK.E			0.071	0.424	31,5	13 : 08	
	K.LGHPDTLNQGEFK.E			0.18	0.513	104.8	14 : 09	
	K.VIEHIMEDLDTNADK.Q			0.192	0.307	130.5	12 : 09	
	R.NIETIINTFHQYSVK.L			0.69	1.173			
IPI00645363 P0DP07	Immunoglobulin heavy variable 4-31	14	1					
	R.VVSVLTVLHQDWLNGK.E			0.0477	0.113	120.9	12 : 10	
IPI00386879 P01876	**Immunoglobulin heavy constant α 1**	13	5					2.4*10-8
	R.WLQGSQELPR.E			0.0183	0.025	91.8	14 : 12	
	R.DASGVTFTWTPSSGK.S			0.0424	0.063	91.7	14 : 12	
	R.VAAEDWK.K			0.016	0.023	42.4	11 : 08	
	K.YLTWASR.Q			0.034	0.05	39.6	13 : 11	
	K.TPLTATLSK.S			0.018	0.039	59.8	13 : 11	
IPI00784865 P01834	**Immunoglobulin κ constant**	10	2					3.4*10-2
	K.DSTYSLSSTLTLSK.A			0.0226	0.1159	78.0	14 : 10	
	K.VDNALQSGNSQESVTEQDSK.D			0.0303	0.014	181.9	12 : 09	
IPI00021841 P02647	Apolipoprotein A-I	12	1					
	R.DYVSQFEGSALGK.Q			0.0255	0.2167	65.9	12 : 11	
IPI00060800 O60844	Zymogen granule protein 16 homolog B	9	2					2.8*10-2
	K.YFSTTEDYDHEITGLR.V			0.0392	0.0387	136.4	12 : 09	
	R.VSVGLLLVK.S			0.0228	0.1202	33.7	11 : 08	
IPI00887169 P01699	Immunoglobulin heavy variable 1-44	8	1					
	K.AGVETTTPSK.Q			0.0305	0.0397	64.7	15 : 11	
IPI00021828 P04080	Cystatin B	6	4					5.6*10-5
	K.SQVVAGTNYFIK.V			0.0413	0.0542	69.6	14 : 11	
	R.VFQSLPHENKPLTLSNYQTNK.A			0.487	0.139	132.0	11 : 09	
	K.VHVGDEDFVHLR.V			0.071	0.048	115.3	12 : 10	
	K.FPVFK.A			0.15	0.147	32.4	10 : 08	
IPI00019038 P61626	Lysozyme C	8	2					9.4*10-3
	R.STDYGIFQINSR.Y			0.0255	0.2167	76.6	12 : 11	
	R.LGMDGYR.G			0.086	0.2395	41.7	10 : 08	
IPI00218918 P10107	Annexin A1	29	1					
	R.ALYEAGER.R			0.0413	0.129	41.2	9 : 09	
IPI00004573 P01833	Polymeric immunoglobulin receptor	21	2					1.1*10-2
	R.LFAEEK.A			0.0297	0.05	43.5	11 : 11	
	K.VYTVDLGR.T			0.0246	0.015	54.8	11 : 10	
IPI00022974 P12273	Prolactin-inducible protein	5	1					
	R.FYTIEILKVE.-			0.0299	0.3216	40.0	9 : 10	

Table 1B. Significantly over-expressed proteins in the saliva pellets of the CP group in comparison with the CF group

Accession Number	Protein (peptide)	Total number of peptides[a]	Number of peptides[b]	P[c]	Average CS : Average CF[d]	MASCOT Score	n CS : n CF[e]	P (Average)[f]
IPI00218918 P04083	**ANXA1 Annexin A1**	44	5					0.282
	R.ALYEAGER.R			0.0379	2.699	40.0	13 : 08	
	K.GDRSEDFGVNEDLADSDAR.A			0.301	1.XI	89.9	12 : 07	
	K.GVDEATIIDILTKR.N			0.172	I.65	104.3	11 : 07	
	R.SEDFGVNEDLADSDAR.A			0.086	I.38	118.3	14 : 08	
	K.GVDEATIIDILTK.R			0.218	0.45	92.8	12 : 07	
IPI00411765 P31947	**SFN 14-3-3 protein ς**	8	1					
	K.SNEEGSEEKGPEVR.E			0.00462	**2.968**	**117.9**	**12 : 07**	
IPI00297056 Q9UBG3	**CRNN Cornulin**	16	0					0.175
	M.PQLLQNINGIIEAFRR.Y			0.041	2.369	88.8	12 : 05	
	R.WMQVSNPEAGETVPGGQAQTGASTESGRQ-EWSSTHPR.R			0.048	2.432	113.5	7 : 05	
	R.ISPQIQLSGQTEQTQK.A			0.349	1.03	145.2	8 : 05	

