12-15-2022 2:00 PM

# Secretion of Salivary Proteins and their Interactions

Karla Tonelli Bicalho Crosara, *The University of Western Ontario*

Supervisor: Siqueira, Walter L., *University of Saskatchewan*
Joint Supervisor: Vieira, Liliani A. C., *The University of Western Ontario*
Joint Supervisor: McKenzie, Charles A., *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Medical Biophysics

# Abstract

Saliva's clinical application for disease diagnosis and monitoring is limited by incomplete knowledge of salivary protein interactions, the effect of stimulation on the salivary proteome, and if these factors impact protein identification. This thesis expands knowledge of the salivary interactome and the effects of stimulation intensity and duration on parotid saliva's proteome. First, previous in vitro studies identified 43 proteins in the histatin 1-protein network and demonstrated histatin 1's increased stability in whole saliva when interacting with amylase. We hypothesized that protein-protein-interaction databases could enlarge the histatin 1-protein network. A comprehensive histatin 1-protein network was created using STRING database, merging previous in vitro complex partners with in-silico interactors. Thirty-seven novel histatin 1 interactors were identified, demonstrating STRING's utility for studying protein-protein networks. Second, heterotypic complexes between amylase and MUC 5B, MUC 7, histatin 1 and histatin 5 have been described. Given amylase's biochemical characteristics and abundance in saliva, we theorized that amylase interacted with other proteins. Affinity chromatography, gel electrophoresis, tryptic in-gel and in-solution digestion, and mass spectrometry were used. Sixty-six proteins were identified in whole saliva's amylase-protein network. Acidic, low molecular weight proteins involved in host protection had preference in amylase's complex formation. An inclusive amylase-protein network was constructed using STRING database, opening avenue for further studies about the amylase interactome. Third, stimulation intensity and duration affect the composition of salivary gland secretions. We questioned if the proteome of the parotid gland's secretion was also affected by stimulation intensity and duration. Continuous parotid saliva secretion (0.25 and 1.00 ml/min) for 30 consecutive minutes was achieved. After in-solution digestion and mass spectrometry, five time points were used for proteome identification. Combining both flows, 169 proteins were identified. Stimulation intensity strongly affected 119 proteins, 44 were affected by both factors, and 4 by neither, suggesting possible protein-specific secretory mechanisms. This thesis demonstrates that salivary proteins participate in large complexes, that can be represented and expanded with aid of protein-protein-interaction databases. It also provides insights into the complexity of factors affecting saliva composition, such as stimulation, and highlights the importance of developing standardized protocols for salivary biomarker research.

# Keywords

Salivary Proteome, Salivary Interactome, Protein-protein-interaction Database, Parotid Gland, Saliva Secretion, Mass Spectrometry.

# Summary for Lay Audience

Saliva is formed mainly by the secretion from salivary glands. However, it also contains elements from the blood, so it might be used to diagnose oral and systemic diseases. Many studies have investigated molecules in saliva that can be used to determine the onset of diseases like tooth decay and cancer. Changes in some salivary proteins may indicate the presence of disease. There is little information about how the interactions among proteins in saliva and the stimulation of saliva production interfere with saliva's protein composition (salivary proteome). This thesis expands the knowledge of interactions between salivary proteins in the formation of complexes and the effects of intensity and duration of stimulation on the proteome of the secretion from salivary glands. The mouth's harsh environment can break proteins into small pieces. The interaction of salivary proteins can protect some proteins from degradation and improve their clinical detection. First, the STRING protein-protein-interactions database was used to construct a comprehensive representation of the proteins that interact with histatin 1. Thirty-seven novel proteins were identified in the novel histatin 1-protein network. Second, our laboratory developed an approach to identify 66 proteins that interact with amylase, the most abundant protein in saliva. A novel inclusive amylase-protein network was created using the STRING database, combining the 66 partners with additional database interactors. Third, variations in the saliva secreted by salivary glands depending on stimulation can impact the composition and diagnostic application of saliva. Differences in the proteins secreted by the parotid glands, the largest human salivary glands, due to stimulation intensity (given by two flow rates) and duration (given by 30 minutes) were demonstrated, suggesting critical implications for the development of protocols for the discovery of biological markers (biomarkers) for diseases in saliva. This thesis shows that salivary proteins participate in large complexes, that can be visualized and enlarged with assistance of protein-protein-interaction databases, like STRING. It also provides an indication of the complexity of factors affecting saliva's composition, such as stimulation, and emphasizes the importance of developing consistent procedures to investigate disease biomarkers in saliva.

# Co-Authorship Statement

Chapters 2 and 3 of this dissertation were based on published journal articles. As the first author of these manuscripts, I led all parts of the work presented. I contributed significantly to experimental design, technological development, subject recruitment, data acquisition and analysis, interpretation of results, and manuscript preparation. Dr. Walter Siqueira supervised the work in both chapters from their theoretical conception to their final publication approval and participated in the manuscript design, writing and review. Dr. Walter Siqueira was also responsible for securing the funding to support all work presented in this dissertation. All listed co-authors contributed significantly to the work presented and approved the manuscript submission. Individual contributions are listed below by chapter.

Chapter 2 was adapted from the publication titled "Merging in-silico and in vitro salivary protein complex partners using the STRING database: A tutorial", published in *Journal of Proteomics* in 2018 by Karla Tonelli Bicalho Crosara, Eduardo Buozi Moffa, Yizhi Xiao, and Walter Luiz Siqueira. Dr. Eduardo Moffa assisted with data analysis, data interpretation, and manuscript preparation and review. Dr. Yizhi Xiao contributed with technical support and data analysis. Dr. Walter Siqueira funded and supervised all the work.

Chapter 3 was adapted from the publication titled "Revealing the amylase interactome in whole saliva using proteomic approaches", published in BioMed Research International in 2018 by Karla Tonelli Bicalho Crosara, David Zuanazzi, Eduardo Buozi Moffa, Yizhi Xiao, Maria Aparecida de Andrade Machado, and Walter Luiz Siqueira. Dr. David Zuanazzi participated in the study design and the development of the in-house chromatography used in this study. Dr. Eduardo Moffa assisted with data interpretation, and manuscript preparation and review. Dr. Maria Aparecida Machado participated in the manuscript preparation and review. Dr. Yizhi Xiao contributed to all the mass spectrometric analysis and assisted with data analysis. Dr. Walter Siqueira funded and supervised all the work.

# Dedication

I would like to dedicate this work to my family for their continuous love, patience, support, and encouragement. To my loving husband, Hewerton, my partner and best friend, thank you for understanding when I couldn't be with you and our family due to work-related activities, for your patience in listening to me as I practiced my presentations, and for your support and encouragement every time I faced technical and personal drawbacks. I couldn't have done this without you. To my dear children, Samuel and Lara, who are now teenagers, thank you for understanding when mom came home disappointed because things weren't going well in the laboratory, and for cheering me up by naively saying "you can do this". Remember that it is never too late to start something new in your lives, and that you can do anything that you put your mind and heart into. I certainly hope to have inspired you with my journey. To our little furry companion, Charlie (dog), who lay by my side as I prepared this thesis. I will forever remember and envy your snores as I tried to concentrate on the work to be done. Thank you for your cheerful company.

To my parents, a constant reminder of strength, and examples of determination and courage to face changes and difficulties. Your achievements have always inspired me to do my best and never be afraid of new challenges.

To my in-laws, for accepting our decisions and cheering for our success.

To my friends and family who helped me maintain my emotional balance throughout this journey. You have no idea how important your words, hugs, and laughters are to me.

Finally, to all my former lab members, my supervisors, and my advisors, thank you for the many moments that we shared inside and outside the laboratory. I learned so much from you all.

# Acknowledgments

Many individuals contributed to my research and personal growth during the time I spent at Western. The work presented in this dissertation was possible because of their support. Their contribution has shaped the person that I am today, and it has inspired the person that I want to be in the future.

Firstly, I want to acknowledge my academic/research supervisors, Drs. Walter Siqueira, Charles McKenzie, and Liliani Vieira. Thank you for your contributions at different stages of my graduate path.

To Dr. Walter Siqueira, for opening the doors of his laboratory to me. The year I spent working in your laboratory as a volunteer research assistant equipped me with the tools to achieve success in my graduate studies. I learned so much about research and academia from you. Thank you for believing in my potential and investing in my research.

To Dr. Charles McKenzie, for always walking the extra mile to assure that I had the necessary support to conclude my graduate studies. Thank you for being the strong helping hand that I needed to stay afloat. Your assistance was critical this past year.

To Dr. Liliani Vieira, for accepting the challenge of guiding me through the last months of my studies. Your strength and determination inspired me. Your humanity and friendship nurtured me. Thank you for believing in me and directing me to the finish line. I couldn't have finished this thesis without your friendship and support.

To my advisors, Drs. Jeff Dixon and Andrew Leask. To Dr. Jeff Dixon for being an example of researcher and mentor. I will forever remember your wise words "90% of the work that we do goes wrong, and the 10% that goes well is what we publish. If it isn't working, you haven't reached that 10% yet.". Thank you for your support and kindness in the moments that I needed the most. To Dr. Andrew Leask, thank you for your contribution during the first years of my graduate studies. Your unconventional views provided interesting perspectives on my research.

Thank you also to Dr. Thomas Drysdale and SGPS for providing the support and means necessary to conclude my graduate studies.

I also thank all former members of my laboratory. Starting with Dr. Yizhi (Cindy) Xiao, our lab manager. Thank you for assuring the maximum quality of all the mass spectrometric results used in this dissertation. Your work ethic and professionalism won't be forgotten. Special thank you also to Dr. David Zuanazzi for initiating the work with amylase and for introducing me to most of the techniques used in this thesis. You were a good teacher. Thank you also for saving me every time I had technical issues with my computer, your help was appreciated. Additionally, big thanks to Cindy and David for collecting saliva countless times for my research. Many thanks also to all my former lab members (Tayebeh Basiri, Flavia Zaidam, Teresa Valente, Vinicius Silva, Mariana Pizza, Eduardo Kohara, Samia Corbi, Paula Jorge, Rajesh Gupta, Lina Martin, Claudia Pellissari, Camilla Esteves, Naima Abousseta, Laura Souto, Gabriela Mancia, Pamela, Natalia Domingues, Camila Martins, Pia Canales, and Woody Neilson), members from the Dixon Lab (Yara Houssein and Brandon) and Beier Lab (Manuela and Melina), and my friends from Medical Biophysics know that each one of you helped me in some way.

Thank you also to all the Medical Biophysics Department. To the former Graduate chairs, Drs. Grace Parraga and Aaron Ward, and the current Graduate chair, Dr. Charles McKenzie, thank you for your continued assistance. To all the administrative team, particularly Wendy Hough and Kathleen Petts, who provided invaluable help with department examinations and scholarship applications.

Finally, thank you to the following sources of funding that have supported me and my research throughout my graduate program: Natural Sciences and Engineering Council (NSERC), Canadian Institutes of Health Research (CIHR), Ontario Graduate Scholarship (OGS), Queen Elizabeth II Graduate Scholarship in Science and Technology (QEII-GSST), Schulich School of Medicine and Dentistry, Western University, and the Department of Medical Biophysics.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| WS | Whole Saliva |
| PPI | Protein-protein Interaction |
| PG | Parotid Gland |
| SMG | Submandibular Gland |
| SLG | Sublingual Gland |
| IPI | International Protein Index |
| GCF | Gingival Crevicular Fluid |
| PTM | Posttranslational Modification |
| PRP | Proline-rich Protein |
| AEP | Acquired Enamel Pellicle |
| SMP | Salivary Mucosal Pellicle |
| MS | Mass Spectrometry |
| PAGE | Polyacrylamide Gel Electrophoresis |
| 1-D | One Dimension |
| 2-D | Two Dimension |
| pI | Isoelectric point |
| DTT | Dithiothreitol |
| SDS | Sodium Dodecyl Sulfate |
| WSS | Whole Saliva Supernatant |

| | |
|---|---|
| MS/MS | Tandem Mass Spectrometry |
| ESI | Electrospray Ionization |
| MALDI | Matrix-assisted Laser Desorption Ionization |
| ESIMS | Electrospray Ionization Mass Spectrometry |
| HPLC | High-pressure Liquid Chromatography |
| m/z | Mass-to-charge |
| LC | Liquid Chromatography |
| IT | Ion Trap |
| TQ | Triple Quadrupole |
| MS1 | Preliminary Mass Spectrum |
| MS2 | Mass Spectrum of Fragment |
| CID | Collision-induced Dissociation |
| HPA | The Human Protein Atlas |
| HSPW | The Human Salivary Proteome Wiki |
| AC | Affinity Chromatography |
| Co-IP | Coimmunoprecipitation |
| SPR | Surface Plasmon Resonance |
| RP | Reverse Phase |
| CIHR | Canadian Institutes of Health Research |
| QEII-GSST | Queen Elizabeth II Graduate Scholarship in Science and Technology |

| | |
|---|---|
| MW | Molecular Weight |
| TFA | Trifluoroacetic Acid |
| BCA | Bicinchoninic Acid |
| RT | Room Temperature |
| ACN | Acetonitrile |
| SPSS | Statistical Package for the Social Sciences |
| T1 | Time Point 1 (First Minute of Collection) |
| T5 | Time Point 5 (Fifth Minute of Collection) |
| T10 | Time Point 10 (Tenth Minute of Collection) |
| T20 | Time Point 20 (Twentieth Minute of Collection) |
| T30 | Time Point 30 (Thirtieth Minute of Collection) |
| AD | Alzheimer's Disease |
| SERS | Surface Enhanced Raman Spectroscopy |

# List of Appendices

# Chapter 1

## 1 Introduction

There is a growing interest in the use of saliva as a diagnostic tool because its collection is painless, cheaper, and quicker than the collection of other biological fluids such as blood [1,2]. In addition, the saliva found in the mouth (whole saliva) is a very complex biological fluid formed by multiple sources, including blood components, which allows the investigation of biomarkers for both oral and systemic diseases [3–6]. Studies have shown that the composition of the whole saliva (WS) is more than simply a mirror of the blood [5–8]. Many salivary components are not identified in blood or are identified at higher concentrations in saliva than in blood [5,6,8–10]. The comprehensive analysis of the different salivary components is called Salivaomics [11,12], and it includes genomics (circulating DNA), epigenomics (epigenetic changes), transcriptomics (RNA), proteomics (proteins), metabolomics (metabolites), and microbiomics (microorganisms). The term PROTEOME is a "contraction of the PROTEins encoded by a given genOME" in a cell or tissue [13–17], and the salivary proteome refers to the mixture of proteins found in WS and the secretion from the salivary glands. Although the proteome derives from the genome, a proteome is richer than a genome since it also accounts for post-transcriptional and post-translational processes, such as complex formation and covalent modifications [2,13,14,16–18]. Moreover, while the genome may remain the same, the salivary proteome varies in different conditions such as physiological and pathological changes [14,18]. Thus, to fully exploit the endless clinical applications of the salivary proteome, it is essential to establish its composition and to understand the factors that influence the availability and stability of the salivary proteins in saliva.

Like most proteins in living organisms, salivary proteins participate in protein-protein interactions (PPIs) to perform their biological functions [19,20]. Salivary Interactomics refers to this new area of Salivary Proteomics that studies the interactions between salivary proteins and their biological impact [2,11]. PPIs can happen transiently or promote the formation of stable complexes, in which the original individual functions of each of the protein partners can be modulated or modified [2,21]. The formation of

protein complexes can impact the stability of the salivary proteome as demonstrated by the interaction between histatin-1 and salivary amylase [21]. Histatin 1 is known for its high susceptibility to proteolysis in the oral cavity and its antifungal activity [22,23], while salivary amylase is recognized for its participation in the initial digestion of carbohydrates in the mouth [24]. Interestingly, when the histatin 1-amylase complex was analyzed in vitro, the resistance of histatin-1 to degradation in WS was increased [21], suggesting that salivary amylase might protect its complex partners from degradation while distributing salivary proteins to different locations in the mouth as an efficient natural delivery system for other proteins and peptides [21,25]. A total of 43 proteins were previously identified in vitro to participate in heterotypic complexes with histatin 1 [21]. Could the histatin 1-protein network be enriched with information from bioinformatics tools such as PPI online databases? Considering that salivary amylase executes the important job of protecting and delivering proteins throughout the oral cavity, what other proteins participate in the amylase heterotypic protein complex? To investigate these questions, the first two objectives of this thesis focused on expanding the knowledge about the salivary interactome by examining heterotypic complexes formed by histatin 1 (Chapter 2) and amylase (Chapter 3), using the STRING protein interaction database to merge in vitro and in-silico information.

The study of the Human Salivary Proteome received great incentive in 2004 when The National Institute of Dental and Craniofacial Research announced an initiative to clarify disease pathogenesis and evaluate the influence of medications on different salivary proteins [9]. As a result, a work published in 2008 by Denny and collaborators described the proteomes of the parotid gland (PG) and submandibular/sublingual gland (SMG/SLG) ductal secretions [5]. This work largely contributed to initiating the construction of the salivary proteome of healthy individuals. This extensive research demonstrated not only differences between the proteome from the major salivary glands, but also highlighted variances in the proteome from the same salivary gland acquired by different research groups, possibly due to the use of diverse methodologies for sample collection, separation strategies, data acquisition and analysis, as well as frequent updates in the International Protein Index (IPI) database, and donor-derived variations. Differences in the composition of the secretions from salivary glands due

to stimulation during sample collection have also been reported in many studies [26–37]. Therefore, being able to anticipate expected changes in the salivary proteomic profile due to aspects related to the stimulation employed in the collection methodology is a determinant for the establishment of robust protocols for biomarker identification. The effects of intensity and duration of stimulation on the main electrolytes and the total protein concentration of the secretion from the major salivary glands, for example, have been demonstrated [31,36,37]; however, the effect of these factors on the proteome of the glandular saliva is unknown. The third objective of this thesis investigates the effect of intensity (reflected by two constant flow rates) and duration (up to 30 minutes) of stimulation on the proteomic composition of the secretion from the PG. Variations in the total protein concentration and the proteomic profile of the PG secretion are investigated in Chapter 4.

In summary, this thesis expands the knowledge about the interactions among salivary proteins and the effect of the stimulation on the secreted proteome, providing important information about factors that can interfere with the availability and stability of salivary proteins. The salivary interactome is discussed in Chapters 2 and 3. Chapter 2 expands the histatin 1-protein network and highlights the advantages of applying bioinformatics tools to combine in-silico strategies with in vitro experimental findings to rapidly advance our understanding of PPIs. Chapter 3 presents an inclusive amylase-protein network merging the salivary proteins identified in vitro in WS using different proteomics approaches, and the in-silico documented and predicted amylase heterotypic complex partners. Chapter 4 demonstrates the effect of intensity and duration of stimulation on the organic composition of the PG saliva (total protein concentration and proteome identification), highlighting the importance of understanding the impact of variations in collection protocols on our ability to establish saliva's proteomic composition. Chapter 5 presents the main conclusions of this thesis, its limitations, and future work that can be done to expand upon our findings.

## 1.1 Whole Saliva and Salivary Glands

Human saliva is usually identified as the fluid found in the oral cavity. In salivary research, however, the saliva found in the mouth is usually termed WS or mixed saliva

since it is formed mostly by the secretions from all salivary glands [38,39]. Almost all the WS volume is secreted by three pairs of major salivary glands (PG, SMG, and SLG), and between 600 and 1000 minor salivary glands [3,39]. The PGs, the largest of all salivary glands, are each located on the sides of the face, in front of the ears, and their secretion reaches the mouth via a single parotid duct (also called Stensen's duct), which opens at the level of the second upper molar [39,40]. The SMGs are located under the lower jaw and their secretions reach the floor of the oral cavity via the submandibular duct (also called Wharton's duct) [39,40]. The SLGs are located under the tongue, in the floor of the mouth, and their secretions reach the oral cavity through the major sublingual duct (also called Bartholin's duct) or directly into the mouth via various excretory ducts opening into the sublingual area (ducts of Rivinus) [39,40]. The minor salivary glands are located throughout the oral cavity (palatine, retromolar, buccal, labial, and lingual glands) and their secretion empties into the mouth via small separate ducts in the oral epithelium [39,40]. Recently, an additional pair of macroscopic salivary glands were identified in the nasopharynx near the torus tubarius, the tubarial glands [41]. Unless collected directly from the excretory ducts, WS also contains elements from the gingival crevicular fluid (GCF), cells and fluids from the oral mucosa, many microorganisms and their metabolites, food debris [3,4], and nasopharyngeal secretion [39]. The fluid secreted by each salivary gland can also be termed saliva; however, this term will be preceded by the name of the gland responsible for this secretion, for example, parotid saliva, submandibular saliva, or sublingual saliva [38,42].

### 1.1.1 Anatomy of Salivary Glands

The salivary glands are formed by a parenchymal and a stromal component [43]. The parenchyma contains the functional secretory end of the salivary gland, the acini, while the stroma includes the supportive tissues, such as blood vessels, connective tissue, myoepithelial cells, and secretory ducts [43]. The acinar cells secrete the salivary fluid into the acinar lumen of the salivary glands. From there, the fluid flows through the salivary ductal system until it reaches the mouth [43] (Figure 1.1).

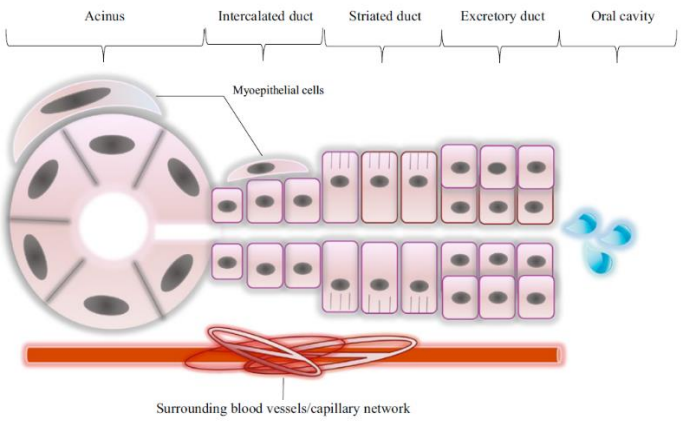



**Figure 1.1 Schematic representation of the secretory end piece of salivary glands (acinus), followed by the ductal system (intercalated, striated, and excretory duct), and the opening into the oral cavity (used with permission from [43])**

The fluid initially secreted into the glandular lumen is isotonic [4,44]. As saliva flows through the ductal system of the major salivary glands, salt is removed and saliva becomes hypotonic, allowing the tasting of salt in food [4,44]. The flow rate of saliva secretion influences the hypotonicity of saliva, with higher salt concentration present in saliva secreted at an increased flow rate [4,44]. In the ductal system, ductal cells remove sodium and chloride from the saliva. Contrarily, ductal cells play an important role in the secretion of bicarbonate in saliva, in addition to some secretion of bicarbonate by the salivary acinar cells [4,44]. The concentration of bicarbonate in the saliva secreted under stimulation is many times higher than that of unstimulated saliva [4,44,45]. The bicarbonate present in saliva is crucial in buffering the pH of saliva near neutrality, thus preventing the demineralization of tooth structure due to salivary acidic pH [4,44]. Calcium, phosphate, thiocyanate, iodide, and nitrate are also transported by salivary gland cells into saliva, with little influence of different stimulation conditions in the salivary calcium concentration [44].