Accession Number – IPI database and Uniprot; results in bold – these peptides fulfilled significance (ratios of their intensities are statistically significant); [a] – the number of valid peptides of the protein; [b] – the number of peptides that were found in at least 50 % of the samples in the group; [c] – significance; [d] – the ratio of the CS and CF averages obtained from the individual intensity data for the particular peptide for each sample of the Profile Analysis program; [e] – the number of samples in the group where the peptide was found; [f] – *t*-test of averages of all peptides that were found in at least 50 % of samples in the group (Average CS : Average CF)

Discussion

The present study provides a proteomic comparison of the saliva fractions obtained from caries-positive and caries-free males. It was observed that the protein composition differs between the groups (caries-positive vs caries-free) and also between the saliva fractions (supernatant vs pellet), which is a unique observation to date.

1/ Comparison of saliva fractions (S vs P) via criterion protein fraction uniqueness

Our comparison of the two fractions (supernatant and pellet) was carried out from three perspectives. First, we compared all supernatants vs all pellets from both the CP and CF groups together, identifying 207 different proteins of which 25 fulfilled our criterion for protein fraction uniqueness (circle "A" in Fig. 1). Second, a comparison was made of the supernatants vs pellets obtained from the CP group. A total of 145 proteins were identified in the CP group, of which 25 were classified as unique to a fraction (circle "B" in Fig. 1). Third, a comparison was made of the supernatants vs pellets of the CF group. A total of 185 proteins were identified in the CF group, of which 30 were classified as unique to a fraction (circle "C" in Fig. 1). We found 12 common proteins among the three comparisons (intersection of circles "A", "B" and "C" in Fig. 1), which is surprisingly only 29 % out of all unique proteins. One of the study's aims was to identify proteins that were found only in the pellet fraction, because this fraction has been virtually neglected to date in the literature. In total, 15 unique proteins were found in the pellet fractions. The results of the intragroup comparisons of both the CP (circle "B") and CF (circle "C" in Fig. 1) groups were remarkable in terms of the protein distribution observed between the pellet and the supernatant. We found specific protein distributions for the two separate groups of subjects. In the CP group, 15 proteins were uniquely detected only in the pellet, and only 10 were unique to the supernatant (Fig. 1). The opposite situation occurred in the CF group, where only three proteins were uniquely detected in the pellet, but 27 proteins were specifically detected in the supernatant (Fig. 1). These results indicate very specific protein compositions of the saliva in the different groups (CP vs CF), showing that the saliva of the CP group mainly contains pellet proteins and the saliva of the CF group mainly contains supernatant proteins. These findings indicate the need for further study in this area to gain a better understanding of the specific mechanisms of anti-caries protection. It appears that CF individuals accumulate their specific proteins in the supernatant fraction (possibly by blocking aggregation) and CP individuals accumulate their specific proteins in the pellet fraction (possibly by enhanced aggregation).

We found 12 proteins in the pellet fraction that fulfilled our criterion for protein fraction uniqueness (circle "B" = CP group in Fig. 1) and 17 proteins in the supernatant fraction that fulfilled our criterion for protein faction uniqueness (circle "C" = CF group in Fig. 1). Proteins such as protein S100-A14 (S100A14), cornulin (CRNN), protein-glutamine γ-glutamyltransferase E (TGM3), junction plakoglobin (JUP), deleted in malignant brain tumours 1 protein (DMBT1) are calcium ion binding, calcium-dependent protein binding or cadherin binding, which are dependent on calcium Ca^{2+} ions to function. These proteins were found uniquely in the pellet (caries-positive), and calcium-binding proteins might