The protein composition of the secretion from salivary glands depends on the characteristics of the gland acinar cells. The acinus is formed predominantly by either serous or mucous cells, or mucous cells coated by serous demilunes, and they are distributed around a central lumen [43]. According to their structural composition and their secretions, the salivary glands are classified histologically as serous, mucous, or

mixed [3,4,43,46–49](Table 1.1). The PGs are formed predominantly by serous cells and produce a watery amylase-rich fluid [43], the SMGs and SLGs contain mixed cell types (serous and mucous) and secrete a more viscous mucin-rich fluid [43]. In the recently discovered tubarial glands, both mucous and serous cell types were identified, with predominance for mucous secretion [41]. The secretions from the minor salivary glands vary. Palatine and retromolar glands secrete mucin-rich saliva; buccal and labial glands present mixed cell types, with the predominance of mucous cells; and lingual glands contain serous cells and produce a watery, lipase-rich fluid [43]. For that reason, many studies have investigated the composition of the secretion from each salivary gland separately [5,42,50–63].

**Table 1.1 Salivary glands parasympathetic innervation, histological classification, excretory ducts, secretion characteristics, flow rates and contribution to WS volume under unstimulated and stimulated collections (modified from [43] with permission)**

| *Salivary glands* | *Innervation* | *Acini cell type* | *Excretory ducts* | *Secretion* | *Flow rates (ml/min)*** | *Contribution (%) to WS volume* |
|---|---|---|---|---|---|---|
| *Parotid glands* | Glossopharyngeal nerve | Serous | Stensen's duct | Watery, amylase-rich | Resting:0.04<br>Stimulated:1.0-1.5 | Resting:25%<br>Stimulated:50% |
| *Submandibular glands* | Facial nerve | Mixed, mainly serous | Wharton's duct | Viscous, mucin-rich | Resting:0.1<br>Stimulated:0.8 | Resting:60%<br>Stimulated:35% |
| *Sublingual glands* | Facial nerve | Mixed, mainly mucous | Ducts of Rivinus and Bartholin's duct | Viscous, mucin-rich | Resting:0.14<br>Stimulated:0.12 | Resting:7%-8%<br>Stimulated:7%-8% |
| *Tubarial glands* * | | Mixed, mainly mucous | Multiple draining ducts | Mucin-rich | | |
| *Minor salivary glands* | | | Individual small ducts | | Resting:<0.05<br>Stimulated:<0.1 | Resting:8%<br>Stimulated:8% |
| *Palatine glands* | Facial nerve | Mucous | | Mucin-rich | | |
| *Buccal glands* | Facial nerve | Mixed, mainly mucous | | Mucin-rich | | |
| *Labial glands* | Facial nerve | Mixed, mainly mucous | | Mucin-rich | | |
| *Lingual glands* | Glossopharyngeal nerve | Serous | | Watery, lipase-rich | | |
| *Retromolar glands* | Glossopharyngeal and Facial nerves | Mucous | | Viscous, mucin-rich | | |

* Tubarial glands are predominantly mucous gland tissue, with multiple macroscopic draining ducts. No amylase expression was found in the tubarial glands tissue, consistent with the very low number of serous acini. Compared to the known major salivary glands, the tubarial glands are more like the sublingual glands [41].

** Flow rates of salivary glands were acquired from previous publications [4,27,34,44,50–52]

Most salivary proteins are secreted by exocytosis of protein storage granules in acinar cells [44]. Mucins, which are large highly glycosylated proteins, are examples of salivary proteins secreted via storage granules [64]. Alternatively, some salivary proteins are secreted into saliva by vesicular transport, which can also happen without fluid secretion and can cause an accumulation of proteins in the ductal system of salivary glands [44]. Both mechanisms are controlled via autonomic stimulation; however, the composition of the proteins secreted by the two mechanisms differs, and the reason for this selective protein secretion is still to be clarified[44].

### 1.1.2 Salivary Reflex and Innervation of Salivary Glands

As described by Proctor (2016)[44], the salivary reflex begins with the perception of food and tastants, such as acid or salt, by taste buds on the tongue (gustatory stimulus) and mechanoreceptors in the periodontal ligament (masticatory stimulus) around the teeth and in the oral mucosa. Afferent sensory nerves transmit signals from the tongue (facial and glossopharyngeal nerves), and from the periodontal ligaments and oral mucosa (trigeminal nerve) to the salivary centers, from where efferent parasympathetic nerves conduct signals to the salivary glands. Sympathetic efferent nerves from the thoracic spinal cord also conduct signals to the salivary glands. Nerves in the central nervous system also innervate the salivary centers, influencing nerve-mediated signals to the salivary glands (Figure 1.2).

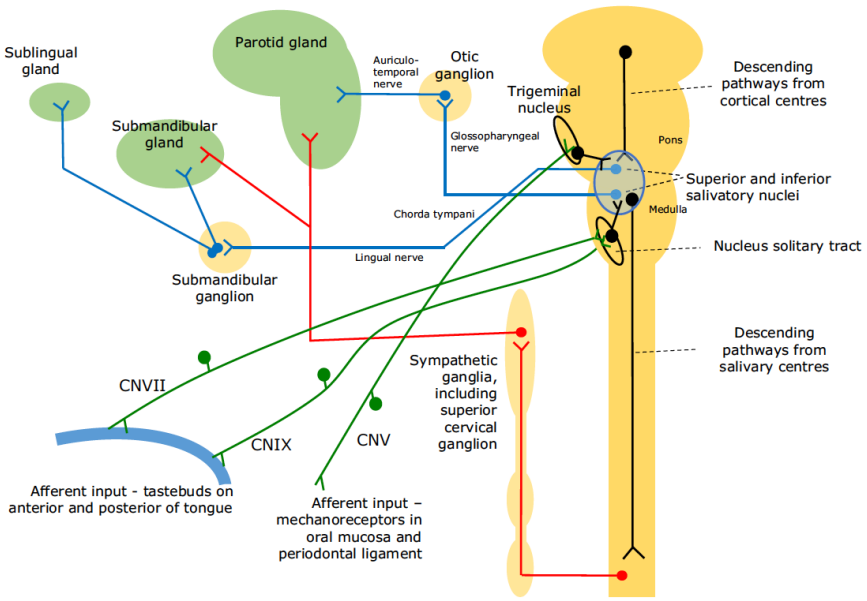



**Figure 1.2 Salivary reflex and the specific neural control of the secretion from the major salivary glands (used with permission from [44]).**

The cells from the salivary glands are closely connected with the autonomic nervous system [44,65]. Parasympathetic and sympathetic nerves contact different cell types in the salivary glands, involving acinar, ductal, myoepithelial cells and blood vessels [44]. There are important differences in the amount of innervation of the salivary glands by sympathetic nerves, with an extensive sympathetic innervation of PGs and SMGs, and little adrenergic innervation of SLGs and minor salivary glands [44]. In general, sympathetic nerve stimulation promotes a protein-rich glandular secretion, while parasympathetic nerve stimulation produces the secretion of large volumes of saliva [44].

The secretions from major and minor salivary glands can be activated differently in response to stimulation [44]. The interaction with receptors from the taste buds, present mostly on the tongue, and the activation of mechanoreceptors in the periodontal ligament and mucosae, activate secretion from the major salivary glands [44]. On the other hand, movement and tactile stimulation of the mucosa are more effective in the secretion from labial and palatine minor glands, although minor salivary gland secretion can also be increased due to taste stimulation [42,44]. Different food smells can increase secretion from SMGs and SLGs, but not from PGs [44,66].

### 1.1.3 Salivary Flow Rate

Saliva secretion is continuous during the whole day, with a total volume of secreted WS around 500 to 600 ml/day on average in a healthy person [4]. Salivary flow rates are lower during sleep and virtually absent during anesthesia [4,44]. Very high saliva secretion is observed during eating and drinking [4,44]. Even in the absence of any exogenous stimulus (unstimulated saliva), the mean flow rate of saliva entering the mouth is about 0.3 to 0.4 ml/min [4,44,67,68], and it is influenced by the circadian cycle, with a lower flow around 6:00 AM and a higher flow around 6:00 PM [4,44,69]. Circannual rhythms are also observed, with higher secretion rates at lower ambient temperatures [70,71]. Despite large variation among individuals [4], an average maximum flow rate of around 7 ml/min can be observed when WS secretion is stimulated [72]. Each salivary gland displays its flow rate (Table 1.1), which changes differently in response to stimulation. The mean unstimulated PG flow rate in healthy individuals is around 0.04 ml/min, but it can increase to about 1.0 to 1.5 ml/min in response to different stimuli [4]. A significant increase in the PG stimulated flow is also observed in response to increased water intake [73]. The unstimulated flow rate of the SMG in healthy individuals is around 0.1 ml/min, and it increases to about 0.8 ml/min under stimulation [4]. The flow rate of the SLGs remains about the same regardless of stimulation [52].

The rate at which saliva enters the mouth varies between individuals [2], and the properties and composition of WS differ based on the characteristics related to the stimulation, including the presence (stimulated/unstimulated) [74], the type (e.g., food tasting [56,72,75], diet [76], smelling [34], or chewing [32,72]), and the duration and intensity of stimulation [28,31,77,78]. Effects of the intensity of gustatory stimulation (sour lemon) and duration of stimulation on the organic and inorganic composition of the secretion from the major salivary glands (PG and SMG) have been demonstrated [31,77,78]. The effect of the intensity of masticatory stimulation (rubber bands or number of strokes per minute) on the PG and WS flow rates have also been demonstrated [79,80]. Additionally, effects of the nature of stimulation on the glandular and WS flow rates in response to masticatory (rubber bands) [79] and gustatory [81] stimuli versus pilocarpine have also been demonstrated [79,81].

Since the secretions from different salivary glands present specific compositions [47,62,63,82,83], variations in the volume secreted by each gland type due to changes in the

stimulation can impact the final WS composition [84]. Although secretions from SMG, PG, minor salivary glands and SLG, represent approximately 65, 20, 10, and 5%, respectively of the total volume of the unstimulated WS, the contribution of the PG increases to about 50% of the total WS volume during stimulated flow [5,48]. Contrarily, the contribution of the SMG to WS decreases to 35% with stimulation, and the secretions from the SLG and minor glands remain stable regardless of stimulation [43]. Because WS is more appropriate than the secretion from salivary glands in the search for biomarkers due to its easy collection, it is imperative to be attentive to possible influences of the stimuli on the final WS composition, and, consequently, on the abundance of some salivary proteins specific to certain salivary glands.

Moreover, differences in the unstimulated WS flow rate were also observed based on sex. Unstimulated WS flow rates in females were approximately 70% of the flow rates observed in males [44,85] possibly due to gland size, which is another important factor related to differences in gland secretions [86]. The sensation of cold temperature, bitter taste, or cooling agents (such as menthol) in the mouth can also increase salivation [44,75,87,88]. Other factors that may affect a person's flow rate include psychological stress [89], intense physical activity [90,91], environmental temperature [71], hydration level [73,92–94], diseases [51], and medications [95–98]. The effects of age on the salivary flow rate are still debatable, with some evidence suggesting that the flow rate decreases in older age [99,100]. This observation is challenged based on the probable influence of medications often used by older people [101,102]. Anticholinergic muscarinic (M3) receptor blockers used to treat irritable bladders (urinary incontinence) and some antidepressants are among the medications that can reduce the secretion of saliva [102]. The most severe forms of hyposalivation are observed in patients subjected to irradiation of the head and neck to treat squamous cell carcinoma [41], and in patients with Sjogren's Syndrome [51], an autoimmune disease that destroys the salivary gland secretory cells, causing chronic inflammation and the interruption of normal secretory signalling in salivary glands. In such populations, the reduction in the salivary flow rate itself can be considered a biomarker for the severity of the condition [44].

### 1.1.4 Changes in Salivary Proteins Before and After Secretion into the Oral Cavity

The proteins found in the WS (WS proteome) are greatly vulnerable to several physiological and biochemical processes. In addition to the dependence on the stimulus duration, intensity, and nature for its resulting volume and composition, the salivary proteome is affected by the many modifications that salivary proteins suffer from their biosynthesis until they reach the oral environment [103]. The biosynthesis of salivary proteins begins with the transcription and translation of their genes in the salivary glands, followed by posttranslational modifications (PTMs) before their secretion into the ductal system, as well as during their transit in the ductal system, and after their release into the oral cavity, including acylation, deamination, sulfation, glycosylation, phosphorylation, and proteolysis; with glycosylation, phosphorylation and proteolysis being the most common PTMs [2,103].

Glycosylation is the most common form of PTM in salivary proteins and it represents the covalent attachment of glycans (also known as carbohydrates, saccharides, or sugars) to a protein. Glycosylated salivary proteins include mucins (MG1 and MG2), agglutinin, glycosylated proline-rich proteins (PRPs), secretory immunoglobulin A, and glycosylated salivary alpha-amylase [103]. In general, glycosylated proteins contribute to the lubrication of oral surfaces and the clearance of microorganisms from the oral cavity [2,103].

Phosphorylation is a PTM of proteins where a phosphate group is covalently bound to a protein's amino acid residue. Examples of phosphorylated salivary proteins are acidic PRPs, statherin, histatin 1, and some cystatin family members (cystatins S and SA-III) [2,103]. Phosphorylation adds a negative charge to proteins promoting the affinity of the phosphoproteins for the enamel surface of the tooth, explaining the participation of phosphorylated proteins in mineral homeostasis [103–106] and the initial formation of a thin protein layer on the enamel surface called acquired enamel pellicle (AEP) [2].

Proteolysis is the breakdown of proteins into smaller peptides or amino acids. Proteolysis can be complete, where no trace of the original protein is found, or partial, providing a mixture of intact and cleaved products [103,107]. The cleavage of the original product often allows the fragments to retain their original activity [103,108], contribute to a functional

improvement in the activity of the resulting peptide [103,109,110], or give birth to different fragments with unique functions [103,108,111]. Proteolysis of salivary proteins can happen intracellular, in the ductal system, or after entering the oral cavity. Examples of intracellular proteolytic modifications of salivary proteins include the hypothesized fragmentation of histatin 3 into histatin 5 inside protease-containing vesicles [22,103,112]. The absence of the intact primary translation products of basic PRPs in salivary secretions is another indication of the intense proteolysis that occurs before these proteins reach the oral cavity [103].

After reaching the oral cavity, the non-sterile oral environment exposes proteins to several proteases, glycosidases, and transferases [103], including those secreted by the host and by microorganisms, causing salivary proteins to undergo several changes including severe proteolytic cleavage, partial deglycosylation, and the formation of protein complexes, all while in a dynamic environment with a continuous turnover of newly secreted proteins and constant removal by swallowing [103]. It is understood that the proteomes of the secretion collected from the duct of distinct salivary glands are different from each other and that they also differ from that of the WS [5,42]. There is also acceptable knowledge about the modifications that happen to most salivary proteins from their biosynthesis in the cells until they leave the glandular duct [103]. However, the knowledge about the modifications of proteins in the dynamic oral environment, such as the formation of protein-protein complexes, for example, is still limited. Additionally, the higher complexity of WS comes with a price associated with the lower stability of some of its constituents, a fact that still hinders the full application of salivary proteins in WS for biomarker identification [103].

### 1.1.5 Functions of Whole Saliva

The WS is a biological fluid with extremely important roles in the maintenance of the integrity of hard and soft tissues in the oral cavity [44,113–115]. Many of such roles depend on the interaction of saliva with the different surfaces found in the oral cavity [44,113–115]. Examples include the soft epithelial tissue surfaces with different degrees of keratinization and roughness and the mineralized hard tooth surfaces. These two unique surfaces interact with specific salivary components due to their distinct textures, composition, and polarity [44]. In contact with the soft tissues, saliva assists in the hydration of the oral mucosa [44] and promotes the adhesion of specific proteins in the formation of a thin layer, the salivary

mucosal pellicle (SMP) [44,116–118]. The SMP is a viscoelastic gel matrix that protects the oral mucosa from mechanical and chemical damage, and the entry of microorganisms and toxic materials [118]. Large highly glycosylated proteins termed mucins are the main component of the SMP [118]. The thin protein layer formed over the hard surface of the tooth (enamel) is called AEP [119]. The AEP offers a lubricating layer that reduces wear and attrition of surfaces, and reduces acid-induced enamel demineralization, suggesting that individuals with thicker AEP might be less susceptible to enamel loss [44]. Tooth brushing and tooth polishing can remove the AEP completely; however, the AEP is reconstructed within minutes of tooth exposure to saliva in response to the strong interaction with calcium-binding proteins, phosphate-binding proteins, and PPIs, as observed with MUC5B, salivary alpha-amylase, histatin and statherin [44,119–122]. Although proteins with affinity for calcium and phosphate are more abundant in the early stages of AEP formation, this profile changes over time with an increase in the participation of proteins that interact with other proteins in the later stages of AEP development [121]. The AEP serves as the base for the selective development of dental plaque (biofilm). Therefore, modulating the AEP could be an interesting avenue for the prevention of dental caries [121].

Moreover, saliva participates in the taste sensation, and facilitates chewing, swallowing, and speaking by lubricating all oral surfaces and facilitating food bolus formation [123]. In addition to initiating food digestion in the mouth, salivary proteins assist in the clearance of microorganisms by swallowing and display antibacterial, antifungal, and antiviral properties [123,124]. Furthermore, saliva protects the tooth surface from demineralization [123,125], and participates in the process of tooth remineralization [123,124], while dynamically maintaining the salivary pH via different buffering systems (bicarbonate, phosphate, and protein buffer systems) [123,124]. Wound healing capacity is also attributed to a family of salivary proteins (histatins) [126–131]. Most of the functions credited to the human saliva are executed by salivary proteins or with their assistance [123,124,132] (Figure 1.3).

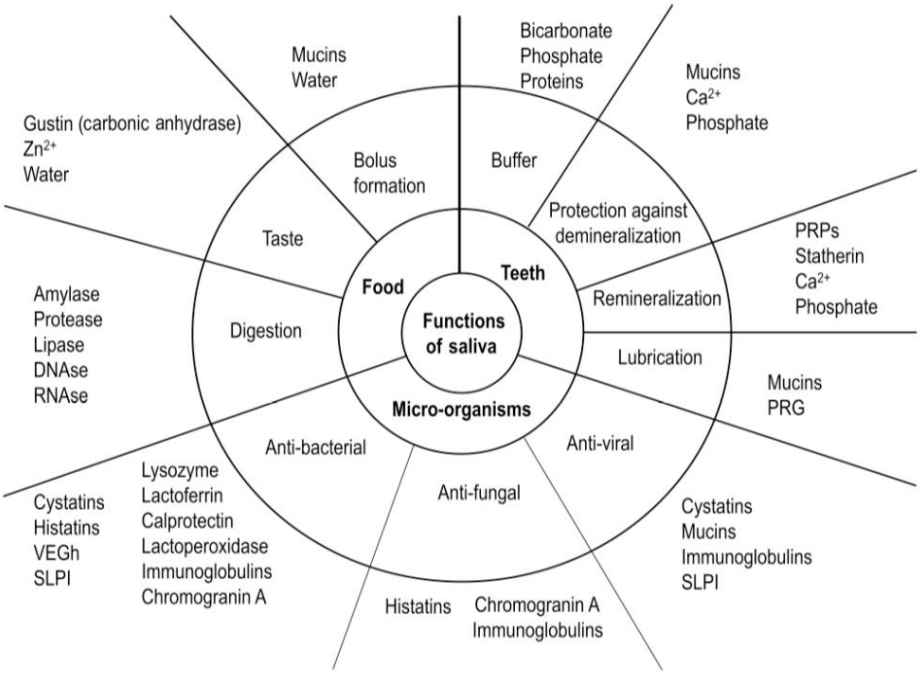



**Figure 1.3 Schematic representation of some salivary components in relation to the main functions of saliva (used with permission from [123]).**

## 1.2 Salivaomics

Currently, saliva can successfully be used for the detection of non-host-derived elements like drugs [133], different viruses (human immunodeficiency virus, hepatitis C, human papilloma virus [134] and the novel coronavirus covid-19 [135]), and genetic analysis [136]. Additionally, many host-derived disease biomarkers have also been studied in saliva. For example, many salivary constituents have been investigated for cancer diagnostics including circulating tumor DNA (ctDNA) [137–142], which have been found in blood, urine, and saliva [143,144]; microRNA biomarker for oral cancer [145]; piwi-interacting RNAs [146,147]; salivary proteins related to oral cancer [12,148–150], lung cancer [151–153], and breast cancer [154,155]; and, interestingly, a panel of exRNA biomarkers for oral cancer detection which is present only in saliva (IL-1β, OAZ1, SAT, and IL-8) [156]. The large number of components that can be assessed in saliva for one single disease, such as cancers[12], is descriptive of the complexity and richness of this biological fluid and emphasizes the need for an integrative study of saliva and its constituents. The study of all the different salivary elements is called Salivaomics [11]. The different types of salivary components with the potential to be used

for diagnosis come from various sources and they include electrolytes and ions (Na+, Cl-, Ca2+), cells and particles (epithelial cells, neutrophils, microorganisms, microparticles, and exosomes), lipids (triglycerides and cholesterol), steroid hormones (estrogen, testosterone, and cortisol), nucleic acid containing molecules (DNA, mRNA, noncoding RNA, microRNA) and many proteins and peptides (mucin glycoproteins, statherin, proline-rich proteins, carbonic anhydrase 6, histatins, secretory IgA, IgG, albumin, lysozyme, lactoferrin, matrix-metalloproteinase-8, interleukin 8, nerve growth factor, leptin, LL37, alpha-defensin) [44].

### 1.2.1 Salivary Proteomics

The Human Genome Project [16,157,158] was created to explain and map the diseases that affect humans. However, following the conclusion of this project, many questions remained suggesting that a better understanding of the Human Proteome was necessary [16,159]. Since the proteome derives from the genome, all proteomics efforts aim to complete the genomic information [15–17]. With that goal, the presence and/or abundance of certain proteins, due to physiological and pathological conditions, have been studied extensively, with protein expression in a biological sample varying from none to abundant in certain conditions [17]. As stated in the report [17] from the "Defining the Mandate of Proteomics in the Post-Genomics Era" symposium, held at the National Academy of Sciences on February 25, 2002, Proteomics represents "the effort to establish the identities, quantities, structures, and biochemical and cellular functions of all proteins in an organism, organ, or organelle, and how these properties vary in space, time, or physiological state". Based on the above definition, proteomics investigations usually start with a broad exploration of complex mixtures, followed by quantitative analysis of the abundance of certain proteins, the determination of the molecular structure of a specific protein/peptide or protein complex, and the investigation of the biochemical characteristics and biological functions of a specific protein, peptide or complex.

A work by Rosa et al. (2012) [10] aimed at creating a comprehensive database about the salivary proteome by providing a reviewed compilation of all the proteins identified by proteomics studies in the oral cavity. They called this comprehensive database "OralOme". A total of 53 studies were included in this review with healthy subjects and subjects with several oral and systemic conditions. Their results demonstrated that 3397 non-redundant

proteins could be found in the oral cavity (Figure 1.4). Of these, 3115 proteins were from saliva, 990 from the oral mucosa, and 1929 were proteins from the plasma. They separated the proteins identified in subjects with oral and systemic conditions from the "normal proteome". Overall, 707 proteins were identified in patients with different oral and systemic conditions, 637 of them in saliva samples (73 in PG, 431 in minor salivary glands, 74 in GCF, and 228 in WS), and 118 in the oral mucosa. Fifty-one proteins were exclusively identified in diseases. There were 1193 proteins found in the secretion from the PG and 999 proteins in the secretion from the SMG/SLG, with 701 proteins common in both secretions. The salivary minor gland contributed 554 proteins and the GCF with 100 proteins, with a total of 2206 proteins found in WS. Many proteins were unique to one of the sources, with 1283 proteins (40%) found only in saliva, being absent from plasma and oral mucosa samples. The highest percentage of proteins identified in only one source came from the PG samples (15.3%), followed by SMG/SLG (8.5%), minor salivary glands (5.4%), and GCF (8%). In summary, this study documented the diversity of the proteome found in the oral cavity, reported its various sources, and reinforced the diagnostic capabilities of the salivary proteome due to its unique composition.

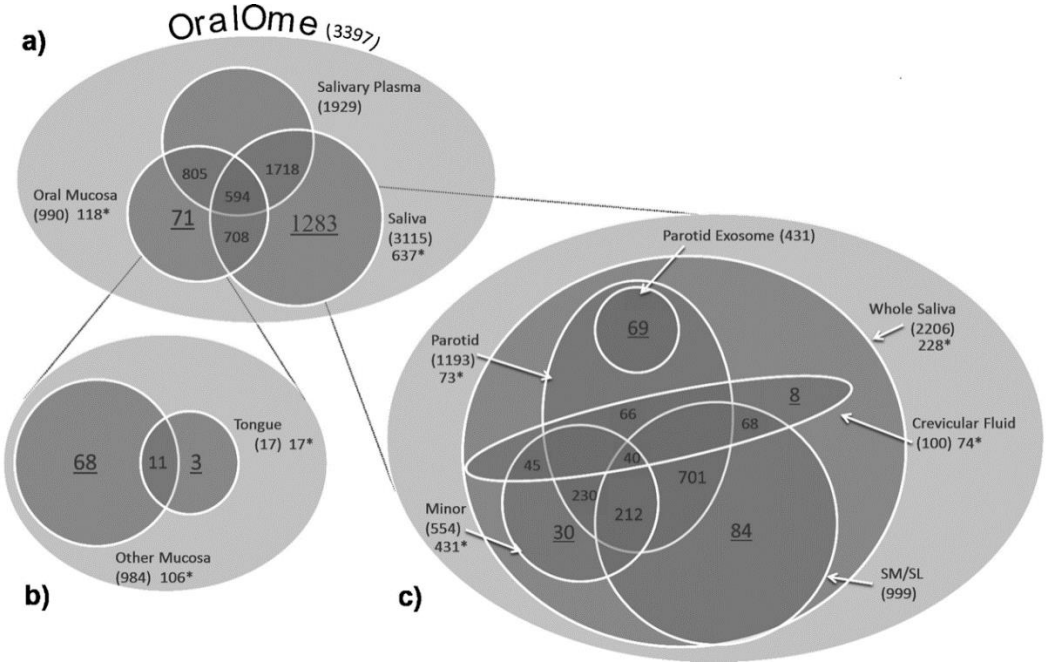



**Figure 1.4 Illustration of the Human OralOme: (a) OralOme main protein sources; (b) oral mucosa protein sources; (c) saliva protein sources. Parenthesis marks the number of proteins from a specific source. Underline marks the number of proteins**

**unique to one source. Asterisk marks the number of proteins obtained from patients with oral or systemic diseases. (Used with permission from [10])**

Despite its diversity, a significant portion of the total protein concentration in saliva is formed by a group of major protein families (proteins with closely related structures) that include acidic PRPs, basic PRPs, amylase, high-molecular-weight glycoprotein MUC5B, low-molecular-weight MUC7, agglutinin, cystatins, histatins, and statherin [2,103]. The 20 most abundant proteins found in WS represent about 40% of the salivary proteome [2,6]. Many other proteins in saliva have been detected in trace amounts [2].

In a review paper, Ruhl (2012) [160] stated that the exploration of the salivary proteome was "essentially completed" at that time. The paper listed important advancements made in the field of salivary proteomics, with the identification of many diagnostic markers among salivary proteins and peptides. However, most of the disease markers identified until that time were serum components and not intrinsic proteins produced by the salivary glands. Ruhl's work also indicated promising future applicability of studies about PTMs of proteins (glycoproteome and phosphoproteome) and PPIs (interactome) in the diagnosis and prevention of oral diseases. Eight years later, Ruhl and collaborators (2020)[63] continued their exploration into the salivary proteome to determine the precise origins of the proteins intrinsic to human saliva. Their results identified hundreds of transcripts specifically expressed in salivary glands, reinforcing saliva's unique capabilities for diagnostic purposes and the distinctive importance of the secretion from each salivary gland.

### 1.2.2 Salivary Interactomics

Although some proteins may act primarily as single monomeric units, a significant percentage of them, if not the majority, act in complexes with partners to execute their functions [2,161]. The fact that the original individual functions of proteins involved in the formation of protein complexes can be increased, reduced, or altered because of the PPIs, opens an entirely new avenue of functional studies [2,21,25,162]. Moreover, many processes related to health and disease are dependent on PPIs [161,163,164]. Therefore, understanding how salivary proteins interact with each other is just as important as documenting the

biological functions of individual salivary proteins. The study of PPIs in saliva is called Salivary Interactomics.

An example of the functional changes that can be observed when salivary proteins interact in the formation of complexes is the increased resistance of histatin 1 to proteolytic degradation in the harsh WS environment when participating in a protein complex with salivary amylase, as demonstrated by Siqueira and collaborators (2012) [165]. By interfering with the stability and availability of certain salivary proteins, PPIs might affect the composition of the salivary proteome and the proper evaluation of potential disease biomarkers, making the exploration of the salivary interactome of great importance for the use of saliva as a diagnostic tool [2]. Moreover, this same interaction between histatin 1 and amylase suggests a new biological function for salivary amylase as a potential carrier for proteins and peptides throughout the oral cavity; thus, providing information for the development of effective systems for the delivery of therapeutic peptides and proteins inspired by the histatin 1-amylase complex. Mapping the Human Interactome is also extremely important in the success of many drugs since the activation/deactivation of signalling cascades and the transport of proteins and other metabolites in the body rely on PPIs [161].

Currently, more than 70,000 human proteins have been identified, but less than 30,000 of them have had their interactome investigated, with almost 350,000 interactions documented [163,166,167]. Different databases have been created to assist and accelerate the identification and characterization of PPIs and their biological functions [162,163,168–177]. Such web tools allow public access to integrated known and predicted PPIs enabling fast integration of novel, current, and predicted knowledge [25,162].

## 1.3 Methods Used to Explore the Salivary Proteome and Interactome

### 1.3.1 Saliva Collection Techniques

Generally, saliva can be collected from specific salivary glands (major and minor salivary glands) or by collecting the fluid found in the mouth (WS). Saliva can be secreted with stimulation (stimulated saliva) or without stimulation (unstimulated or resting saliva). The volume of saliva sampled usually depends on the analytical procedures to be used. An

alternative to a fixed volume is the pre-determination of a fixed duration of collection, allowing a variation in the final volume collected. A fixed intensity of stimulation can also be used via maintenance of continuous secretion rates. All the factors mentioned above can influence the final proteome [178–180].

#### 1.3.1.1 Whole Saliva Collection

WS can be collected without stimulation (unstimulated or resting saliva) or with stimulation (stimulated WS). WS is usually collected using a graded Falcon tube on ice by passive drooling and draining method (unstimulated), or by spitting (stimulated). Swab-based sampling can also be used; however, the choice of absorbent material should consider the subject tolerability and the capability to preserve the analytes of interest [26,181]. Also, the amount of sample collected using swabs or other absorbent materials is limited. Suction with a pump is an alternative for WS collection, and it is usually the choice for some patients with special needs (infants, young children, patients with cerebral palsy or syndromes that interfere with their motor or mental capacity, and elders with neurological or motor impairment) [26]. The work presented in Chapter 3 of this dissertation was performed with WS collected with masticatory stimulation by spitting into a graded Falcon tube [25] (Figure 1.5).

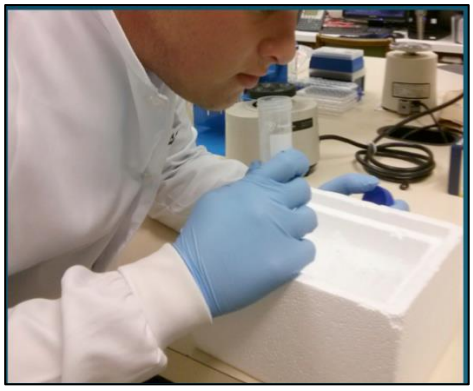


**Figure 1.5 Collection of WS by spitting into a graded Falcon tube.**

Physiologic stimulated WS is secreted in response to masticatory and gustatory stimulations during food intake. For clinical routine, stimulated WS can be achieved by mastication on inert materials such as parafilm, unflavored chewing gum, and rubber bands, since the stimulant does not mix with the composition of the WS [181]. Gustatory

stimulation of WS secretion can be obtained using citric acid, sour candy, or other tastants of interest [181]. Stimulation increases the WS flow rate, protein output, and salivary pH [181]. Both mastication and gustatory stimulation have an impact on the composition of the WS [26], highlighting the importance of standardization of the sampling procedures to make saliva analysis more robust, reproducible, and accurate.

#### 1.3.1.2 Glandular Saliva Collection

For the collection of the secretion from salivary glands, gustatory stimulation is usually preferred to masticatory stimulation to avoid the displacement of the devices. When saliva is collected as the secretion from a particular salivary gland, specific devices are used [42,181].

The secretion from the PGs can be collected by intraductal cannulation, or with the assistance of a Lashley cup device [182], also known as the Carlson-Crittenden collector. The intraductal cannulation consists of inserting a polyethylene tube or a tapered sialography cannula into the gland. This approach is slow and invasive and not suitable for routine clinical use [181]. A non-invasive alternative for parotid saliva collection consists of using a Lashley cup device (Figure 1.6).

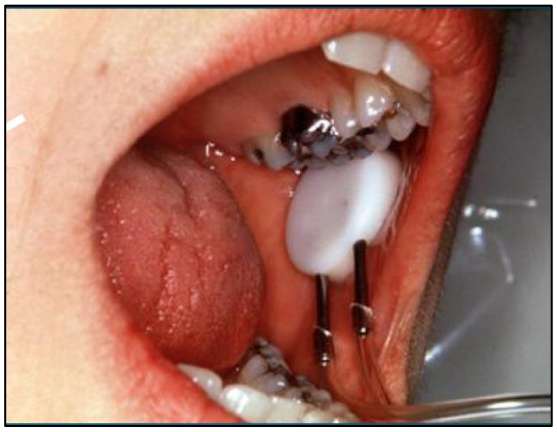


**Figure 1.6 Collection of the secretion from the left PG using a Lashley cup device.**

The Lashley cup consists of two concentric chambers communicating with the exterior by two metal cannulas. The inner chamber is placed around the opening of Stensen's duct, and a vacuum is created in the outer chamber by a suction pump. A tube connected to the inner chamber allows the exit of the free-flowing parotid saliva into graded tubes placed on ice [181]. The work presented in Chapter 4 of this dissertation used stimulated PG saliva,

collected using a Lashley cup device, via gustatory stimulation, with continuous flow rates, which resulted in a fixed duration of collection for all samples, and fixed volumes within each tested flow rate.

To assure the separate collection from the SMGs or the SLGs, duct cannulation can be used. However, SMG/SLG duct cannulation is an invasive and complex procedure, which can result in gland injury, and that is not suitable for routine clinical use [181]. Alternatively, gentle suction of the saliva accumulated in the floor of the mouth, or the use of custom-made devices placed at the openings of the Wharton's and Bartholin's ducts, such as Pickerill's device, Block-Brottman collection devices, polymethylmethacrylate devices, and the Wolff saliva collector, are acceptable approaches [181].

Some alternatives to collecting secretion from minor salivary glands include the use of capillary tubes, paper strip filters and a Periotron, and a microliter plastic pipette tip connected to a P10 Gilson pipette [59,181]. However, the clinical application of the secretions from the minor salivary glands is limited by the arduous collection procedure and the small amount of secretion collected, which is often not enough for a proper chemical analysis [181].

### 1.3.2 Protein Purification and Separation

The purification strategies applied to samples before mass spectrometry (MS)-based investigation are important factors in the final sensitivity of this technique [183]. Classical separation methods can be initially used such as centrifugation, column chromatography, and affinity-based procedures [183]. Normally, protein purification starts with whole-cell lysate (for tissue samples), or saliva collection (for salivary research), and ends with a gel-separated protein band or spot. Polyacrylamide gel electrophoresis (PAGE) can be used to separate native [184] or denatured proteins [53]. Examples of gel electrophoresis strategies include one-dimension (1-D) and two-dimension (2-D) gels. In the 2-D gels, the proteins are separated based on their isoelectric point (pI) and their size (molecular mass) simultaneously [53,185]. Proteins in the 2-D PAGE must be denatured and solubilized, which is usually achieved with high concentrations of urea, thiourea, and reducing agents like dithiothreitol (DTT) [13,53]. In the 1-D gels, on the other hand, proteins are separated according to only one chemical characteristic, their molecular mobility [186]. Examples of 1-

D gels include the sodium dodecyl sulfate (SDS) PAGE, the native PAGE, and the cationic PAGE. Coomassie staining, silver staining, or radioactive labelling can be used to visualize the proteins in gels [183]. For the work presented in Chapter 3 of this dissertation, centrifugation was initially used to separate the WS supernatant (WSS), which was further separated using an in-house affinity chromatography (AC) column, followed by 1-D gel electrophoresis (SDS-PAGE and native PAGE) (Figure 1.7). After tryptic in-gel or in-solution digestion, all samples were further purified using reversed-phase chromatography for the removal of salts and detergents, before MS analysis [25].

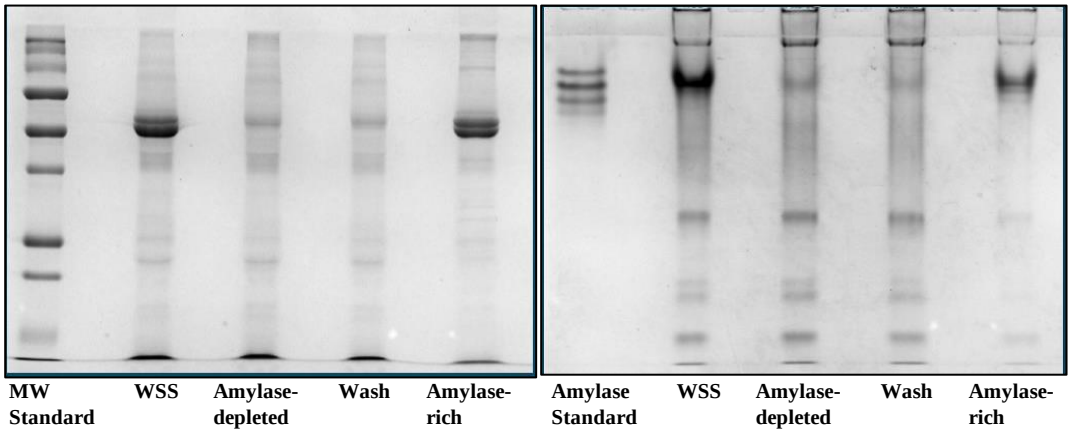



**Figure 1.7 (A)** **Figure 1.7 (B)**

**Figure 1.7 Examples of PAGE separation methods: (A) SDS-PAGE, and (B) native PAGE. The first column to the left of the gels represents the protein standard to which the samples loaded in the subsequent columns are compared to for identification.**

In the SDS-PAGE, a detergent bind to the proteins conferring a negative charge to them [186]. The proteins are loaded into a polyacrylamide gel and allowed to run toward a positive charge. Smaller proteins will run faster to the bottom of the gel, while larger proteins will run slower and remain in the top part of the gel. A protein standard is included in every run for molecular weight reference. Contrarily, the proteins are separated respecting their natural charge in the native PAGE [184] and cationic PAGE [187], and a sample of the protein(s) of interest is used as standard. MS analysis is usually applied on peptides recovered from gel-separated proteins after enzymatic degradation. Alternatively, the gel electrophoresis

step can be skipped. Regardless of the selected preliminary steps, the final sample must be free of detergents and salts before MS analysis [183].

### 1.3.3 Mass Spectrometry-based Proteomics

MS is a powerful tool in protein analysis and the key technology in the field of proteomics [183]. MS combines high sensitivity with high specificity to allow the identification of human proteins directly from genome databases based on minimal sample amount [183]. Additionally, PPIs can also be analyzed by precipitation of a tagged bait followed by MS identification of its binding partners. As a result, entire protein complexes, signaling pathways, and whole organelles can be characterized [183]. Furthermore, new MS-based approaches are being used to further explore protein PTMs, with important advances in the fields of phosphoproteomics [152,188,189], glycoproteomics [189,190], and degradome/peptidomics [119,189].

#### 1.3.3.1 MS-based Strategies

There are two main strategies used in MS-based proteomics: top-down proteomics and bottom-up proteomics. All work reported in this dissertation was acquired using bottom-up proteomics.

When proteins are submitted to enzymatic degradation (tryptic digestion) before MS analysis, the strategy is called bottom-up proteomics. Bottom-up proteomics is the most used approach. In bottom-up "shotgun" proteomics, a complex protein mixture is digested into fragments called peptides (tryptic digestion). The peptide mixture is often fractionated and analyzed by tandem MS (MS/MS). Peptide identification of tandem MS happens by matching mass measurements of the precursor peptide and the MS/MS fragment ions with theoretical sequences from the genome. Bioinformatics tools assist in the interpretation of the data [188,189,191]. The proteins are identified by cross-referencing the mass-to-charge ratio (m/z) of the fragments measured by MS/MS to a computed proteomic database. This approach is usually preferred when analyzing complex mixtures as it allows the identification of many proteins at once. However, since this approach is not targeted, proteins that are more abundant in the mixture tend to be identified more often than low-abundance proteins [189]. To reduce this technical limitation, gel separation techniques are often used before MS/MS analysis (Figure 1.8).

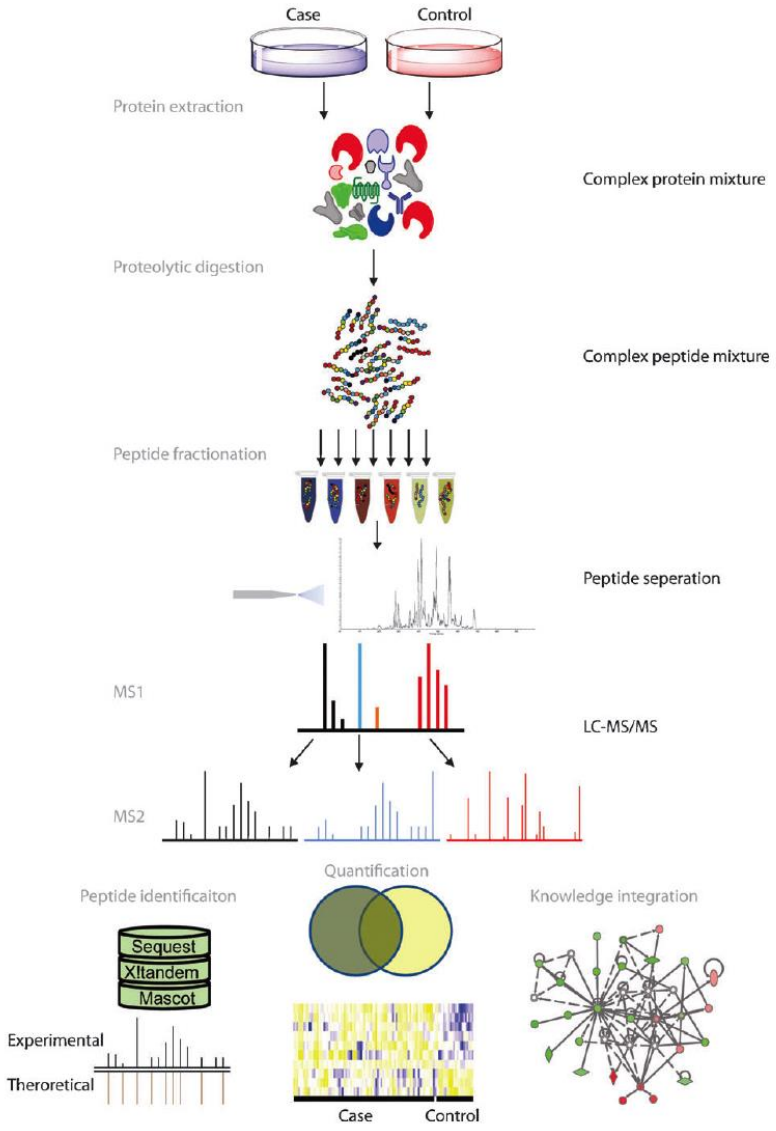



**Figure 1.8 Schematic representation of MS-based bottom-up "shotgun" proteomics strategy. (Used with permission from [189])**

The MS-based strategy where intact proteins (without proteolytic digestion) are analyzed is termed top-down proteomics. Top-down proteomics is usually employed in studies looking at characteristics of the proteins present in the sample in their natural state (PPIs, protein structure, naturally occurring peptides, or complex formation). This approach is more commonly used in samples with less complexity. The intact proteins are fragmented in gas-phase [188,189,191].

#### 1.3.3.2 Ionization Techniques

To be submitted to the mass analyzer, the biomolecules need to be transferred into the gas phase and ionized. The two main ionization techniques used in the life sciences are electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI). All MS data used in this dissertation were acquired through ESI.

ESIMS was developed for biological MS [183]. In this technique, the liquid containing the analyte is pumped at low microliter-per-minute flow rates through a hypodermic needle at high voltage to electrostatically disperse (electrospray) small micrometer sized droplets, which quickly evaporate and provide charge onto the analyte molecules [189]. This ionization process is gentle (soft ionization), it does not fragment the analyte ions in the gas phase, and the charged molecules are transferred into the mass spectrometer with high efficiency. A wide range of compounds can be analyzed with ESIMS if they are sufficiently polar to allow attachment of a charge, including proteins, oligonucleotides, sugars, and polar lipids [183]. The mass spectrometer measures the m/z rather than the mass of the analyte ion. Large ions are usually multiply charged allowing the analysis of analytes with virtually no upper mass limit by ESIMS [183]. Generally, high-pressure liquid chromatography (HPLC) (also known as high-performance liquid chromatography) is coupled with an ESI source [188]. Higher-efficiency peptide separation can be achieved by using higher liquid chromatography (LC) operating pressure, longer columns, and reducing the particle sizes of the material used in the chromatography column [188,191,192]. HPLC coupled with MS enables the analysis of thousands of proteins per measurement [189]. In the HPLC the peptides are usually separated according to their hydrophobicity and the peptides eluting from the column are directly ionized by ESI before entering the mass spectrometer [188,191].

Matrix-assisted laser desorption ionization (MALDI) is also considered a soft ionization method. In this technique, the molecules to be analyzed are coprecipitated with a large excess of matrix material by pipetting a sub microliter volume of the mixture onto a metal substrate and allowing it to dry. The solid matrix, which is typically a small organic molecule, is irradiated by nanosecond laser pulses generating gas-phase protonated molecules [183]. MALDI is mostly applied to the analysis of peptides, since proteins generally undergo some fragmentation with this technique, resulting in broad peaks and loss in sensitivity [183].

#### 1.3.3.3 Mass Analyzers

Mass spectrometers usually include an ion source and optics, a mass analyzer, and data processing electronics. Mass analyzers can store ions and separate them based on the m/z of proteins, peptides, and peptide fragments [183]. Different types of mass analyzers present specific properties related to mass range, analysis speed, resolution, sensitivity, ion transmission, dynamic range, and applications [188]. The mass spectrometers mostly used in proteomics include ion traps (IT), like quadrupole IT, linear IT, and IT-time of flight; and triple quadrupoles (TQ), like linear TQ-Orbitrap, linear TQ-Fourier transform ion cyclotron resonance, and TQ-Fourier transform ion cyclotron resonance [188].

The mass measurement accuracy power of mass spectrometers is increased due to their ability to perform MS/MS measurements, which provide additional information specific to the peptide amino acid sequence [188,189]. Typically, LC-MS/MS relies on the acquisition of a preliminary mass spectrum (MS1) of the intact (precursor) peptide, the dissociation of the isolated precursor ion of interest into smaller fragments, and subsequent mass analysis of the fragment (MS2) [188,189]. This process is repeated until all the peptide mixture is separated in the LC column. Peptide fragmentation is usually achieved by collision-induced dissociation (CID) [189].