be associated with dental caries through the mechanism of calcium-dependent aggregation. These proteins might have been detected in caries-positive individuals because these proteins are better retained in the saliva. There may just be an association between the disease state rather than mechanistic contribution to the disease state. Proteins such as lactotransferrin (LTF), protein S100-A8 (S100A8), BPI fold-containing family B member 1 (BPIFB1), BPI fold-containing family B member 2 (BPIFB2) and lysozyme (LYZ) help to provide an antimicrobial humoral response. These proteins were unique to the supernatant of the CF group. The proteins lactotransferrin, immunoglobulin λ variable 1-44 (IGLV1-44), protein S100-A8 and lysozyme support the immune response. Overall, the roles of these proteins (antimicrobial humoral response and immune response) may play a major part in the caries prevention. One quarter (~40) of the total identified proteins were found only in the pellet fraction but did not meet our criterion. These are from interesting groups of proteins such as plakins, keratins and histones. The detected keratins, plakins and tubulins belong to the group of cytoskeleton proteins. We also found periplakin, desmoplakin, plakophilin 1, envoplakin, and fillaggrin, which are intermediate filament-binding proteins (Ruhrberg et al., 1997). These findings indicate that saliva pellets should also be analysed because they probably contain proteins that are important for the aetiology of tooth caries. In contrast, our results suggest that proteins supporting anti-caries defence are mainly present in the saliva supernatant.

2/ Comparison of caries-positive and caries-free groups

The primary intention of this study was to compare the saliva samples obtained from caries-positive and caries-free males using three methods (two-dimensional electrophoresis; two-dimensional difference gel electrophoresis; and label-free quantification). However, we found that two-dimensional electrophoresis was not appropriate for these comparisons due to the insufficient repeatability of the method, the small number of samples, and their differing concentrations. One other reason was the incomplete dissolution of the samples, mainly the high abundance of protein in the lysis buffer. Only ~40 spots were evaluated additionally with label-free quantification (comparison of supernatant fraction – 162 valid proteins; comparison of the pellet fraction – 73 valid proteins). We observed 14 proteins in the supernatant fraction with significantly higher expression in the CF group (comparison of samples of CP supernatants vs CF supernatants by label-free quantification) (Table 1A). Half the number (7) of these proteins are immuno- and/or antibacterial-proteins (lysozyme C, protein S100-A8, immunoglobulin heavy variable 1-34, immunoglobulin heavy constant α 1, immunoglobulin κ constant, immunoglobulin λ variable 1-44, polymeric immunoglobulin receptor). These significantly up-regulated proteins in the CF group could play an important role in the caries prevention.

Three proteins in the pellet fractions with significantly higher expression levels were observed in the caries-positive group (comparison of samples of caries-positive pellets vs caries-free pellets by label-free quantification) (Table 1B). The concentration of annexin A1 was significantly higher in the CF supernatant and also in the CP pellet (by both methods: MS and 2DE). These results appear to be contrasting; however, they are understandable in the context of the main results of this study showing possible specific protein aggregation according to the caries susceptibility (different ratio of proteins (pellet/supernatant) in the CP and CF groups). In this context, annexin could display higher protein aggregation in the saliva of the caries-positive group or lower protein aggregation in the saliva of the CF group. We have no simple explanation for this possible sample-specific aggregation. Vitorino et al. (2012) described a saliva sample preparation method, including the pellet fraction, and demonstrated it using three saliva samples obtained from males. In contrast, our study presents the results of analysis of both the supernatant and pellet fractions. These results were eventually used to investigate the differences in the protein (peptide) expression according to the dental caries status (in the pellet for the first time).

Our results are partly consistent with the work of Wang et al. (2018a). They analysed the saliva supernatant using iTRAQ-coupled LC-MS/MS. They found 14 age-specific proteins associated with caries and verified non-age-specific proteins (histatin-1) and BPI (fold-containing family B member 1) as important candidate biomarkers for common dental caries. In the current study, we found that cornulin (in the pellet fraction) was more highly expressed in the caries-positive group. Protein S100A9 was more highly expressed in the caries-free group and was found in both the supernatant and pellet fractions. These results are in agreement with those of Wang et al. (2018a). The observed differences in results may be due to the analysis of salivary samples (only male volunteers vs female and male volunteers together, age average of sample donations, different mass spectrometry analysis approach). Another i-TRAQ-coupled LC-MS/MS methods was used in the Wang's studies characterizing the healthy and cariogenic salivary proteome and determining the changes in salivary protein expression of children with varying degrees of active caries (Wang et al., 2018b).