All work presented in this thesis was done with the use of an LTQ-Velos mass spectrometer (Thermo Scientific, San Jose, CA, USA). The LTQ instruments are formed by three parts: two-mass separating quadrupole sections separated by a central quadrupole section to contain the ions during fragmentation [183]. The quadrupole is a mass filter, containing four rods to which an oscillating electric field is applied to allow only a certain mass to pass through. The scan of the amplitude of the electric field and the ions that reach the detector provide a mass spectrum [183].

More recently, hybrid instruments like the LTQ-Orbitrap (Thermo Scientific, San Jose, CA, USA), which combines a dual-pressure LTQ ion trap mass analyzer with an Orbitrap, have allowed the simultaneous identification of low-level analytes in complex samples with accurate mass and ultra-high resolution [183,193]. The LTQ-Orbitrap hybrid instrument combines the speed and sensitivity of the LTQ with the high resolution and mass accuracy of the Orbitrap. In such instrument, the LTQ is used to control the number of ions going

into the Orbitrap, and to perform CID to fragment the ions. The Orbitrap relies on the orbital trapping of ions in its static electrostatic fields causing the ions to orbit around a central electrode and oscillate in an axial direction [188].

### 1.3.4 Bioinformatics Tools for Protein Identification and Analysis

Peptide identification using MS/MS normally uses genomics data by matching mass measurements for intact peptides and MS/MS fragment ions to theoretical sequences derived from genome sequence data [189]. The database matching strategies are usually done with the assistance of bioinformatics tools such as Mascot, Sequest, and Xltandem [189], and the MS/MS spectra are matched with human protein databases (Swiss-Prot and TrEMBL) [5,25]. All protein identification done in this dissertation was processed with the use of Sequest.

In addition to protein/peptide identification, a relative quantification of the abundance of the identified proteins/peptides can be achieved from LC-MS/MS [189], which is often estimated by counting the number of times a peptide mass spectrum is measured and identified (spectral counting) [189]. Determination of the biological significance of the qualitative and quantitative data acquired with LC-MS/MS is usually achieved with the assistance of publicly available information about the protein's tissue origin, functionality, and role in biochemical processes [189]. Due to the immense volume of information associated with proteomics measurements, different bioinformatics tools can be used to provide accurate and up-to-date information and functional enrichment of protein lists in a practical and time-efficient manner. Examples of some key bioinformatics tools and databases used in the studies included in this thesis are STRING Protein-Protein Interaction Networks Functional Enrichment Analysis (https://string-db.org/) [176], PANTHER Database (https://www.panther.org) [173,194], DAVID Bioinformatics Database (https://david.ncifcrf.gov/) [170], Bgee Gene Expression Database (https://bgee.org/) [195], The Human Protein Atlas (HPA) (https://www.proteinatlas.org/) [196], and The Human Salivary Proteome Wiki (HSPW) (https://www.salivaryproteome.org/) [197]. Initiatives such as the HSPW aim at reducing the gap in knowledge by compiling data from various publications, allowing data review by experts, and making original data publicly available for re-examination [9,197].

### 1.3.5 Protein Interactions and Protein Complexes

Studies about PPIs usually start with a general search for novel binding partners. Techniques used in this initial step include affinity chromatography (AC), pull-down and/or coimmunoprecipitation (Co-IP) followed by MS for the identification of PPIs [19].

The AC technique relies on the specific affinity of the target protein for the substrate present in a column to achieve the enrichment of the sample. After non-binding proteins are rinsed during the intermediate washing step, the target protein is eluted along with the other proteins involved in the complex [19,198]. An in-house starch AC column was used for this procedure in Chapter 3 of this thesis [25].

In the pull-down experiments, a "bait" protein is used to capture the interacting proteins from the complex mixture. In the Co-IP experiments, the "bait" protein, along with its complex partners, is captured from the mixture using a specific antibody. The antibody is immobilized using proteins A or G, which are covalently attached to sepharose beads. The beads are washed to remove proteins not pertaining to the complex, and the antibody and bait protein are eluted together with other proteins associated with the complex [19,198]. Co-IP and pull-down experiments were employed in the initial study about the histatin 1 interactome [21], which was used to demonstrate the use of the STRING database to merge in vitro and in-silico results for the construction of a protein-protein network shown in Chapter 2 of this thesis [162].

The final enriched mixture from AC, pull-down or Co-IP is submitted to MS and/or immunoblotting for protein identification [19]. AC, pull-down and Co-IP have the advantage of not requiring the use of artificial tags [19]. Additionally, in-silico approaches are important computational alternatives to model protein-protein complexes (e.g., Haddock)[199], or to complement the results from in vitro experiments and expand a protein complex network [25,162]. By assessing public PPI databases, researchers can save both time and resources when analyzing extensive data and selecting the most promising protein partners for subsequent experiments. Because protein complexes include both proteins that interact directly with the target protein being investigated (direct interaction), as well as proteins that interact indirectly with the target protein via a common interaction partner [19], it is important to conduct experiments to verify the nature of the interaction. Examples of

strategies that can be used to verify PPIs include confocal microscopy for intracellular localization of proteins, Co-IP, surface plasmon resonance (SPR) and spectroscopic studies [19].

## 1.4 Thesis Objective

The aim of this thesis was to investigate aspects that can affect saliva's final proteomic composition. Two main factors were elected for this investigation: (1) the interaction among salivary proteins, and (2) the effect of intensity and duration of stimulation on the proteomic composition of the secretion from salivary glands. This aim has been addressed through the following:

1. Demonstrate the usefulness of the STRING database for studying PPI in the histatin1-protein network.
2. Reveal the salivary proteins that interact with amylase forming heterotypic complexes in WS.
3. Investigate the effects of intensity and duration of stimulation on the proteome of the secretion from the human PG under continuous flow rates.

Chapters 2 through 4 address these objectives in order.

## 1.5 Thesis Outline

This Chapter provides a background about saliva's composition and secretion, the functions of its components, and the analysis techniques commonly employed in the study of salivary proteins. Chapters 2 and 3 were based on published work. Chapter 4 demonstrates the effect of intensity and duration of the stimulation on the proteome of the secretion from the human PG. Chapter 5 summarizes the advances presented in this work and the perspectives on future use of our discoveries for the advancement of salivary research.

### 1.5.1 Merging in-silico and in vitro salivary protein complex partners using the STRING database (Chapter 2)

An in-silico approach was used to successfully perform a fast simulation of a novel constructed histatin 1 protein-protein network, including both known and predicted

interactors, along with in vitro complex partners identified in a previous publication. This Chapter is based on a publication titled "Merging in-silico and in vitro salivary protein complex partners using the STRING database: A tutorial.", published in the *Journal of Proteomics* in 2018 by Karla Tonelli Bicalho Crosara, Eduardo Buozi Moffa, Yizhi Xiao, and Walter Luiz Siqueira [162].

### 1.5.2 Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches (Chapter 3)

This study identified several salivary proteins that interact with salivary amylase forming heterotypic complexes in saliva, characterized the high-confidence interactors, and constructed a simulated novel amylase-protein hub network with the aid of bioinformatics tools. This chapter is based on a publication titled "Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches" published in *BioMed Research International* in 2018 by Karla Tonelli Bicalho Crosara, David Zuanazzi, Eduardo Buozi Moffa, Yizhi Xiao, Maria Aparecida de Andrade Moreira Machado, and Walter Luiz Siqueira [25].

### 1.5.3 Effects of Intensity and Duration of Stimulation on the Proteome of the Secretion from the Human Parotid Glands (Chapter 4)

This study demonstrated that the proteome of the secretion from the human PG is affected by both intensity and duration of stimulation. Our results indicate the importance of understanding physiologic factors that influence the composition of the secreted saliva for the development of robust saliva collection protocols. Additionally, our analysis suggests that there may also be a protein-specific secretory mechanism affecting the secretion of PG proteins differently in response to the two stimulation factors: intensity and duration.

### 1.5.4 Conclusion and Future Directions (Chapter 5)

This chapter summarizes the main findings of the work presented in this thesis and elaborates on future directions for salivary research. Conclusions and closing remarks are discussed.

## 1.6 References


1. Tabak, L. A. A Revolution in Biomedical Assessment: The Development of Salivary Diagnostics. *J Dent Educ* **65**, (2001).

2. Siqueira, W. L. & Dawes, C. The salivary proteome: Challenges and perspectives. *Proteomics - Clinical Applications* vol. 5 (2011).

3. Edgar, W. M. Saliva: Its secretion, composition and functions. *Br Dent J* **172**, (1992).

4. Edgar, W., O'Mullane, D. & Dawes, C. *Saliva and oral health.* (British Dental Association, 2004).

5. Denny, P. *et al.* The proteomes of human parotid and submandibular/sublingual gland salivas collected as the ductal secretions. *J Proteome Res* **7**, (2008).

6. Loo, J. A., Yan, W., Ramachandran, P. & Wong, D. T. Comparative human salivary and plasma proteomes. *Journal of Dental Research* vol. 89 (2010).

7. Yan, W. *et al.* Systematic comparison of the human saliva and plasma proteomes. *Proteomics Clin Appl* **3**, (2009).

8. Nagler, R. M., Hershkovich, O., Lischinsky, S., Diamond, E. & Reznick, A. Z. Saliva analysis in the clinical setting: Revisiting an underused diagnostic tool. *Journal of Investigative Medicine* **50**, 214–225 (2002).

9. Dawes, C. & Wong, D. T. W. Role of Saliva and Salivary Diagnostics in the Advancement of Oral Health. *J Dent Res* **98**, (2019).

10. Rosa, N. *et al.* From the salivary proteome to the OralOme: Comprehensive molecular oral biology. *Arch Oral Biol* **57**, (2012).

11. Wong, D. T. W. Salivaomics. *Journal of the American Dental Association* **143**, (2012).

12. Aro, K., Kaczor-Urbanowicz, K. & Carreras-Presas, C. M. Salivaomics in oral cancer. *Current Opinion in Otolaryngology and Head and Neck Surgery* vol. 27 (2019).

13. Esteves, C. V. *et al.* Diagnostic potential of saliva proteome analysis: A review and guide to clinical practice. *Braz Oral Res* **33**, (2019).

14. Gooley, A. & Packer, N. *Proteome Research: New Frontiers in Functional Genomics*. *Gooley, A. & Packer, N. in Proteome Research: New Frontiers in Functional Genomics (eds. Wilkins, W. et al.) 65–91 (Springer, New York, 1997).* (1997).

15. Wilkins, M. R. *et al.* From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Bio/Technology* **14**, (1996).

16. Wilkins, M. R. *et al.* Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* **13**, (1996).

17. Kenyon, G. L. *et al.* Defining the mandate of proteomics in the post-genomics era: workshop report. in *Molecular & cellular proteomics : MCP* vol. 1 (2002).

18. Simonian, M. Proteomics in Medicine. *J Data Mining Genomics Proteomics* **7**, (2016).

19. Berggård, T., Linse, S. & James, P. Methods for the detection and analysis of protein-protein interactions. *Proteomics* vol. 7 (2007).

20. Zhou, M., Li, Q. & Wang, R. Current Experimental Methods for Characterizing Protein-Protein Interactions. *ChemMedChem* vol. 11 (2016).

21. Siqueira, W. L., Lee, Y. H., Xiao, Y., Held, K. & Wong, W. Identification and characterization of histatin 1 salivary complexes by using mass spectrometry. *Proteomics* **12**, (2012).

22. Oppenheim, F. G. *et al.* Histatins, a novel family of histidine-rich proteins in human parotid secretion. Isolation, characterization, primary structure, and fungistatic effects on Candida albicans. *Journal of Biological Chemistry* **263**, (1988).

23. Xiuli, S., Erdjan, S., Oppenheim, F. G. & Helmerhorst, E. J. Kinetics of histatin proteolysis in whole saliva and the effect on bioactive domains with metal-binding, antifungal, and wound-healing properties. *The FASEB Journal* **23**, (2009).

24. Contreras-Aguilar, M. D. *et al.* Changes occurring on the activity of salivary alpha-amylase proteoforms in two naturalistic situations using a spectrophotometric assay. *Biology (Basel)* **10**, (2021).

25. Crosara, K. T. B. *et al.* Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches. *Biomed Res Int* **2018**, (2018).

26. Michishige, F. *et al.* Effect of saliva collection method on the concentration of protein components in saliva. *Journal of Medical Investigation* **53**, (2006).

27. Windeler, A. S. & Shannon, I. L. Effect of flow rate on parotid fluid calcium, magnesium, and protein concentrations in man. *Arch Oral Biol* **12**, (1967).

28. Dawes, C. Stimulus effects on protein and electrolyte concentrations in parotid saliva. *J Physiol* **346**, (1984).

29. Shannon, I. L., Suddick, R. P. & Chauncey, H. H. Effect of atropine-induced flow rate depression on the composition of unstimulated human parotid fluid. *Arch Oral Biol* **14**, (1969).

30. Dawes, C., Dowse, C. M. & Knull, H. R. Stop-flow effects on human salivary composition and hydrostatic pressures. *Arch Oral Biol* **25**, (1980).

31. Dawes, C. The effect of flow rate and length of stimulation on the protein concentration in human parotid saliva. *Arch Oral Biol* **12**, (1967).

32. Dawes, C. & Kubieniec, K. The effects of prolonged gum chewing on salivary flow rate and composition. *Arch Oral Biol* **49**, (2004).

33. Dawes, C. & Jenkins, G. N. The effects of different stimuli on the composition of saliva in man. *J Physiol* **170**, (1964).

34. Shannon, I. L. Effects of Visual and Olfactory Stimulation on Parotid Secretion Rate in the Human (38259). *Proceedings of the Society for Experimental Biology and Medicine* **146**, (1974).

35. Chen, C. L., Su, J. Z. & Yu, G. Y. Effects of acid stimulation on saliva flow rate and compositions of human parotid and submandibular glands. *Beijing da xue xue bao. Yi xue ban = Journal of Peking University. Health sciences* **54**, (2022).

36. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human submandibular saliva. *Arch Oral Biol* **19**, 887–895 (1974).

37. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human parotid saliva. *Arch Oral Biol* **14**, 277–294 (1969).

38. Guidelines for Saliva Nomenclature and Collection. *Ann N Y Acad Sci* **694**, (1993).

39. Ekström, J., Khosravani, N., Castagnola, M. & Messana, I. Saliva and the control of its secretion. in *Medical Radiology* vol. 0 (2019).

40. Hollins, C. *Basic Guide to Anatomy and Physiology for Dental Care Professionals*. *Basic Guide to Anatomy and Physiology for Dental Care Professionals* (2012). doi:10.1002/9781118702789.

41. Valstar, M. H. *et al.* The tubarial salivary glands: A potential new organ at risk for radiotherapy. *Radiotherapy and Oncology* **154**, (2021).

42. Veerman, E. C. I., van den Keybus, P. A. M., Vissink, A. & Nieuw Amerongen, A. v. Human glandular salivas: Their separate collection and analysis. *Eur J Oral Sci* **104**, (1996).

43. Pedersen, A. M. L., Sørensen, C. E., Proctor, G. B., Carpenter, G. H. & Ekström, J. Salivary secretion in health and disease. *Journal of Oral Rehabilitation* vol. 45 (2018).

44. Proctor, G. B. The physiology of salivary secretion. *Periodontology 2000* vol. 70 (2016).

45. Lee, M. G., Ohana, E., Park, H. W., Yang, D. & Muallem, S. Molecular mechanism of pancreatic and salivary gland fluid and HCO 3- secretion. *Physiological Reviews* vol. 92 (2012).

46. Proctor, G. B. & Carpenter, G. H. Regulation of salivary gland function by autonomic nerves. *Autonomic Neuroscience: Basic and Clinical* vol. 133 (2007).

47. Pedersen, A. M. L., Sørensen, C. E., Dynesen, A. W. & Jensen, S. B. Salivary gland structure and functions and regulation of saliva secretion in health and disease. in *Salivary Glands Anatomy, Functions in Digestion and Role in Disease* (2013).

48. Humphrey & Williamson, R. T. A review of saliva Normal composition, flow, and function. Humphrey, Williamson. 2001. Journal of Prosthetic Dentistry.pdf. *J Prosthet Dent* **85**, (2001).

49. de Almeida, P. D. V., Grégio, A. M. T., Machado, M. Â. N., de Lima, A. A. S. & Azevedo, L. R. Saliva composition and functions: A comprehensive review. *Journal of Contemporary Dental Practice* vol. 9 (2008).

50. Ericson, S. & Wiberg, A. The secretory rate of the human submandibular glands in response to parasympathicomimetic and gustatory stimulation. *Arch Oral Biol* **17**, (1972).

51. Pijpe, J. *et al.* Progression of salivary gland dysfunction in patients with Sjögren's syndrome. *Ann Rheum Dis* **66**, (2007).

52. Dawes, C. Factors influencing salivary flow rate and composition. in *Saliva and Oral Health* vol. 56 (2004).

53. Hardt, M. *et al.* Toward defining the human parotid gland salivary proteome and peptidome: Identification and characterization using 2D SDS-PAGE, ultrafiltration, HPLC, and mass spectrometry. *Biochemistry* **44**, (2005).

54. Hardt, M. *et al.* Assessing the effects of diurnal variation on the composition of human parotid saliva: Quantitative analysis of native peptides using iTRAQ reagents. *Anal Chem* **77**, (2005).

55. Ryu, O. H., Atkinson, J. C., Hoehn, G. T., Illei, G. G. & Hart, T. C. Identification of parotid salivary biomarkers in Sjögren's syndrome by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry and two-dimensional difference gel electrophoresis. *Rheumatology* **45**, (2006).

56. Neyraud, E., Sayd, T., Morzel, M. & Dransfield, E. Proteomic analysis of human whole and parotid salivas following stimulation by different tastes. *J Proteome Res* **5**, (2006).

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

58. Ramachandran, P. *et al.* Comparison of N-linked glycoproteins in human whole saliva, parotid, submandibular, and sublingual glandular secretions identified using hydrazide chemistry and mass spectrometry. *Clin Proteomics* **4**, (2008).

59. Siqueira, W. L., Salih, E., Wan, D. L., Helmerhorst, E. J. & Oppenheim, F. G. Proteome of human minor salivary gland secretion. *J Dent Res* **87**, (2008).

60. Gonzalez-Begne, M. *et al.* Proteomic analysis of human parotid gland exosomes by multidimensional protein identification technology (MudPIT). *J Proteome Res* **8**, (2009).

61. Jasim, H., Olausson, P., Hedenberg-Magnusson, B., Ernberg, M. & Ghafouri, B. The proteomic profile of whole and glandular saliva in healthy pain-free subjects. *Sci Rep* **6**, (2016).

62. Schulte, F., Hasturk, H. & Hardt, M. Mapping Relative Differences in Human Salivary Gland Secretions by Dried Saliva Spot Sampling and nanoLC–MS/MS. *Proteomics* **19**, (2019).

63. Saitou, M. *et al.* Functional Specialization of Human Salivary Glands and Origins of Proteins Intrinsic to Human Saliva. *Cell Rep* **33**, (2020).

64. Ambort, D. *et al.* Calcium and pH-dependent packing and release of the gel-forming MUC2 mucin. *Proc Natl Acad Sci U S A* **109**, (2012).

65. Garrett, J. R. The Proper Role of Nerves in Salivary Secretion: A Review. *Journal of Dental Research* vol. 66 (1987).

66. Lee, V. & Linden, R. An olfactory-submandibular salivary reflex in humans. *Exp Physiol* **77**, (1992).

67. Shannon, I. L. & Frome, W. J. Enhancement of salivary flow rate and buffering capacity. *J Can Dent Assoc (Tor)* **39**, (1973).

68. Becks, H. & Wainwright, W. W. Human Saliva:XIII. Rate of Flow of Resting Saliva of Healthy Individuals. *J Dent Res* **22**, (1943).

69. Dawes, C. Circadian rhythms in human salivary flow rate and composition. *J Physiol* **220**, (1972).

70. Kariyawasam, A. P. & Dawes, C. A circannual rhythm in unstimulated salivary flow rate when the ambient temperature varies by only about 2°C. *Arch Oral Biol* **50**, (2005).

71. Ligtenberg, A. J. M., Meuffels, M. & Veerman, E. C. I. Effects of environmental temperature on saliva flow rate and secretion of protein, amylase and mucin 5B. *Arch Oral Biol* **109**, (2020).

72. Watanabe, S. & Dawes, C. A comparison of the effects of tasting and chewing foods on the flow rate of whole saliva in man. *Arch Oral Biol* **33**, (1988).

73. Shannon, I. L. & Chauncey, H. H. Hyperhydration and Parotid Flow in Man. *J Dent Res* **46**, (1967).

74. Sas, R. & Dawes, C. The intra-oral distribution of unstimulated and chewing-gum-stimulated parotid saliva. *Arch Oral Biol* **42**, (1997).

75. Lorenz, K. *et al.* Orosensory stimulation effects on human saliva proteome. *J Agric Food Chem* **59**, (2011).

76. Dawes, C. Effects of Diet on Salivary Secretion and Composition. *J Dent Res* **49**, (1970).

77. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human parotid saliva. *Arch Oral Biol* **14**, (1969).

78. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human submandibular saliva. *Arch Oral Biol* **19**, (1974).

79. Shannon, T. L. Parotid fluid flow rate as related to whole saliva volume. *Arch Oral Biol* **7**, (1962).

80. Rayment, S. A., Liu, B., Offner, G. D., Oppenheim, F. G. & Troxler, R. F. Immunoquantification of human salivary mucins MG1 and MG2 in stimulated whole saliva: Factors influencing mucin levels. *J Dent Res* **79**, (2000).

81. Dawes, C. The composition of human saliva secreted in response to a gustatory stimulus and to pilocarpine. *J Physiol* **183**, (1966).

82. Wu, S. *et al.* Quantitative analysis of human salivary gland-derived intact proteome using top-down mass spectrometry. *Proteomics* **14**, (2014).

83. Jensen, J. L., Lamkin, M. S. & Oppenheim, F. G. Adsorption of Human Salivary Proteins to Hydroxyapatite: A Comparison Between Whole Saliva and Glandular Salivary Secretions. *J Dent Res* **71**, (1992).

84. Dawes, C. Physiological Factors Affecting Salivary Flow Rate, Oral Sugar Clearance, and the Sensation of Dry Mouth in Man. *J Dent Res* **66**, (1987).

85. Heintze, U., Birkhed, D. & Björn, H. Secretion rate and buffer effect of resting and stimulated whole saliva as a function of age and sex. *Swed Dent J* **7**, (1983).

86. Dawes, C., Cross, H. G., Baker, C. G. & Chebib, F. S. The influence of gland size on the flow rate and composition of human parotid saliva. *Dent J* **44**, (1978).

87. Wang, B., Danjo, A., Kajiya, H., Okabe, K. & Kido, M. A. Oral epithelial cells are activated via TRP channels. *J Dent Res* **90**, (2011).

88. Lee, A., Guest, S. & Essick, G. Thermally evoked parotid salivation. *Physiol Behav* **87**, (2006).

89. Bosch, J. A. *et al.* Psychological stress as a determinant of protein levels and salivary- induced aggregation of Streptococcus gordonii in human whole saliva. *Psychosom Med* **58**, (1996).

90. Dawes, C. The effects of exercise on protein and electrolyte secretion in parotid saliva. *J Physiol* **320**, (1981).

91. Ligtenberg, A. J. M., Brand, H. S., van den Keijbus, P. A. M. & Veerman, E. C. I. The effect of physical exercise on salivary secretion of MUC5B, amylase and lysozyme. *Arch Oral Biol* **60**, (2015).

92. Fortes, M. B., Diment, B. C., di Felice, U. & Walsh, N. P. Dehydration decreases saliva antimicrobial proteins important for mucosal immunity. *Applied Physiology, Nutrition and Metabolism* **37**, (2012).

93. Ship, J. A. & Fischer, D. J. The relationship between dehydration and parotid salivary gland function in young and older healthy adults. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* **52**, (1997).

94. Walsh, N. P., Montague, J. C., Callow, N. & Rowlands, A. v. Saliva flow rate, total protein concentration and osmolality as potential markers of whole body hydration status during progressive acute dehydration in humans. *Arch Oral Biol* **49**, (2004).

95. Villa, A. *et al.* World Workshop on Oral Medicine VI: a systematic review of medication-induced salivary gland dysfunction. *Oral Diseases* vol. 22 (2016).

96. Wolff, A., Zuk-Paz, L. & Kaplan, I. Major salivary gland output differs between users and non-users of specific medication categories. *Gerodontology* **25**, (2008).

97. Wu, A. J. & Ship, J. A. A characterization of major salivary gland flow rates in the presence of medications and systemic diseases. *Oral Surgery, Oral Medicine, Oral Pathology* **76**, (1993).

98. Wolff, A. *et al.* A Guide to Medications Inducing Salivary Gland Dysfunction, Xerostomia, and Subjective Sialorrhea: A Systematic Review Sponsored by the World Workshop on Oral Medicine VI. *Drugs in R and D* vol. 17 (2017).

99. Dodds, M. W. J., Johnson, D. A. & Yeh, C. K. Health benefits of saliva: A review. *J Dent* **33**, (2005).

100. Percival, R. S., Challacombe, S. & Marsh, P. D. Flow Rates of Resting Whole and Stimulated Parotid Saliva in Relation to Age and Gender. *J Dent Res* **73**, (1994).

101. Proctor, G. B. Medication-Induced Dry Mouth. in *Dry Mouth* (2015).

102. Carpenter, G. Dry mouth: A Clinical Guide on Causes, Effects and Treatments. *Mayo Clinic* (2016).

103. Helmerhorst, E. J. & Oppenheim, F. G. Saliva: A dynamic proteome. *Journal of Dental Research* vol. 86 (2007).

104. Featherstone, J. D. B. The caries balance: the basis for caries management by risk assessment. *Oral Health Prev Dent* **2 Suppl 1**, (2004).

105. Featherstone, J. D. B. The continuum of dental caries - Evidence for a dynamic disease process. in *Journal of Dental Research* vol. 83 (2004).

106. Featherstone, J. D. B. Dental caries: A dynamic disease process. in *Australian Dental Journal* vol. 53 (2008).

107. Cai, K. & Bennick, A. Processing of acidic proline-rich proprotein by human salivary gland convertase. *Arch Oral Biol* **49**, (2004).

108. Lu, Y. & Bennick, A. Interaction of tannin with human salivary proline-rich proteins. *Arch Oral Biol* **43**, (1998).

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

110. Xu, T., Levitz, S. M., Diamond, R. D. & Oppenheim, F. G. Anticandidal activity of major human salivary histatins. *Infect Immun* **59**, (1991).

111. Gibbons, R. J., Hay, D. I. & Schlesinger, D. H. Delineation of a segment of adsorbed salivary acidic proline-rich proteins which promotes adhesion of Streptococcus gordonii to apatitic surfaces. *Infect Immun* **59**, (1991).

112. Oppenheim, F. G. *et al.* The primary structure and functional characterization of the neutral histidine-rich polypeptide from human parotid secretion. *Journal of Biological Chemistry* **261**, (1986).

113. Hannig, M. The protective nature of the salivary pellicle. *Int Dent J* **52**, (2002).

114. Dawes, C., Watanabe, S., Biglow-Lecomte, P. & Dibdin, G. H. Estimation of the Velocity of the Salivary Film at Some Different Locations in the Mouth. *J Dent Res* **68**, (1989).

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

116. Pramanik, R., Osailan, S. M., Challacombe, S. J., Urquhart, D. & Proctor, G. B. Protein and mucin retention on oral mucosal surfaces in dry mouth patients. *Eur J Oral Sci* **118**, (2010).

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

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

119. Zimmerman, J. N. *et al.* Proteome and peptidome of human acquired enamel pellicle on deciduous teeth. *Int J Mol Sci* **14**, (2013).

120. Carpenter, G. *et al.* Composition of enamel pellicle from dental erosion patients. *Caries Res* **48**, (2014).

121. Lee, Y. H. *et al.* Proteomic Evaluation of Acquired Enamel Pellicle during In Vivo Formation. *PLoS One* **8**, (2013).

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

123. Nieuw Amerongen, A. v. & Veerman, E. C. I. Saliva - The defender of the oral cavity. *Oral Dis* **8**, (2002).

124. Dawes, C. *et al.* The functions of human saliva: A review sponsored by the World Workshop on Oral Medicine VI. *Archives of Oral Biology* vol. 60 (2015).

125. Dawes, C. Factors Influencing Protein Secretion in Human Saliva. in *The environment of the teeth* vol. 3 125-137 (1981).

126. Oudhoff, M. J. *et al.* Histatins are the major wound-closure stimulating factors in human saliva as identified in a cell culture assay. *The FASEB Journal* **22**, (2008).

127. Torres, P., Castro, M., Reyes, M. & Torres, V. A. Histatins, wound healing, and cell migration. *Oral Diseases* vol. 24 (2018).

128. Oudhoff, M. J. *et al.* Structure-activity analysis of histatin, a potent wound healing peptide from human saliva: Cyclization of histatin potentiates molar activity 1000-fold. *FASEB Journal* **23**, (2009).

129. Cheng, L. *et al.* Histatin 1 enhanced the speed and quality of wound healing through regulating the behaviour of fibroblast. *Cell Prolif* **54**, (2021).

130. Shah, D. *et al.* Wound Healing Properties of Histatin-5 and Identification of a Functional Domain Required for Histatin-5-Induced Cell Migration. *Mol Ther Methods Clin Dev* **17**, (2020).

131. Oudhoff, M. J. *et al.* The role of salivary histatin and the human cathelicidin LL-37 in wound healing and innate immunity. *Biol Chem* **391**, (2010).

132. van Nieuw Amerongen, A., Bolscher, J. G. M. & Veerman, E. C. I. Salivary proteins: Protective and diagnostic value in cariology? in *Caries Research* vol. 38 (2004).

133. Simonsen, K. W. *et al.* Presence of psychoactive substances in oral fluid from randomly selected drivers in Denmark. *Forensic Sci Int* **221**, (2012).

134. Corstjens, P. L. A. M., Abrams, W. R. & Malamud, D. Detecting viruses by using salivary diagnostics. *Journal of the American Dental Association* **143**, (2012).

135. Ramanujam, A., Almodovar, S. & Botte, G. G. Ultra-fast electrochemical sensor for point-of-care covid-19 diagnosis using non-invasive saliva sampling. *Processes* **9**, (2021).

136. Regalado, A. 2017 was the year consumer DNA testing blew up. *MIT Technology Review* (2018).

137. Jahr, S. *et al.* DNA fragments in the blood plasma of cancer patients: Quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res* **61**, (2001).

138. Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L. & Quake, S. R. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc Natl Acad Sci U S A* **105**, (2008).

139. Diehl, F. *et al.* Circulating mutant DNA to assess tumor dynamics. *Nat Med* **14**, (2008).

140. Diehl, F. *et al.* Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc Natl Acad Sci U S A* **102**, (2005).

141. Diaz, L. A. & Bardelli, A. Liquid biopsies: Genotyping circulating tumor DNA. *Journal of Clinical Oncology* vol. 32 Preprint at https://doi.org/10.1200/JCO.2012.45.2011 (2014).

142. Mouliere, F. *et al.* High fragmentation characterizes tumour-derived circulating DNA. *PLoS One* **6**, (2011).

143. Overman, M. J. *et al.* Use of research biopsies in clinical trials: Are risks and benefits adequately discussed? *Journal of Clinical Oncology* **31**, (2013).

144. Haber, D. A. & Velculescu, V. E. Blood-based analyses of cancer: Circulating tumor cells and circulating tumor DNA. *Cancer Discovery* vol. 4 Preprint at https://doi.org/10.1158/2159-8290.CD-13-1014 (2014).

145. Park, N. J. *et al.* Salivary microRNA: Discovery, characterization, and clinical utility for oral cancer detection. *Clinical Cancer Research* **15**, (2009).

146. Spielmann, N. & Wong, D. Saliva: Diagnostics and therapeutic perspectives. *Oral Diseases* vol. 17 (2011).

147. Bahn, J. H. *et al.* The landscape of MicroRNA, piwi-interacting RNA, and circular RNA in human saliva. *Clin Chem* **61**, (2015).

148. Kaur, J., Jacobs, R., Huang, Y., Salvo, N. & Politis, C. Salivary biomarkers for oral cancer and pre-cancer screening: a review. *Clinical Oral Investigations* vol. 22(2018).

149. Khurshid, Z. *et al.* Role of Salivary Biomarkers in Oral Cancer Detection. *Adv Clin Chem* **86**, (2018).

150. Li, Q., Ouyang, X., Chen, J., Zhang, P. & Feng, Y. A review on salivary proteomics for oral cancer screening. *Current Issues in Molecular Biology* vol. 37 (2020).

151. Xiao, H. *et al.* Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. *Molecular and Cellular Proteomics* **11**, (2012).

152. Wang, C. *et al.* Affinity chromatography assisted comprehensive phosphoproteomics analysis of human saliva for lung cancer. *Anal Chim Acta* **1111**, (2020).

153. H., X. *et al.* Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. *Molecular and Cellular Proteomics* vol. 11 (2012).

154. Porto-Mascarenhas, E. C. *et al.* Salivary biomarkers in the diagnosis of breast cancer: A review. *Critical Reviews in Oncology/Hematology* vol. 110 (2017).

155. Zhang, L. *et al.* Discovery and preclinical validation of salivary transcriptomic and proteomic biomarkers for the non- invasive detection of breast cancer. *PLoS One* **5**, (2010).

156. Li, Y. *et al.* Salivary transcriptome diagnostics for oral cancer detection. *Clinical Cancer Research* **10**, (2004).

157. Birney, E. The International Human Genome Project. *Human Molecular Genetics* vol. 30 (2021).

158. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, (2012).

159. Moraes, F. & Góes, A. A decade of human genome project conclusion: Scientific diffusion about our genome knowledge. *Biochemistry and Molecular Biology Education* **44**, (2016).

160. Ruhl, S. The scientific exploration of saliva in the post-proteomic era: From database back to basic function. *Expert Review of Proteomics* vol. 9 (2012).

161. Kastritis, P. L. & Bonvin, A. M. J. J. On the binding affinity of macromolecular interactions: Daring to ask why proteins interact. *Journal of the Royal Society Interface* vol. 10 (2013).

162. Crosara, K. T. B., Moffa, E. B., Xiao, Y. & Siqueira, W. L. Merging in-silico and in vitro salivary protein complex partners using the STRING database: A tutorial. *J Proteomics* (2017)

163. de Las Rivas, J. & de Luis, A. Interactome data and databases: Different types of protein interaction. *Comparative and Functional Genomics* vol. 5 (2004).

164. Das, J. & Yu, H. HINT: High-quality protein interactomes and their applications in understanding human disease. *BMC Syst Biol* **6**, (2012).

165. Siqueira, W. L., Lee, Y. H., Xiao, Y., Held, K. & Wong, W. Identification and characterization of histatin 1 salivary complexes by using mass spectrometry. *Proteomics* **12**, 3426–3435 (2012).

166. D204-D212. UniProt: a hub for protein information The UniProt Consortium. *Nucleic Acids Res* **43**, (2015).

167. Apweiler, R. *et al.* UniProt: The universal protein knowledgebase. *Nucleic Acids Res* **32**, (2004).

168. Chatr-Aryamontri, A. *et al.* The BioGRID interaction database: 2015 update. *Nucleic Acids Res* **43**, D470–D478 (2015).

169. Orchard, S. *et al.* The MIntAct project - IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res* **42**, (2014).

170. Dennis, G. *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* **4**, (2003).

171. Kotlyar, M., Pastrello, C., Sheahan, N. & Jurisica, I. Integrated interactions database: Tissue-specific view of the human and model organism interactomes. *Nucleic Acids Res* **44**, D536–D541 (2016).

172. Schmitt, T., Ogris, C. & Sonnhammer, E. L. L. FunCoup 3.0: Database of genome-wide functional coupling networks. *Nucleic Acids Res* **42**, (2014).

173. Mi, H. *et al.* The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res* **33**, (2005).

174. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research* vol. 28 (2000).

175. Szklarczyk, D. *et al.* The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**, D362–D368 (2017).

176. Szklarczyk, D. *et al.* The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* **49**, (2021).

177. Turner, B. *et al.* iRefWeb: interactive analysis of consolidated protein interaction data and their supporting evidence. *Database (Oxford)* **2010**, (2010).

178. Rudney, J. D. Implications of a model from olfactory research for the use of secretion rates in salivary studies. *Arch Oral Biol* **35**, (1990).

179. Rudney, J. D. Relationships between human parotid saliva lysozyme lactoferrin, salivary peroxidase and secretory immunoglobulin A in a large sample population. *Arch Oral Biol* **34**, (1989).

180. Rudney, J. D., Kajander, K. C. & Smith, Q. T. Correlations between human salivary levels of lysozyme, lactoferrin, salivary peroxidase and secretory immunoglobulin A with different stimulatory states and over time. *Arch Oral Biol* **30**, (1985).

181. Bellagambi, F. G. *et al.* Saliva sampling: Methods and devices. An overview. *TrAC - Trends in Analytical Chemistry* vol. 124 (2020).

182. Shannon, I. L., Prigmore, J. R. & Chauncey, H. H. Modified Carlson-Crittenden Device for the Collection of Parotid Fluid. *J Dent Res* **41**, (1962).

183. Mann, M., Hendrickson, R. C. & Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annual Review of Biochemistry* vol. 70 (2001).

184. Arndt, C., Koristka, S., Feldmann, A. & Bachmann, M. Native polyacrylamide gels. in *Methods in Molecular Biology* vol. 1855 (2019).

185. Lee, P. Y., Saraygord-Afshari, N. & Low, T. Y. The evolution of two-dimensional gel electrophoresis - from proteomics to emerging alternative applications. *J Chromatogr A* **1615**, (2020).

186. Ramos, Y. *et al.* Sodium dodecyl sulfate free gel electrophoresis/electroelution sorting for peptide fractionation. *J Sep Sci* **42**, (2019).

187. Buxbaum, E. Cationic electrophoresis. in *Methods in Molecular Biology* vol. 1855 (2019).

188. Yates, J. R., Ruse, C. I. & Nakorchevsky, A. Proteomics by mass spectrometry: Approaches, advances, and applications. *Annual Review of Biomedical Engineering* vol. 11 (2009).

189. Angel, T. E. *et al.* Mass spectrometry-based proteomics: Existing capabilities and future directions. *Chem Soc Rev* **41**, (2012).

190. Sondej, M. *et al.* Glycoprofiling of the human salivary proteome. *Clin Proteomics* **5**, (2009).

191. Gillet, L. C., Leitner, A. & Aebersold, R. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annual Review of Analytical Chemistry* vol. 9 (2016).

192. Nagaraj, N. *et al.* System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top orbitrap. *Molecular and Cellular Proteomics* **11**, (2012).

193. Mann, K. & Mann, M. In-depth analysis of the chicken egg white proteome using an LTQ Orbitrap Velos. *Proteome Sci* **9**, (2011).

194. Mi, H. *et al.* PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* **45**, (2017).

195. Bastian, F. B. *et al.* The Bgee suite: Integrated curated expression atlas and comparative transcriptomics in animals. *Nucleic Acids Res* **49**, (2021).

196. Digre, A. & Lindskog, C. The Human Protein Atlas—Spatial localization of the human proteome in health and disease. *Protein Science* **30**, (2021).

197. Lau, W. W., Hardt, M., Zhang, Y. H., Freire, M. & Ruhl, S. The Human Salivary Proteome Wiki: A Community-Driven Research Platform. *J Dent Res* **100**, (2021).

198. Goldring, J. P. D. Concentrating proteins by salt, polyethylene glycol, solvent, sds precipitation, three-phase partitioning, dialysis, centrifugation, ultrafiltration, lyophilization, affinity chromatography, immunoprecipitation or increased temperature for protein isolation, drug interaction, and proteomic and peptidomic evaluation. in *Methods in Molecular Biology* vol. 1855 (2019).

199. van Zundert, G. C. P. & Bonvin, A. M. J. J. Modeling protein–protein complexes using the HADDOCK webserver "modeling protein complexes with HADDOCK". *Methods in Molecular Biology* **1137**, (2014).

# Chapter 2

# 2 Merging *In-silico* and *In Vitro* Salivary Protein Complex Partners Using the STRING Database

This chapter has been adapted from the publication titled "Merging in-silico and in vitro salivary protein complex partners using the STRING database: A tutorial.", published in *Journal of Proteomics* in 2018 by Karla Tonelli Bicalho Crosara, Eduardo Buozi Moffa, Yizhi Xiao, and Walter Luiz Siqueira.

## 2.1 Abstract

Protein-protein interaction (PPI) is a common physiological mechanism for the protection and actions of proteins in an organism. The identification and characterization of PPIs in different organisms is necessary to better understand their physiology and to determine their efficacy. In a previous *in vitro* study using mass spectrometry, we identified 43 proteins that interact with histatin 1. Six previously documented interactors were confirmed, and 37 novel partners were identified. In this chapter, we demonstrate the usefulness of the STRING database for studying protein-protein interactions. We used an in-silico approach along with the STRING database (http://string-db.org/) and successfully performed a fast simulation of a novel constructed histatin 1 protein-protein network, including both the previously known and the predicted interactors, along with our newly identified interactors. Our study highlights the advantages and importance of applying bioinformatics tools to merge in-silico tactics with experimental *in vitro* findings for the rapid advancement of our knowledge about PPIs. Our findings also indicate that bioinformatics tools such as the STRING protein network database can help predict potential interactions between proteins and thus serve as a guide for future steps in our exploration of the Human Interactome.

## 2.2 Introduction

The study on the biological roles of proteins in a system begin with the characterization of individual proteins. However, proteins usually interact with other proteins *in vivo* for many purposes and in different ways. Proteins are likely to interact directly by binding to other proteins, as a defense mechanism against degradation; to modulate the function of one or both partners; or to rely on a partner as a delivery system for transport. In addition, proteins also communicate through indirect interactions to regulate each other's production and half-life, exchange reaction products, and activate/deactivate different signaling pathways, thus contributing to the functioning of the whole organism. The broad combination of these direct and indirect interactions determines the functional associations among proteins [1–6].Different experimental methods can be used to identify proteins that interact directly to form heterotypic complexes. Co-immunoprecipitation (Co-IP) and/or pull-down assays associated with mass spectrometry (MS) are most commonly used for this purpose. Both these approaches are very efficient for the recovery of protein complexes when the protein partners exhibit strong and stable interactions. In Co-IP, the target protein precipitated by the antibody is used to co-precipitate a binding partner/protein complex from a mixture. On the other hand, in the pull-down assay, a "bait" protein is used instead of an antibody to extract proteins that bind to it from the mixture or protein complexes that contain proteins that bind to the "bait" protein. The combination of both these methods helps efficiently identify heterotypic complex partners, which helps avoid false positive results.

After separating the proteins of interest from a mixture and identifying each amino acid constituent and post-translational modification, all the efforts turn to the difficult task of characterizing PPIs and their biological roles. Many databases and online resources have been created to assist in this titanic mission. First, evidence-based PPIs are curated from the published literature by members of the UniProt [7] and IMEx [8,9] consortia. Based on this information, databases such as BioGRID [10], HINT [11], iRefWeb [12], and APID[13]collect all data about the interactions proven in the experiments. Finally, databases designed to build on top of the data obtained about direct PPIs in the experiments add data about indirect and predicted PPIs to create a more comprehensive network. Some examples from this last group of bioinformatics tools are GeneMANIA [14], Integrated Multi-species Prediction [15],

Integrated Interactions Databases [16], HumanNet [17], FunCoup [18], and the STRING database [19,20]. In these latter databases, scores are provided to weigh interactions based on their confidence.

In the STRING database, for example, each stored PPI has a score (between zero and one) representing its confidence. The supporting evidence for each interaction is divided into seven "evidence channels." The *experiments* channel is formed by results from the lab that proved protein interactions (including biochemical, biophysical, and genetic experiments), mostly data from the IMEx consortium and BioGRID. The *database* channel is manually curated and imported from pathway databases. The *textmining* channel indicates possible interactions between proteins that are mentioned in the same PubMed abstracts, in an in-house selection of more than 3 million full-text articles, and in other text collections [21,22]. The evidence is considered stronger if a concept such as "binding" or "phosphorylated by" is found to connect the mentioned proteins. The *coexpression* channel shows normalized, pruned, and correlated gene expression data from many experiments [23]. The *neighborhood* channel is a genome-based prediction channel, where genes are given scores if they are frequently seen in each other's genome neighborhood. In the *fusion* channel, pairs of proteins are given association scores if there is at least one organism in which their respective orthologs have fused into a single, protein-coding gene. The *co-occurrence* channel evaluates the phylogenetic distribution of orthologs of all proteins in an organism; two proteins with high similarity in this distribution are assigned a score [24]. Finally, in addition to the seven listed evidence channels, the STRING database also benefits from the transfer of evidence from one organism to another, since orthologs of interacting proteins in one organism often also interact in other organisms; this is called "interolog" transfer [25,26].

In a previous study, our group identified 43 proteins in human saliva that participate in heterotypic complexes with histatin 1 [27], a salivary protein with many functions in the oral cavity, including strong antibacterial and antifungal functions. To exemplify the use of the STRING database in salivary proteomics research, in this chapter, we provide a detailed guide on how we used the STRING protein network database to combine the results of an in-silico approach to identify the proteins that interact with histatin 1 with the results of our

previous *in vitro* experiments, where Co-IP and pull-down assays, followed by MS, were used to identify the interactors of histatin 1 in saliva. Herein, we demonstrate how this skill web tool, STRING database, can be used to simulate protein-protein networks. We also discuss the importance and advantages of combining in-silico and *in vitro* approaches to provide a more comprehensive view of the real protein hub.

## 2.3 Methods and Results

### 2.3.1 Identification of the Complex Partners of Histatin 1

Heterotypic complexes of histatin 1 with other salivary proteins in parotid saliva were identified using classical PPI methods in combination with MS [27]. A co-IP assay using magnetic beads and a pull-down assay with immunopure immobilized streptavidin beads were used to separate the histatin 1 complexes from the mixture. In-solution tryptic digestion was performed, and the proteins present in the complex were identified using Reverse Phase Liquid Chromatography Electrospray Ionization Tandem MS (RP LC-ESI-MS/MS). Positive identification and characterization of the proteins/peptides were determined by the appearance of the same constituent in both methods. In total, 43 proteins were confirmed in the two approaches and identified as partners of histatin 1 in heterotypic complexes in saliva.

### 2.3.2 Representation of the First Shell Interactors in the Histatin 1 Network Using the STRING Database

The STRING database was used to simulate the protein-protein network with interactions between histatin 1 and other proteins. First, a representation was made with all the proteins known to interact with histatin 1 in the first shell of the protein complex, along with the proteins predicted to belong to the first shell of the complex based on the seven evidence channels used in this database. On the first page of the STRING database (http://string-db.org/), we selected **Protein by name** from the left menu bar, entered the protein name **Histatin 1**, selected organism **Homo sapiens**, and clicked **SEARCH**. From the top menu bar, by selecting **VIEWER**, the user can access more detailed information on one of the seven cited evidence channels. Here, we opted to use **NETWORK**, since it provides a summary of all the evidence channels used to create the link between the nodes (proteins)

(Figures 2.1 A and 2.1 B).