Compared to our previous study (Kulhavá et al., 2018), we found two disagreements (α-amylase, protein S100A9) and one agreement (lysozyme). The reason for this could be that the samples were prepared by different procedures (purification conditions and sampling). Additionally, the samples were analysed by MS, and the evaluation of the results was also performed by a different approach compared to our previous publication (Kulhavá et al., 2018). Cornulin, 14-3-3 protein ς and annexin A1 were found with significantly higher ex-

pression in samples of the pellet in the caries-positive group (comparison of samples of caries-positive pellets vs samples of caries-free pellets). Cornulin is a calcium ion-binding protein (Contzler et al., 2005). Cornulin expression is associated with lesion progression in oral epithelial dysplasia (Santosh et al., 2019) and with the grade of dysplasia in predicting malignant progression of oral leukoplakia (Schaaij-Visser et al., 2010). Protein 14-3-3 (SFN) is a cadherin-binding protein. SFN was found to be decreased in the whole unstimulated saliva of subjects with generalised aggressive periodontitis compared to healthy volunteers (Wu et al., 2009). Annexin A1 binds to phospholipids in a Ca^{2+}-dependent manner (Mailliard et al., 1996). These results complement our findings that more calcium-binding proteins (calcium-dependent) are found in the pellet fraction of the CP group (Fig. 1, Table 1B), which could be a reason for the decalcification of enamel.

Annexin A1 is an anti-inflammatory and inflammation-resolving protein that is likely involved in a number of chronic conditions such as salivary gland disorders, periodontal disease and diabetes mellitus. The differences in annexin expression may also be related to the periodontal status. This aspect must be included in the experimental design (the effect on the results with respect to other diseases must be excluded or clarified; to be considered in the evaluation in future study; however, persons with diabetes mellitus were not included in the study).

3/ Comparison of significant and unique fraction proteins by database for annotation

We observed some significantly enriched biological processes when comparing the significantly changed and also unique proteins of both groups (CP and CF). First of all, these concern specific immune processes in the CF group: defence response to bacterium (GO:0042742), which confirms our results from label-free quantification (7 significantly up-regulated immune and/or antibacterial proteins in the CF group, Table 1A) and declares their important defence role against caries. Another significantly enriched biological process in the CF group is receptor-mediated endocytosis (GO:0006898). This process, by which cells absorb proteins, hormones, metabolites and viruses, may play an important role in oral health; however, there are not enough data for its better characterization. We also observed a significantly enriched source of the protein origin: "trachea" in the CF group. The tracheal origin in the CF group may just represent the lack of comprehensive investigation in the oral cavity and not a transfer. It is not uncommon for tracheal and nasal origins to show up in the oral cavity, and they may actually not be transient proteins but exist in individuals at all times.

This study shows the protein specificity of the saliva fractions after centrifugation in relationship to the dental caries experience of the saliva sample donor. We surmise that specific proteins facilitating formation of the tooth caries are found both in the saliva supernatant and pellet fractions. We detected seven significantly up-regulated specific immune proteins in the saliva supernatant of the CF group that could play an important role in the caries prevention. In the pellet of the CP group, we observed five specific calcium-binding proteins that could have decalcified enamel. The particular protein composition of the saliva pellet and supernatant in the groups with different susceptibilities to tooth decay is a promising finding for future research.

References

Al-Tarawned, S. K., Border, M. B., Dibble, C. F., Bencharit, S. (2011) Defining salivary biomarkers using mass-spectrometry-based proteomics: a systematic review. *OMICS* **15**, 353-361.

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., Clark, N. R., Ma´ayan, A. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128.

Contzler, R., Favre, B., Huber, M., Hold, D. (2005) Cornulin, a new member of the „fused gene" family, is expressed during epidermal differentiation. *J. Invest. Dermatol.* **124**, 990-997.