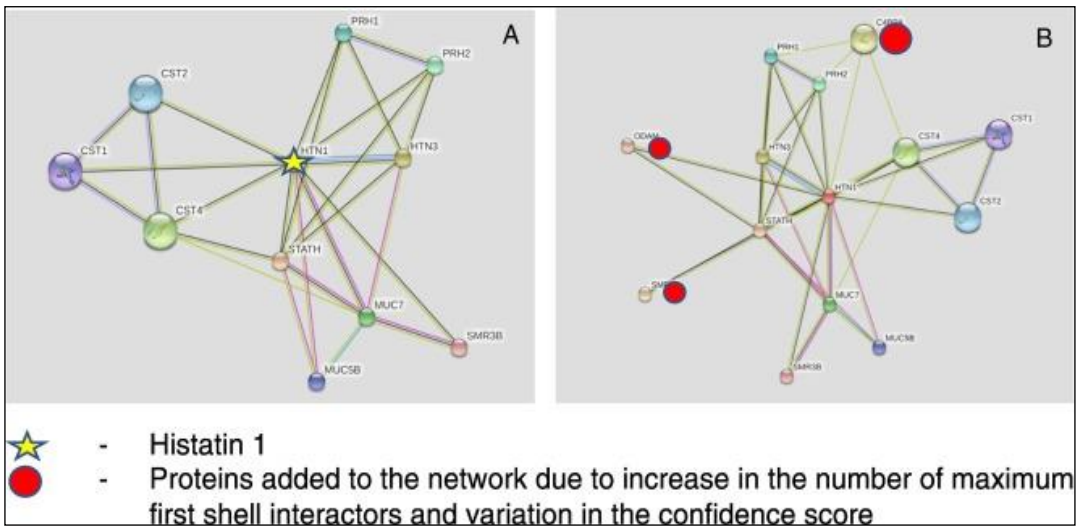



**Figure 2.1 PPI network of histatin 1 based on STRING database. (A) Histatin 1 PPI network with number of first shell interactors limited to 10 and using a medium confidence (0.400). (B) Histatin 1 PPI network with number of first shell interactors limited to 50 and using medium confidence (0.400).**

Under the option **LEGEND**, a brief description of the color code used to draw the network is provided, along with the description of the representative nodes for each protein in the network (Figure 2.2). The colored lines linking the interacting nodes highlight the evidence channel used to determine the relationship. Next, by choosing **SETTINGS**, and by increasing **the maximum number of first shell interactors,** all the possible partners of histatin 1 listed in the database were included (Figure 2.1 B). This step is very important, since limiting the number of confirmed/predicted first shell interactors to a maximum of 5 or 10, for example, may not be enough to fully represent all the listed elements of a protein network. If suspecting that more interactors may be added, the viewer may include a custom number of displayed interactors greater than 50. Here, by increasing the limit for displayed interactors to up to 50, 3 additional histatin 1 fist shell interactors were added to the network (marked in red in Figure 2.1 B).

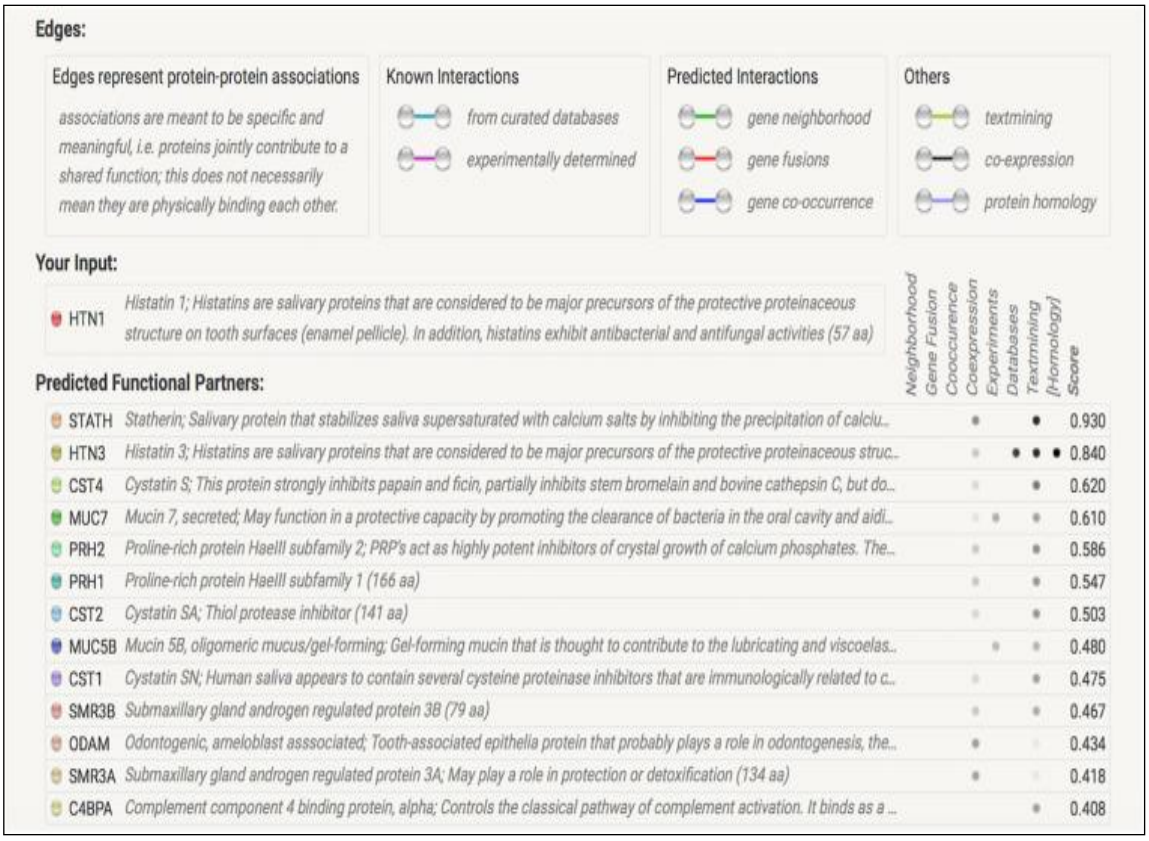



**Figure 2.2 Legends representing color code used to differentiate evidence channels used to establish the link between proteins represented in the histatin 1 network.**

The number of interactors can also be increased by lowering the setting of the minimum required confidence score used for the display. The confidence score is a number between zero and one given to a functional protein association considering the seven evidence channels previously described. The stronger the evidence that two proteins interact with each other, the higher the confidence score for that specific situation. However, although applying a lower confidence score as the cut-off for including possible partners broadens the inclusion criteria, it decreases the confidence of the network. For this reference exploration, the **confidence score** was set to **medium** (score of **0.400).** A list with all the proteins identified as first shell interactors of histatin 1 using the STRING database for an in-silico approach is shown in Table 2.1.

**Table 2.1 List of all proteins identified *in-silico* by the STRING database as first shell interactors of Histatin 1. Proteins marked in red were listed in the simulation using an increased number of maximum first shell interactors to be displayed.**

| Accession | | Name |
|---|---|---|
| **P15516** | HTN3 | Histatin 3 |
| **P02808** | STATH | Statherin |
| **Q8TAX7** | MUC7 | Mucin 7 |
| **Q9Hc84** | MUC5B | Mucin 5B |
| **P02814** | SMR3B | Submaxillary gland androgen-regulated protein 3B |
| **Q99954** | **SMR3A** | **Submaxillary gland androgen-regulated protein 3A** |
| **P01037** | CST1 | Cystatin-SN |
| **P09228** | CST2 | Cystatin-SA |
| **P01036** | CST4 | Cystatin-S |
| **P02810** | PRH1/PRH2 | Salivary acidic proline-rich phosphoprotein 1/2 |
| **P04003** | **C4BPA** | **Complement component 4 binding protein, alpha chain** |
| **A1E959** | **ODAM** | **Odontogenic ameloblast-associated protein** |

### 2.3.3 Use of the STRING Database for Network Simulation of the Histatin 1 Heterotypic Complexes Identified *In Vitro*

Next, a simulated protein-protein network was created with the 43 proteins identified in our previous study (Figure 2.3). Table 2.2 shows all the constituents of this network. To construct the simulated protein network in the STRING database, **Multiple Proteins** was

selected from the left menu bar, the accession numbers of all the identified proteins were copied into the correspondent box, and **Homo sapiens** was selected as the organism. Under the **SETTINGS** menu, the maximum number of first shell interactors was limited to the query proteins. To exemplify the effect of the confidence score on the construction of a protein-protein network, the level of confidence was changed to provide a more complete representation of the wide range of possibilities for the use of this bioinformatics tool. As a result, the simulated Histatin 1 interactome varied from having one single partner (statherin; Figure 2.3 A) to including all except one partner (Figure 2.3 C). For this demonstration, the simulated network for our identified partners of histatin 1 in saliva was constructed with a confidence score set to **highest** (0.900), **medium** (0.400), and **low** (0.150) (Figures 2.3 A, 2.3 B, and 2.3 C, respectively). The confidence score represents the probability that the marked interaction is biologically meaningful, specific, and reproducible based on the supporting evidence [28]. First, every interaction is divided into one or more evidence channels depending on its origin and type. Next, each evidence channel is scored separately and represented in different colors in the network. Depending on the user's need, each evidence channel can be disabled individually. Usually, stronger interactions not only present a higher score but are also supported by more than one evidence channel. A final confidence score is built as the result of the combination of all evidence channels. This "combined score" is used to construct networks and to filter interactions, thus acting as a cut-off measurement determining the inclusion/exclusion of partners from the represented network. Carefully varying the combined confidence score can provide interesting insights into the predicted interactions that have not yet been confirmed by *in vitro* experiments, thus suggesting a promising avenue to be explored in future studies. On the other hand, lowering the confidence score without caution may lead to a noisy network with many false associations.

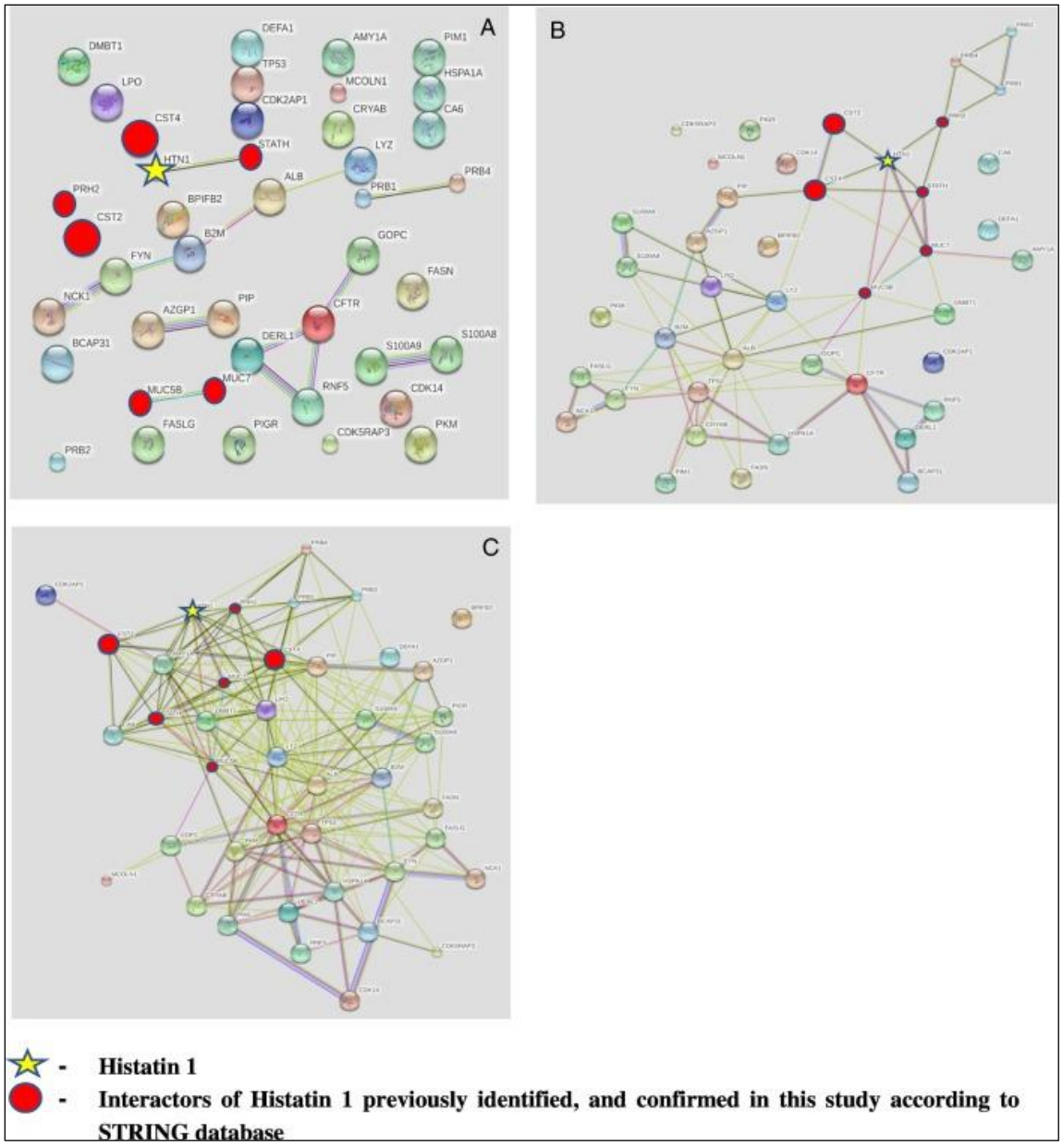



**Figure 2.3 Representation of protein-protein network of the *in vitro* identified histatin 1 interactors using STRING database. Figures show simulations for the same group of proteins with changes in the confidence score adopted: (A) highest confidence-0.900, (B) medium confidence – 0.400, and (C) low confidence – 0.150.**

**Table 2.2 List of the 43 proteins that make heterotypic complexes with histatin 1, identified after Co-IP and pull-down assay, followed by MS.**

| Accession number | | Name |
|---|---|---|
| **Q9HC84** | MUC5B | Mucin 5B |
| **P02808** | STATH | Statherin |
| **P04745** | AMY1A | Alpha-amylase |
| **P02768** | ALB | Serum albumin |
| **P02812** | PRB2 | Basic salivary proline-rich protein 2 |
| **P04280** | PRB1 | Basic salivary proline-rich protein 1 |
| **P23280** | CA6 | Carbonic anhydrase 6 |
| **P01876** | IGHA1 | Ig alpha-1 chain C region |
| **P61769** | B2M | Beta-2-microglobulin |
| **P02810** | PRH2 | Salivary acidic proline-rich phosphoprotein % (acidic PRP/2) |
| **P22079** | LPO | Lactoperoxidase (salivary peroxidase) |
| **P05109** | S100A8 | Protein S100-A8 (calgranulin-A) |
| **P01833** | PIGR | Poly Ig receptor |
| **P25311** | AZGP1 | Zinc-alpha-2-glycoprotein |
| **P09228** | CST2 | Cystatin-SA |
| **P13569** | CFTR | Cystic fibrosis transmembrane conductance regulator |
| **Q9UGM3** | DMBT1 | Deleted in malignant brain tumors 1 protein (glycoprotein 340) |

| Accession number | | Name |
|---|---|---|
| **P61626** | LYZ | Lysozyme C |
| **P48023** | FASLG | Tumor necrosis factor ligand superfamily member 6 |
| **P02511** | CRYAB | Alpha-crystallin B chain (HSP beta-5) |
| **P08107** | HSPA1A | Heat Shock 70 kDa protein 1A/1B |
| **P12273** | PIP | Prolactin inducible protein |
| **P04637** | TP53 | Cellular tumor antigen p53 |
| **Q8N4F0** | BPIFB2 | Bactericidal/permeability-increasing protein-like 1 |
| **P11309** | PIM1 | Proto-oncogene serine/threonine-protein kinase pim-1 |
| **P49327** | FASN | Fatty acid synthase |
| **P01834** | IGKC | Ig kappa chain C region |
| **P16333** | NCK1 | Cytoplasmic protein NCK1 |
| **P06241** | FYN | Tyrosine-protein kinase Fyn |
| **P06702** | S100A9 | Protein S100-A9 (calgranulin-B) |
| **O94921** | CDK14 | Cyclin-dependent kinase 14 |
| **Q9GZU1** | MCOLN1 | Mucolipin-1 (MG-2) |
| **P14618** | PKM | Pyruvate kinase isoenzymes MI/M2 |
| **O14519** | CDK2AP1 | Cyclin-dependent kinase 2-associated protein 1 |
| **Q96JB5** | CDK5RAP3 | CDK5 regulatory subunit-associated protein 3 |
| **Q9BUN8** | DERL1 | Derlin-1 |

| Accession number | | Name |
|---|---|---|
| **P51572** | BCAP31 | B-cell receptor associated protein 31 |
| **Q9HD26** | GOPC | Golgi-associated PDZ and coiled-coil motif-containing protein |
| **Q99942** | RNF5 | E3 ubiquitin-protein ligase RNF5 |
| **P59665** | DEFA1 | Neutrophil defensin 1 |
| **P10163** | PRB4 | Basic salivary proline-rich protein 4 allele S precursor |
| **Q8TAX7** | MUC7 | Mucin-7 (low molecular weight salivary mucin MG2) |
| **P01036** | CST4 | Cystatin-S |

### 2.3.4 Merging an *In-silico* Approach and the Histatin 1 Complex Partners Identified *In Vitro* Using the STRING Database

Finally, a new comprehensive histatin 1 interactome was constructed, merging the known and the predicted first shell interactors of histatin 1 identified in the STRING database (*in-silico*) with the proteins that were identified *in vitro* to form heterotypic complexes with histatin 1 in saliva (Figure 2.4). Under the menu **SETTINGS**, the lines between nodes in this representation were set to represent the confidence in the association occurrence, instead of the evidence channels involved. In this way, the results from the seven evidence channels are combined and represented as one single line. The thickness of the line linking two nodes that represent the interacting proteins varies based on the combined confidence score for the protein association. Higher combined confidence scores are represented by thicker lines. By clicking on the line between two interacting proteins, short information regarding the biological function of each participating protein is provided, along with details regarding the evidence that suggests the interaction between them, and a link to the supporting literature, this was exemplified with histatin 1 (Figure 2.4). Information about the proteins in the network is also available by clicking directly on the nodes, illustrated with the association between histatin 1 and statherin.

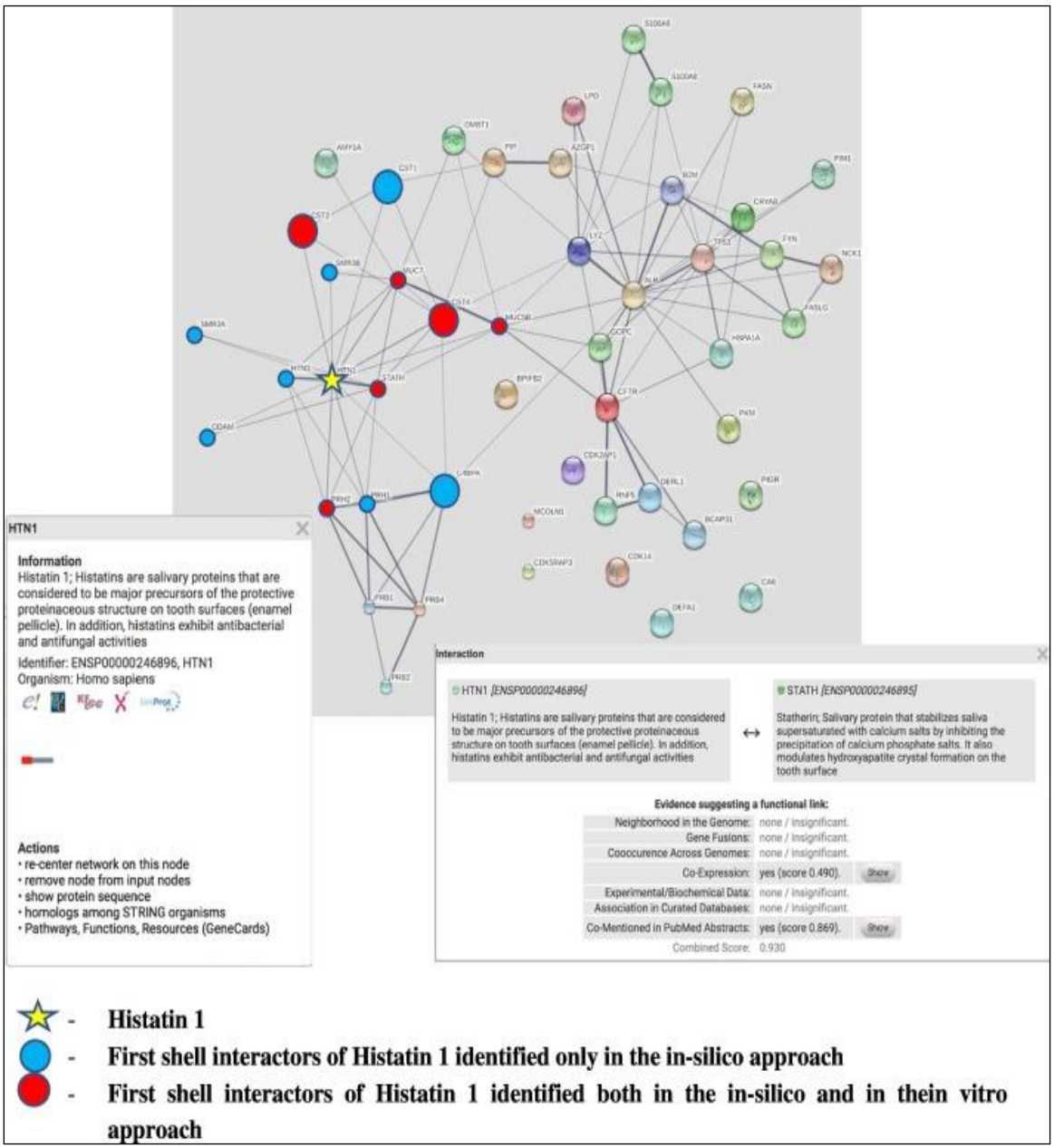



**Figure 2.4 Inclusive protein-protein network simulation merging *in-silico* and *in vitro* identified histatin 1 interactors, using STRING database. Note. Connecting lines represent interactions with at least medium confidence score (0.400). The thickness of the line linking two nodes that represent the interacting proteins varies based on the combined confidence score for the protein association. Higher combined confidence scores are represented by thicker lines.**

Evaluation of the combined simulated histatin 1 protein network showed that 6 proteins identified by the in-silico approach were confirmed in the combined approach demonstrating a strong correlation among the in-silico and the *in vitro* methods, nodes marked red in Figure 2.4. Also, 4 out of the 7 first shell proteins that were identified only in-silico, marked blue in Figure 2.4, appeared to interact with proteins listed in the *in vitro* approach showing the complementation between the two methods.

Under **ANALYSIS**, a summary containing a descriptive evaluation of the proteins present in your network is provided. Elements are listed according to the biological processes that they are involved in, their molecular function, cellular component, KEGG pathways, PFAM protein domains, and INTERPRO protein domains and features. By selecting any of the listed attributes, the proteins that display such characteristics are marked in red. In the example shown in Figure 2.5, all proteins related to salivary secretion were marked.

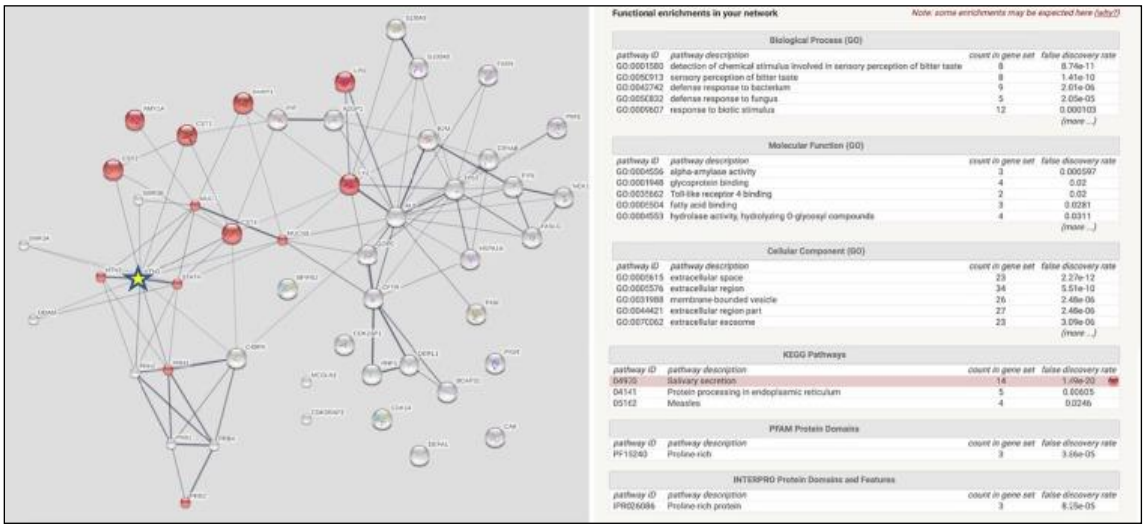



**Figure 2.5 Functional classification of the enrichments in the merged histatin 1 network where proteins related to salivary secretion are highlighted in red.**

## 2.4 Discussion

According to the PathGuide resource (http://pathguide.org), almost 300 PPI databases are available [29], one of which is The STRING protein-protein network database, which has

been maintained since the year 2000. Some qualities that distinguish STRING from other databases include its comprehensiveness, usability, quality control, and traceability [28]. Moreover, STRING covers the largest number of organisms and uses many input sources (called evidence channels) which include textmining and computational predictions. Besides, STRING's intuitive interface allows easy navigation and integration with Cytoscape software [30], thus facilitating the representation of even large-scale networks. Each interaction is also annotated with its confidence score, divided by evidence type, which is linked to the supporting literature that can be accessed by clicking on the line between the nodes. Moreover, using STRING, a basic literature search for interactors of a protein can be performed within minutes. Further, the representation of the protein network permits a simulation of the whole complex, with its many interacting shells. However, the network representation offered by STRING should be treated as a scaffold over which novel discoveries can be integrated for visualization by merging the already publicly available *in-silico* identification and the *in vitro/in vivo* novel identified observations.

In this chapter, we provided an overview of the many functions of the STRING protein network database which can be used to study PPIs. First, we simulated an in-silico identification with known and predicted first shell interactors of histatin 1. Second, using data from a previous publication, a network was created using all the proteins that were identified *in vitro* to interact with histatin 1 in saliva to form heterotypic complexes. Third, a new histatin 1-protein network was designed by combining elements that were identified by the in-silico approach with those identified by our *in vitro* experiments.

In the in-silico identification of first shell interactors with histatin 1, we emphasized the importance of increasing the maximum number of first shell interactors to assure inclusiveness. Using histatin 1 as a sample, we demonstrated that setting the maximum number of first shell interactors to 10 was not sufficient to achieve a complete representation at a medium confidence score. Moreover, by increasing the number of interactors to be displayed, three proteins were added to the network representation. Therefore, when building a comprehensive representation, the viewer should expand the number of first or second shell interactors as needed. In addition, lowering the confidence score can also be used to expand the number of interactors, as shown in the network

simulation using the complex partners of histatin 1 identified *in vitro*. Lowering the confidence score of the connections included in the network can also help increase the number of projected interactions due to genome-based prediction channels and other evidence channels used to predict associations such as textmining, co-expression, and protein homology. However, even though the last method results in high coverage, it often produces an increased number of poorly substantiated or false-positive interactions. This approach can provide insights for further exploration on yet to be experimentally confirmed protein-protein relations. Therefore, we advise caution when altering confidence levels and suggest that comparisons between networks should always consider this factor. On the other hand, lowering the confidence score of a network will help expand possibilities in the exploration to verify *ad hoc* predictions inferred from the STRING simulation. In contrast, increasing the combined confidence score is a good approach to treating noisy networks by filtering the most reliable data. Data can also be filtered or pruned by selectively reducing the number of evidence channels to be included in the network.

Another very interesting feature provided in the STRING network simulation is found in the **ANALYSIS** menu. There, all constituents of the network are classified based on their biological characteristics. We exemplified this utility in the merged in-silico/*in vitro* histatin 1 network, where the proteins associated with salivary gland secretion were marked in red. The viewer can filter the results using any of the listed aspects. The summarized biological information about all network constituents allows a fast characterization and qualitative evaluation of the group of proteins being analyzed, a very important tool for the evaluation of enriched samples. In addition, establishing biological similarities among protein-protein associations is also extremely important in research related to diseases. In the presented combined final network, the user can easily highlight the four proteins identified that participate in the process related to measles (Cellular tumor antigen p53, Tyrosine-protein kinase Fyn, Heat Shock 70 kDa protein 1A/1B, Tumor necrosis factor ligand superfamily member 6). This observation is well-aligned with the idea of using saliva for the diagnostic of the disease via salivary measles IgM and Measles Virus RNA very early after the onset of symptoms [31,32]. However, because we selected Homo sapiens

as searched organism in this representation, the proteins marked as related to measles were host proteins, not virus proteins.

Finally, by combining the in-silico identified first shell interactors of histatin 1 with histatin 1's *in vitro* identified heterotypic complex partners, a merged simulation of the inclusive histatin 1 interactome was achieved using STRING database. In the resultant merged network, the co-occurrence of 6 listed first shell interactors in both in-silico and *in vitro* approaches were used to validate our findings. Interestingly, 3 of the other 7 first shell proteins that were identified exclusively in-silico, presented association with other proteins found in the *in vitro* experiments, demonstrating the interconnectivity between the two approaches. This achieved comprehensive network strengthens the importance of combining approaches to bridge the gap that still exists in the knowledge about PPIs and suggests STRING's importance as one of the databases of choice for the study of protein-protein associations and for protein complex visual representation.

## 2.5 Conclusion

Due to the increasing complexity of functional associations among proteins, databases have become essential instruments in the study of PPIs. Huge benefits unfold from the correct use of these fantastic bioinformatics helpers. The STRING database fulfills the actual need for a tool that can collect and integrate data about known and predicted protein-protein associations from many organisms, including both direct (physical) and indirect (functional) interactions, in an easy-to-use interface.

## 2.6 Acknowledgment

Canadian Institutes of Health Research (CIHR grants # 106657 and 97577). W.L.S. is a recipient of a CIHR New Investigator Award (grant # 113166). K.T.B.C. holds a Queen Elizabeth II Graduate Scholarship in Science and Technology (QEII-GSST).

## 2.7 References


1. Dannenfelser, R., Clark, N. R. & Ma'ayan, A. Genes2FANs: Connecting genes through functional association networks. *BMC Bioinformatics* **13**, (2012).

2. de Las Rivas, J. & de Luis, A. Interactome data and databases: Different types of protein interaction. *Comparative and Functional Genomics* vol. 5 (2004).

3. Enright, A. J. & Ouzounis, C. A. Functional associations of proteins in entire genomes by means of exhaustive detection of gene fusions. *Genome Biol* **2**, (2001).

4. Rives, A. W. & Galitski, T. Modular organization of cellular networks. *Proc Natl Acad Sci U S A* **100**, (2003).

5. Snel, B., Bork, P. & Huynen, M. A. The identification of functional modules from the genomic association of genes. *Proc Natl Acad Sci U S A* **99**, (2002).

6. Studham, M. E., Tjärnberg, A., Nordling, T. E. M., Nelander, S. & Sonnhammer, E. L. L. Functional association networks as priors for gene regulatory network inference. *Bioinformatics* **30**, (2014).

7. D204-D212. UniProt: a hub for protein information The UniProt Consortium. *Nucleic Acids Res* **43**, (2015).

8. Orchard, S. *et al.* Protein interaction data curation: The International Molecular Exchange (IMEx) consortium. *Nat Methods* **9**, 345–350 (2012).

9. Orchard, S. *et al.* The MIntAct project - IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res* **42**, (2014).

10. Chatr-Aryamontri, A. *et al.* The BioGRID interaction database: 2015 update. *Nucleic Acids Res* **43**, D470–D478 (2015).

11. Das, J. & Yu, H. HINT: High-quality protein interactomes and their applications in understanding human disease. *BMC Syst Biol* **6**, (2012).

12. Turner, B. *et al.* iRefWeb: interactive analysis of consolidated protein interaction data and their supporting evidence. *Database (Oxford)* **2010**, (2010).

13. Alonso-López, D. *et al.* APID interactomes: Providing proteome-based interactomes with controlled quality for multiple species and derived networks. *Nucleic Acids Res* **44**, W529–W535 (2016).

14. Zuberi, K. *et al.* GeneMANIA prediction server 2013 update. *Nucleic Acids Res* **41**, (2013).

15. Wong, A. K., Krishnan, A., Yao, V., Tadych, A. & Troyanskaya, O. G. IMP 2.0: A multi-species functional genomics portal for integration, visualization and prediction of protein functions and networks. *Nucleic Acids Res* **43**, W128–W133 (2015).

16. Kotlyar, M., Pastrello, C., Sheahan, N. & Jurisica, I. Integrated interactions database: Tissue-specific view of the human and model organism interactomes. *Nucleic Acids Res* **44**, D536–D541 (2016).

17. Lee, I., Blom, U. M., Wang, P. I., Shim, J. E. & Marcotte, E. M. Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Res* **21**, 1109–1121 (2011).

18. Schmitt, T., Ogris, C. & Sonnhammer, E. L. L. FunCoup 3.0: Database of genome-wide functional coupling networks. *Nucleic Acids Res* **42**, (2014).

19. Snel, B., Lehmann, G., Bork, P. & Huynen, M. A. String: A web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res* **28**, (2000).

20. Szklarczyk, D. *et al.* The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* **39**, (2011).

21. Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F. & Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an Online catalog of human genes and genetic disorders. *Nucleic Acids Res* **43**, D789–D798 (2015).

22. Cherry, J. M. *et al.* Saccharomyces Genome Database: The genomics resource of budding yeast. *Nucleic Acids Res* **40**, (2012).

23. Szklarczyk, D. *et al.* STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43**, D447–D452 (2015).

24. Franceschini, A., Lin, J., von Mering, C. & Jensen, L. J. SVD-phy: Improved prediction of protein functional associations through singular value decomposition of phylogenetic profiles. *Bioinformatics* **32**, 1085–1087 (2016).

25. Walhout, A. *et al.* The C. elegans protein interaction mapping project: a test-case using proteins involved in vulval development. *Science (1979)* **287**, (1999).

26. Yu, H. *et al.* Annotation transfer between genomes: Protein-protein interrologs and protein-DNA regulogs. *Genome Res* **14**, 1107–1118 (2004).

27. Siqueira, W. L., Lee, Y. H., Xiao, Y., Held, K. & Wong, W. Identification and characterization of histatin 1 salivary complexes by using mass spectrometry. *Proteomics* **12**, 3426–3435 (2012).

28. Szklarczyk, D. *et al.* The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**, D362–D368 (2017).

29. Bader, G. D., Cary, M. P. & Sander, C. Pathguide: a pathway resource list. *Nucleic Acids Res* **34**, (2006).

30. Cline, M. S. *et al.* Integration of biological networks and gene expression data using cytoscape. *Nat Protoc* **2**, 2366–2382 (2007).

31. Centers for Disease Control and Prevention (CDC). Recommendations from an ad hoc Meeting of the WHO Measles and Rubella Laboratory Network (LabNet) on use of alternative diagnostic samples for measles and rubella surveillance. *MMWR Morb Mortal Wkly Rep* **57**, (2008).

32. Hutse, V. *et al.* Oral fluid for the serological and molecular diagnosis of measles. *International Journal of Infectious Diseases* **14**, (2010).

# Chapter 3

# 3 Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches

This chapter has been adapted from the publication titled "Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches" published in *BioMed Research International* in 2018 by Karla Tonelli Bicalho Crosara, David Zuanazzi, Eduardo Buozi Moffa, Yizhi Xiao, Maria Aparecida de Andrade Moreira Machado, and Walter Luiz Siqueira.

## 3.1 Abstract

Understanding proteins present in saliva and their function when isolated is not enough to describe their real role in the mouth. Due to protein-protein interactions (PPIs), structural changes may occur in macromolecules leading to functional modulation or modification. Besides amylase's function in carbohydrate breakdown, amylase can delay the proteolytic degradation of protein partners (e.g., histatin 1) when complexed. Due to its biochemical characteristics and high abundance in saliva, amylase probably interacts with several proteins acting as a biological carrier. This study focused on identifying interactions between amylase and other proteins found in whole saliva (WS) using proteomic approaches. Affinity chromatography (AC) was used, followed by gel electrophoresis methods, sodium dodecyl sulfate (SDS) and native, tryptic in-solution and in-gel digestion, and mass spectrometry (MS). We identified 66 proteins that interact with amylase in WS. Characterization of the identified proteins suggests that acidic (pI < 6.8) and low molecular weight (MW < 56 kDa) proteins have a preference during amylase complex formation. Most of the identified proteins present biological functions related to host protection. A new protein-amylase network was constructed using the STRING database. Further studies are necessary to investigate the individualities of the identified amylase interactors. These observations open avenues for more comprehensive studies on the not yet fully characterized biological function of amylase.

## 3.2 Introduction

WS is a complex solution that results from secretions from major and minor salivary glands, oral mucosa cells, microorganisms, and elements from the plasma, which reach saliva via gingival crevicular fluid [1]. WS participates in different mechanisms related to the processing of food, the protection of hard and soft oral tissue, and the oral microorganisms' homeostasis [2]. In fact, most of the functions attributed to WS are executed by the salivary proteins [2]. An example of the functionality of salivary proteins is the formation of the Acquire Enamel Pellicle (AEP), a protein layer formed mainly by salivary proteins with a higher affinity for hydroxyapatite [3]. Primarily, the AEP works as a physical and chemical barrier that protects the teeth. However, oral microorganisms also use the AEP as a platform to selectively adhere to the tooth surface leading to the formation of the oral biofilm (dental plaque) [3–8].

The presence of the oral biofilm is determinant for the development of the two most prevalent oral diseases: dental caries and periodontal disease. These diseases are the result of an unbalanced situation regarding the host's ability, in part provided by the salivary proteins, to control the growth of pathogenic oral bacteria when compared with the presence of indigenous microorganisms [9].

Several salivary proteins have been explored as key factors for the development of oral diseases based on biofilm formation [7,10–16]. For example, carbonic anhydrase VI has been investigated as a potential modulator for dental caries progression [14,15]. This protein is involved in the maintenance of the salivary physiological pH, by the bicarbonate buffer system, and in the neutralization of acid produced by cariogenic microorganisms present in the biofilm [14,17]. It has been suggested that reduced abundance or activity of carbonic anhydrase VI could be associated with a higher risk to develop dental caries [15,18,19].

Salivary amylase is another protein with a potential correlation with oral diseases. Amylase is the most abundant protein found in human saliva. Amylase is also present in the secretion of mammary and lacrimal glands [20]. Despite the vast literature on salivary amylase, the main function of salivary amylase as an efficient initiator of food digestion in the oral cavity is still debatable [21,22]. Mechanisms that associate salivary amylase with the clearance

of microorganisms from the oral cavity [22–24] and participation in the formation of the AEP [3,22] and in the modulation of the oral biofilm via bacteria adhesion [7,23–25] are well explained if considering the protein isolated. However, studies on the salivary proteome have indicated that understanding the individual proteins present in saliva, as well as how they function when isolated, is not enough to describe their real role when in the oral cavity. In fact, most proteins interact with other proteins originating protein complexes. Such interactions may cause structural changes in the macromolecule leading to the modulation or modification of the original individual function of the protein. For instance, when the *in vivo* identified amylase-histatin 1 complex was tested *in vitro*, amylase maintained its enzymatic activity on the hydrolysis of starch, while histatin 1 showed reduced killing activity against *Candida albicans* [26]. Also, it was shown that the lifetime of histatin 1, when complexed with amylase, was significantly increased when exposed to WS [26]. The observation that amylase can delay the proteolytic degradation of salivary protein partners when complexed suggests that this salivary protein may behave as an ideal carrier for important proteins throughout the oral cavity while maintaining their integrity [26–28].

Heterotypic complexes in saliva between amylase and MUC 5B [28], MUC 7 [29], histatin 1 [26], and histatin 5 [27] have been previously described. Due to the biochemical characteristics and abundance of amylase in saliva, it is very likely that amylase interacts with several other proteins forming complexes. The objective of our study was to reveal the interactions among amylase and other salivary proteins in WS. Comprehensive identification of *in vivo* salivary amylase complexes opens new avenues for further studies related to potential protein degradation stability and how these physiological complexes can be translated to an emerging area related to protein/peptide protection and delivery in a target area.

## 3.3 Materials and Methods

### 3.3.1 Ethics Approval for Human Participants

This research was approved by the Research Human Ethics Board of the University of Western Ontario (review number 16181E).

### 3.3.2 Collection of Whole Saliva and Preparation of Sample Pools

Stimulated saliva samples were collected from three healthy, nonsmoking adult volunteers, ranging in age from 38 to 42 years (one male and two females). All volunteers exhibited good oral health and overall good systemic conditions. The collection of WS was done between 10:00 AM and 11:00 AM, to reduce the effects of the circadian cycle. Volunteers chewed on a 5 × 5 cm piece of parafilm until 7 mL of saliva was reached. Centrifugation at 14000 ×g for 20 min at 4°C was used to separate the pellet and the WS supernatant (WSS). Only WSS were pooled together. Pellets were discarded. Each pool was made with 5 mL of WSS from each volunteer. Three pools were prepared, on different dates. The detailed scheme is shown in Figure 3.1. Saliva was used fresh for all experiments and was kept on ice from collection to the preparation of aliquots [30]. No protease inhibitors were added to the saliva samples.

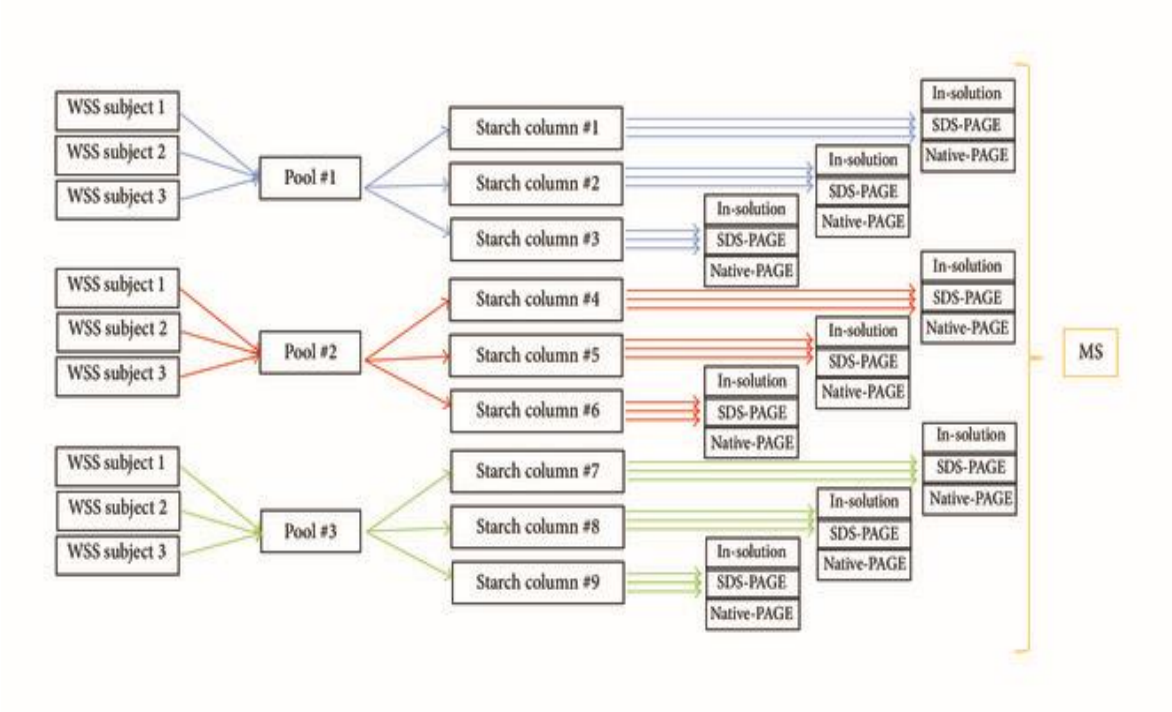



**Figure 3.1 Schematic representation of the adopted methodology.**

### 3.3.3 Separation of Amylase Complex from WSS Using an In-House Affinity Chromatography

Affinity Chromatography (AC) was employed to enrich amylase when complexed with its protein partners. Potato starch (Acros Organics, New Jersey, USA) was used as a ligand and amylase as a target. The used in-house AC method was designed and optimized, inspired by previous study [31]. A sample of 1 mL of pooled WSS was submitted to the column containing 700 $\mu$g of starch and hand-pressed slowly, the column was washed with distilled water, and amylase and its complex partners enriched solution was eluted with 1 mL of 0.1% trifluoroacetic acid (TFA). The eluate was subjected to bicinchoninic acid assay (BCA) (Pierce Chemical, Rockford, USA) for total protein concentration measurement. Bovine serum albumin was used as the protein standard. Aliquots of 20 $\mu$g protein were prepared and subjected to further separation and characterization. Following the enrichment with AC, the amylase-enriched samples were subjected to three distinctive methods: (1) in-solution tryptic digestion, (2) further separation in SDS-PAGE and in-gel tryptic digestion, and (3) confirmation of the complex formation by molecular mobility in the native-PAGE and in-gel tryptic digestion of the amylase complex.

### 3.3.4 In-Solution Digestion

Aliquots of 20 $\mu$g of total protein each were denatured and reduced by the addition of 50 $\mu$L of 4 M urea, 10 mM DTT, and 50 mM $NH_4HCO_3$, pH 7.8, and incubated for 1 hour at room temperature (RT). The solution was diluted with the addition of 150 $\mu$L of 50 mM $NH_4HCO_3$, pH 7.8. After tryptic digestion, carried out for at least 16 hours, at 37°C, with 2% w/w sequencing-grade trypsin (Promega, Madison, WI, USA), samples were desalted (Zip Tip C-18, EMD Millipore Inc., Germany) and submitted to mass spectrometric analysis (LC-ESI-MS/MS).

### 3.3.5 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) Separation Followed by In-Gel Digestion

SDS-PAGE was used to separate our protein mixture based on the individual molecular weight (MW) of our proteins. Before loading in the 12% SDS-PAGE, all samples were resuspended in 20 $\mu$L of sample buffer (0.4 M Tris-HCl pH 6.8, 4% SDS, 10% glycerol,

0.4% bromophenol blue, and 2% 2-mercaptoethanol) and boiled for 5 min. Each sample was loaded in a separate well. From left to right, the first well was loaded with 5 $\mu$L of protein standard (Precision Plus Protein™ All Blue Prestained Protein Standards, Bio-Rad, California, USA), the second well was loaded with a sample from our original solution (WSS), the third well showed the amylase-depleted saliva, the fourth well had a sample from the wash of the column, and the last well was loaded with an aliquot containing our amylase-enriched solution (amylase recovered from the starch column along with its partners) (Figure 3.2 (a)). The voltage was kept constant at 100 V during electrophoresis. Immediately after the run, all gels were stained with Coomassie Blue (40% methanol, 10% acetic acid, and 2 g Coomassie Blue) overnight with shaking at RT. Destaining was done the following morning (40% methanol, 10% acetic acid), for 1 hour and 30 min with shaking at RT. After destaining protocol, the gels were kept in Milli-Q water until scanning.