Gao, X., Jiang, S., Koh, D., Hsu, C. Y. S. (2016) Salivary biomarkers for dental caries. *Periodontology 2000* **70**, 128-141.

Gonçalves, Lda. R., Soare, M. R., Nogueira, F. C. S., Garcia, C. H. S., Camisasca, D. R., Domont, G., Feitosa, A. C. R., Pereira, D. A., Zingali, R. B., Alves, G. (2011) Analysis of the salivary proteome in gingivitis patients. *J. Periodont. Res.* **46**, 599-606.

Jehmlich, N., Dinh, H. D. K., Salazar, M. G., Hammer, E., Steil, L., Dhople, V. M., Schurmann, C., Holtfreter, B., Kocher, T., Völker, U. (2013) Quantitative analysis of the intra- and inter-subject variability of the whole salivary proteome. *J. Periodont. Res.* **48**, 392-403.

Kulhavá, L., Eckhardt, A., Pataridis, S., Bartoš, M., Foltán, R., Mikšík, I. (2018) Differences of saliva composition in relation to tooth decay and gender. *Folia Biol. (Praha)* **64**, 195-203.

Laputková, G., Schwartzová, V., Bániccin, J., Alexovič, M., Sabo, J. (2018) Salivary protein roles in oral health and as predictors of caries risk. *Open Life Sci*. **13**, 174-200.

Mailliard, W. S., Haigler, H. T., Schlaepfer, D. D. (1996) Calcium-dependent binding of S100C to the N-terminal domain of annexin I. *J. Biol. Chem.* **271**, 719-725.

Mikšík, I., Pataridis, S., Eckhardt, A., Sedmera, D. (2018) Analysis of Siamese Crocodile (Crocodylus siamensis) eggshell proteome. *Protein J.* **37**, 21-37.

Podzimek, S., Vondrackova, L., Duskova, J., Janatova, T., Broukal, Z. (2016) Salivary markers for periodontal and general diseases. *Dis. Markers* **2016,** Article ID 9179632.

Proctor, G. B. (2016) The physiology of salivary secretion. *Periodontology 2000* **70**, 11-25.

Ruhrberg, C., Hajibagheri, M. A., Parry, D. A., Watt, F. M. (1997) Periplakin, a novel component of cornified envelopes and desmosomes that belongs to the plakin family and forms complexes with envoplakin. *J. Cell. Biol.* **139**, 1835-1849.

Santosh, N., McNamara, K. K., Beck, F. M., Kalmar, J. R. (2019) Expression of cornulin in oral premalignant lesions. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **127**, 526-534.

Schaaij-Visser, T. B., Bremmer, J. F., Braakhuis, B. J., Heck, A. J., Slijper, M., Van Der Waal, I., Brakenhoff, R. H. (2010) Evaluation of cornulin, keratin 4, keratin 13 expression and grade of dysplasia for predicting malignant progression of oral leukoplakia. *Oral Oncol.* **46**, 123-127.

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

Vitorino, R., Guedes, S., Manadas, B., Ferreira, R., Amodo, F. (2012) Toward a standardized saliva proteome analysis methodology. *J. Proteomics* **75**, 5140-5165.

Wang, K., Wang, X., Zheng, S., Niu, Y., Zheng, W., Qin, X., Li, Z., Luo, J., Jiang, W., Zhou, X., Li, W., Zhang, L. (2018a) iTRAQ-based quantitative analysis of age-specific variations in salivary proteome of caries-susceptible individuals. *J. Transl. Med*. **16**, 293.

Wang, K., Wang, Y., Wang, X., Ren, Q., Han, S., Ding, L., Li, Z., Zhou, X., Li, W., Zhang, L. (2018b) Comparative salivary protemics analysis of children with and without dental caries using the iTRAQ/MRM approach. *J. Transl. Med*. **16**, 11.

Wu, Y., Shu, R., Luo, K. J., Ge, L. H., Xie, Y. F. (2009) Initial comparison of proteomic profiles of whole unstimulated saliva obtained from generalized aggressive periodontitis patients and healthy control subjects. *J. Periodontal Res.* **44**, 636-644.