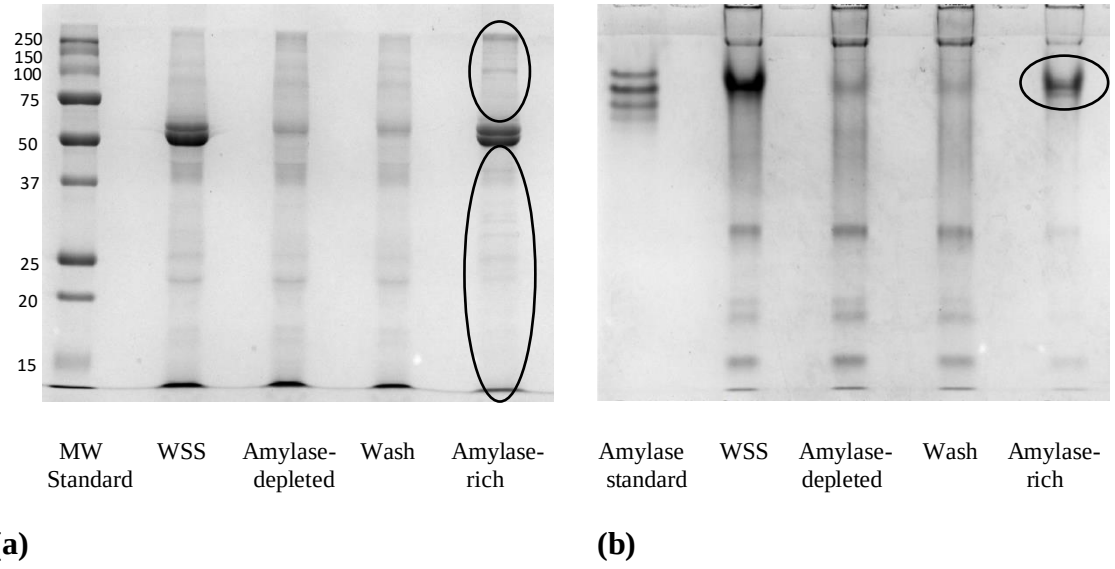



**Figure 3.2 (a) SDS-PAGE 12% and (b) native-PAGE 8% showing areas of interest for identification of protein partners of amylase. The potential protein partners of amylase are expected to be found in the areas marked with an oval shape.**

Using a razor blade, regions containing potential partners of amylase were excised from the gels. In the SDS-PAGE, the partners are expected to be found dispersed in the entire lane representing the "amylase-enriched" solution (Figure 3.2 (a)). Each lane was separated

into six band regions, and a template was used to ensure that the spots from all gels were extracted at the same MW range. After placement in separate polycarbonate tubes, each band region was cut into approximately 1 × 1 mm pieces. Gel pieces were then destained using 25 mM $NH_4HCO_3$ in 50% acetonitrile (ACN), shrunk with 100% ACN, and subjected to in-gel tryptic digestion. The digestion was carried out in 25 mM ammonium bicarbonate solution containing 0.01 $\mu$g/$\mu$L sequencing-grade trypsin (Promega, Madison, WI), for 16 hours at 37°C. Peptide extraction was achieved. Samples were desalted (Zip Tip C-18, EMD Millipore Inc., Germany) before MS.

### 3.3.6 Native-PAGE and In-Gel Digestion

A native gel was used to ensure that amylase would run still complexed with its protein partners. For the 8% native-PAGE, after resuspending the samples of 20 $\mu$g of protein with 20 $\mu$L of sample buffer (0.4 M Tris-HCl pH 6.8, 10% glycerol, and 0.4% bromophenol blue), the same order used in the SDS-PAGE was observed when loading the samples into the wells from left to right (Figure 3.2 (b)). Native-PAGE running buffer was added to the electrophoresis unit, and the voltage was kept constant at 100 V. The same staining method was used with Coomassie Blue overnight as described above. Destaining was done the following morning (40% methanol, 10% acetic acid), for 1 hour with shaking.

For the native-PAGE, the protein partners of amylase are expected to be found in the dark band corresponding to the molecular mobility of the amylase complex (Figure 3.2 (b)). Only the band about the molecular mobility of the amylase-protein complex was studied. As described above, gel bands were cut into small 1 × 1 mm pieces, destained, and subjected to in-gel tryptic digestion. The digestion was carried out in the same manner that was described for the SDS-PAGE. Peptides were recovered and samples were desalted (Zip Tip C-18, EMD Millipore Inc., Germany) before MS.

### 3.3.7 MS Analysis

Samples from all three described approaches were resuspended in 97.5% distilled water/2.4% ACN/0.1% formic acid and then subjected to RP nLC-ESI-MS/MS, using a LTQ-Velos (Thermo Scientific, San Jose, CA, USA) mass spectrometer. LC aligned with

the C18 column of capillary-fused silica (column length 10 mm, column id 75 m, 3 m spherical beads, and 100 A° pores size) was used, linked to the MS through ESI. The survey scan was set in the range of values 390–2000 MS/MS. Peptides were eluted from the nanoflow RP-HPLC over a 65 min period, with linear gradient ranging from 5 to 55% of solvent B (97.5% ACN, 0.1% formic acid), at a flow rate of 300 nL/min, with a maximum pressure of 280 bar. The electrospray voltage was 1.8 kV and the temperature of the ion-transfer capillary was 300°C. After a MS survey scan range within 390–2000 was performed and after selection of the most intense ion (parent ion), MS/MS spectra were achieved via an automated sequential selection of the seven peptides with the most intense ion for CID at 35% normalized collision energy, with the dynamic exclusion of the previously selected ions. The MS/MS spectra were matched with human protein databases (Swiss-Prot and TrEMBL, Swiss Institute of Bioinformatics, Geneva, Switzerland, https://ca.expasy.org/sprot/) using SEQUEST algorithm in Proteome Discoverer 1.3 software (Thermo Scientific, USA). The searches were performed by selecting the following SEQUEST parameters: (1) trypsin as protease enzyme, (2) 2 Da precursor ion mass tolerance, (3) 0.8 Da fragment ion mass tolerance, and (4) dynamic modifications of oxidized cysteine and methionine and phosphorylated serine and threonine. A maximum of four dynamic modifications per peptide were accepted. The SEQUEST score filter criteria applied to the MS/MS spectra for peptides were absolute XCorr threshold 0.4, fragment ion cutoff percentage 0.1, and peptide without protein XCorr threshold 1.5. Any nontryptic peptides passing the filter criteria were discarded. Only proteins for which two or more peptides were identified are reported in this study.

### 3.3.8 Identification of Protein Partners of Amylase

After MS analysis and interpretation, comparison of the common partners among the used methods allowed the construction of a list with proteins that participate with amylase in salivary complexes. Three in-house AC columns (technical triplicate) were used for each one of the 3 saliva pools (biological triplicate) prepared in different dates, making a total of 9 replicates for each one of the used approaches (in-solution digestion, SDS-PAGE followed by in-gel digestion, and native-PAGE followed by in-gel digestion). For the approach using in-solution digestion, the proteins identified by MS for the 9 replicates were

compared, and proteins that were identified in at least 2 of the replicates were listed as common proteins for this first approach. Similarly, the proteins identified by MS from the dark bands of the 9 replicates submitted to native-PAGE, followed by in-gel digestion, were compared and the proteins that appeared in at least two replicates were considered common protein partners for this second method. Last, to identify the amylase partners from the lines representing the amylase-enriched sample in the 9 replicates submitted to the SDS-PAGE approach, followed by in-gel digestion and MS, a template was used to extract the bands from all the 9 gels at approximately the same molecular weight range. The lanes with enriched samples were divided into 6 areas. Each area was analyzed separately and the 6 protein lists for each line were combined into one single protein list for each replicate; duplicate proteins were excluded. Like the other two approaches, proteins identified in at least two of the 9 replicates were deemed common for this third approach. After this triage, a Venn diagram was used to verify similarities among the common proteins listed from each described approach. The inclusion criterion for the positive identification of proteins as complex partners of amylase was that the same protein was found in at least two of the used approaches.

### 3.3.9 Bioinformatics Characterization of Amylase Complex Partners

The proteins identified in at least two of the described approaches were then characterized based on their calculated isoelectric point (pI) and molecular weight (MW). Using the physiologic salivary pH as a reference (pH 6.8), the identified proteins were grouped based on their pI (pI below and above 6.8). In addition, the MW of amylase (56 kDa) was assumed for our MW cut-off and the same proteins were divided into three groups: proteins with 0–20 kDa, proteins with 20–56 kDa, and proteins with MW above 56 kDa. Whenever available, pI and MW were calculated after the removal of the signal peptide given by the UniProt database. Otherwise, pI and MW informed in the MS report were adopted. The identified amylase complex partners were also classified based on their biological functions using data from UniProt (http://uniprot.org) assessed on August 2017. Four major groups were formed including proteins that exhibit antimicrobial activity, protection against chemical aggression, participation in host immune response and/or regulation of inflammation, and physical protection of the oral mucosa and/or wound healing.

### 3.3.10 Simulation of Amylase-Protein Network Using STRING Database

STRING database was used to provide a schematic representation of the interactions among amylase and other proteins found in the human WS as described in the publication[32] included in Chapter 2. First, a comprehensive search was performed in eight different databases (BioGRID, HPRD, APID, EMBL-EBI, FpClass, STRING, IntAct, and BioPlex) (Table 3.1) to provide a solid list with both known and predicted protein-amylase interactions. Second, a simulated amylase hub containing only the 66 proteins identified in this study was constructed using the STRING database. Last, a more inclusive network was created by merging the hub containing the proteins identified in this study with the possible partners of amylase listed in all eight searched databases. The filter was set to match with the human databank, and the confidence score was set to 0.4 (medium) in all representations.

**Table 3.1 List of proteins with known and predicted interactions with amylase, identified by search in eight databases (BioGrid, HPRD, APID, EMBL-EBI, FpClass, STRING, IntAct, and BioPlex).**

| Protein name | Database |
|---|---|
| Sucrase-isomaltase (alpha-glucosidase) | STRING |
| Amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase | STRING |
| Lactase | STRING |
| Bactericidal/permeability-increasing protein | STRING |
| Collagen, type X, alpha 1 | STRING |
| Glycogen phosphorylase, muscle form | STRING |
| Acetyl-CoA carboxylase beta (2) | STRING |
| Uridine monophosphate synthetase | STRING |
| Acetyl-CoA carboxylase alpha | STRING |
| S-phase kinase-associated protein 1 | STRING |

| | |
|---|---|
| Glucan (1,4-alpha-), branching enzyme 1 | STRING |
| Glycogen phosphorylase, liver form | STRING |
| Glycogen phosphorylase, brain form | STRING |
| A kinase (PRKA) anchor protein 8 | HPRD, BioGRID, APID, EMBL-EBI, FpClass, STRING |
| Cyclin-dependent kinase 2 associated protein 1 | BioGRID, String |
| A kinase (PRKA) anchor protein 1 | HPRD, BioGRID, APID, EMBL-EBI, STRING |
| Mucin 5B, oligomeric mucus/gel-forming | BioGRID, APID, EMBL-EBI |
| Ras association (RalGDS/AF-6) domain family member 6 | HPRD, BioGRID, APID, EMBL-EBI |
| Putative oral cancer supressor, deleted in oral cancer 1 | IntAct, HPRD, APID, EMBL-EBI |
| Superoxide dismutase (Mn), mitochondrial | IntAct, BioGRID, APID, EMBL-EBI |
| Uncoupling protein 2 (mitochondrial, proton carrier) | IntAct, APID, EMBL-EBI |
| ARP8 actin-related protein 8 homolog (yeast) | BioGRID, BioPlex, APID, EMBL-EBI |
| Beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group) | BioGRID, BioPlex, APID, EMBL-EBI |
| General transcription factor IIB | BioGRID, BioPlex, APID, EMBL-EBI |
| Killer cell immunoglobulin-like receptor, two domains, short cytoplasmatic tail, 2 | BioGRID, BioPlex, APID, EMBL-EBI |
| Mab-21-like 1 (C. elegans) | BioGRID, BioPlex, APID, EMBL-EBI |
| Starch binding domain 1 | BioGRID, BioPlex, APID, EMBL-EBI |
| Protein kinase, cAMP-dependent, regulatory subunit, type II, beta | BioGRID, BioPlex, APID, EMBL-EBI |
| Peptide (mitochondrial processing) beta | BioGRID, BioPlex, APID, EMBL-EBI |
| Trafficking protein particle complex 12 | BioGRID, BioPlex, APID, EMBL-EBI |
| Ubiquitin-like 7 | BioGRID, BioPlex, APID, EMBL-EBI |
| calcium/calmodulin-dependent protein kinase ID | BioGRID, BioPlex, APID, EMBL-EBI |
| Receptor-interacting serine-theorine kinase 3 | BioGRID, BioPlex, APID, EMBL-EBI |
| Chemokine (C-C motif) receptor-like 2 | BioGRID, BioPlex, APID, EMBL-EBI |

| | |
|---|---|
| Pleckstrin homology domain containing, family G (with RhoGef domain) member 6 | BioPlex, APID |
| Regulator of calcineurin 1 | BioPlex, APID |
| Vasohibin 1 | BioPlex, APID |
| Gastrokine 1 | BioPlex, APID |
| Zinc-finger, B-box domain containing | BioPlex, APID |
| DEAD (Asp-Glu-Ala-Asp) box helicase 17 | BioPlex, APID |
| Rho-related BTB domain containing 1 | BioPlex, APID |
| Ts translation elongation factor, mitochondrial | BioPlex, APID |
| Tumor necrosis factor receptor superfamily, member 19 | BioPlex, APID |
| Neuropeptide B | BioPlex, APID |
| Forkhead box N4 | BioPlex, APID |
| FERM domain containing 1 | BioPlex, APID |
| WD repeat domain 6 | BioPlex, APID |
| DNA replication licensing factor MCM2 | FpClass |
| Probable ATP-dependent RNA helicase DDX5 | FpClass |
| G1/S-specific cyclin-D3 | FpClass |
| cAMP-dependent protein kinase type II-alpha reulatory subunit | FpClass |
| Histatin 1 | FpClass |
| Salivary acidic proline-rich phosphoprotein 1/2 | FpClass |
| Statherin | FpClass |

Websites: STRING: http://www.string-db.org; HPRD: http://www.hprd.org; BioGRID: http://www.thebiogrid.org; APID: http://www.apid.dep.usal.es; EMBL-EBI: http://www.ebi.ac.uk; FpClass: http://www.dcv.uhnres.utoronto.ca; IntAct: http://www.ebi.ac.uk/intact; BioPlex: http://www.bioplex.hms.harvard.edu

## 3.4 Results

The selectivity of our in-house AC starch columns towards amylase is demonstrated in Figure 3.2, where the band related to amylase's MW (56 kDa) in Figure 3.2 (a) and the amylase complex in Figure 3.2 (b) practically disappears in the lines representing saliva depleted from amylase and the wash with distilled water. On the other hand, dark bands are seen in the corresponding areas with the amylase-enriched samples. Although slight bands can be seen in areas besides that of the amylase complex in Figure 3.2 (b) where the enriched sample is represented, such faded bands might be related to proteins that either show weak interaction with the complex which was disrupted during the processing of samples or may be related to "contaminants" that remained in the column after wash. To ensure precise identification of proteins from the complex, only the proteins listed in the dark band in the native-PAGE (Figure 3.2 (b)) were considered.

The data obtained after LC-ESI-MS/MS analysis of samples from the three described approaches identified 66 different proteins found in WS that form complexes with salivary amylase. All identified proteins are listed in Table 3.2, along with the corresponding approach used for the identification, protein MW and pI.

**Table 3.2 List of all identified potential amylase protein partners according to the used proteomic approach.**

| Accession Number | Protein Name | In-solution | In-gel (SDS-PAGE) | In-gel (Native-PAGE) | MW (KDa) | calc. pI |
|---|---|---|---|---|---|---|
| C0JYZ2 | Titin | x | x | x | 3340.16 | 6.09 |
| B4E1M1 | cDNA FLJ60391, highly similar to Lactoperoxidase | x | x | x | 73.88 | 8.15 |
| Q9HC84 | Mucin-5B | x | x | x | 593.84 | 6.20 |
| P04080 | Cystatin-B | x | x | x | 11.14 | 6.96 |
| B4DVQ0 | cDNA FLJ58286, highly similar to Actin, cytoplasmic 2 | x | x | x | 37.30 | 5.71 |
| P01037 | Cystatin-SN | x | x | x | 14.32 | 6.92 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Q6PJF2 | IGK@ protein | **x** | **x** | **x** | 23.32 | 6.98 |
| Q0QET7 | Glyceraldehyde-3-phosphate dehydrogenase (Fragment) | **x** | **x** | **x** | 24.60 | 8.51 |
| A0A075B6K9 | Ig lambda-2 chain C regions (Fragment) | **x** | **x** | **x** | 11.30 | 7.24 |
| P05109 | Protein S100-A8 | **x** | **x** | **x** | 10.70 | 6.57 |
| P12273 | Prolactin-inducible protein | **x** | **x** | **x** | 13.52 | 5.40 |
| Q96DR5 | BPI fold-containing family A member 2 | **x** | **x** | **x** | 25.05 | 5.19 |
| A0A0C4DGN4 | Zymogen granule protein 16 homolog B | **x** | **x** | **x** | 17.21 | 5.39 |
| Q9UGM3 | Deleted in malignant brain tumors 1 protein | **x** | **x** | **x** | 258.66 | 5.19 |
| P01833 | Polymeric immunoglobulin receptor | **x** | **x** | **x** | 81.35 | 5.59 |
| P01876 | Ig alpha-1 chain C region | **x** | **x** | **x** | 37.66 | 6.51 |
| P23280 | Carbonic anhydrase VI | **x** | **x** | **x** | 33.57 | 6.41 |
| C8C504 | Beta-globin | **x** | **x** | **x** | 15.87 | 7.98 |
| A7Y9J9 | Mucin 5AC, oligomeric mucus/gel-forming | **x** | **x** | **x** | 645.90 | 6.27 |
| P01834 | Ig kappa chain C region | **x** | **x** | **x** | 11.60 | 5.87 |
| H6VRF8 | Keratin 1 | **x** | **x** | **x** | 66.00 | 8.12 |
| P13645 | Keratin, type I cytoskeletal 10 | **x** | **x** | **x** | 58.83 | 5.13 |
| P01036 | Cystatin-S | **x** | **x** | **x** | 14.19 | 4.83 |
| B2R4M6 | Protein S100 | **x** | **x** | **x** | 4.31 | 4.55 |
| P35908 | Keratin, type II cytoskeletal 2 epidermal | **x** | **x** | **x** | 65.43 | 8.07 |
| B1APF8 | cAMP-dependent protein kinase catalytic subunit beta (Fragment) | **x** | **x** | **x** | 20.56 | 9.56 |
| B5ME49 | Mucin-16 | **x** | **x** | **x** | 1519.17 | 5.13 |
| P25311 | Zinc-alpha-2-glycoprotein | **x** | **x** | | 32.14 | 5.58 |
| F6KPG5 | Albumin (Fragment) | **x** | **x** | | 66.49 | 6.04 |
| B2R7Z6 | cDNA, FLJ93674 | **x** | **x** | | 50.34 | 7.05 |

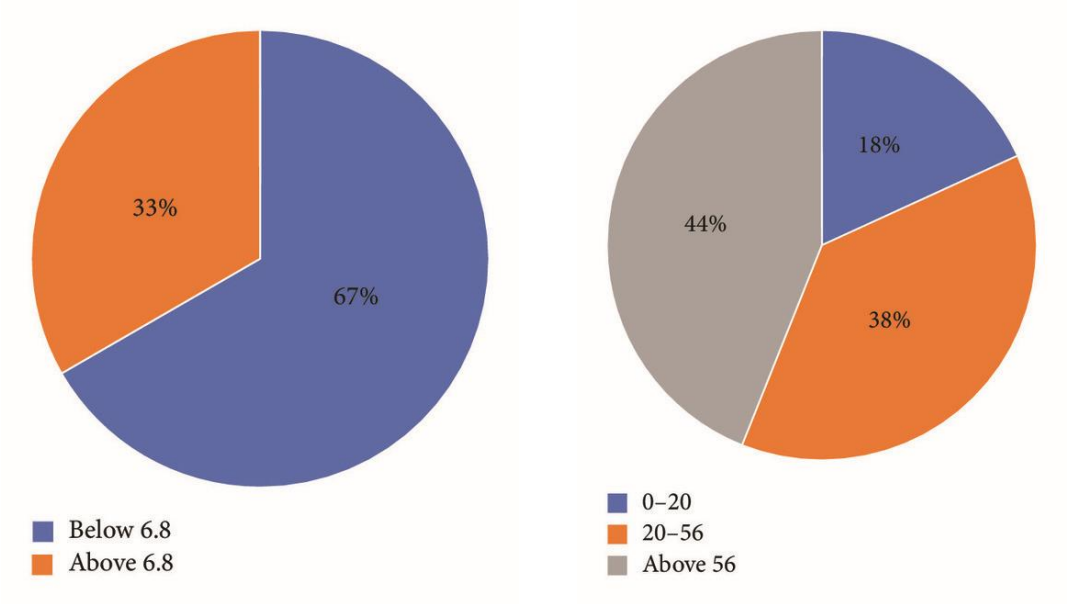
| | | | | | | |
|---|---|---|---|---|---|---|
| E9PKG6 | Nucleobindin-2 | **x** | **x** | | 37.50 | 5.01 |
| P02647 | Apolipoprotein A-I | **x** | **x** | | 28.08 | 5.27 |
| Q9Y6V0 | Protein piccolo | **x** | **x** | | 553.28 | 6.09 |
| P09228 | Cystatin-SA | **x** | **x** | | 14.35 | 4.85 |
| A0A024R9Y3 | HECT, UBA and WWE domain containing 1, isoform CRA_a | **x** | **x** | | 479.90 | 5.21 |
| E7ETD6 | Nucleosome-remodeling factor subunit BPTF | **x** | **x** | | 307.90 | 6.04 |
| Q8TAX7 | Mucin-7 | **x** | **x** | | 36.81 | 9.30 |
| P06733 | Alpha-enolase | **x** | **x** | | 47.04 | 6.99 |
| P10599 | Thioredoxin | **x** | **x** | | 11.61 | 4.82 |
| Q9UPN3 | Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5 | **x** | **x** | | 838.31 | 5.28 |
| O95661 | GTP-binding protein Di-Ras3 | **x** | | **x** | 25.50 | 9.46 |
| A7E2D6 | NAV2 protein | **x** | | **x** | 261.56 | 8.98 |
| G3CIG0 | MUC19 variant 12 | **x** | | **x** | 802.68 | 4.96 |
| Q8N4F0 | BPI fold-containing family B member 2 | | **x** | **x** | 47.13 | 8.48 |
| P01024 | Complement C3 | | **x** | **x** | 184.95 | 6.00 |
| H7BY35 | Ryanodine receptor 2 | | **x** | **x** | 562.25 | 6.19 |
| Q07869 | Peroxisomeproliferator-activated receptor alpha | | **x** | **x** | 52.23 | 5.86 |
| B4E1T1 | cDNA FLJ54081, highly similar to Keratin, type II cytoskeletal 5 | | **x** | **x** | 58.81 | 5.97 |
| A8K2U0 | Alpha-2-macroglobulin-like protein 1 | | **x** | **x** | 159.33 | 5.50 |
| Q6P5S2 | Protein LEG1 homolog | | **x** | **x** | 35.86 | 5.79 |
| B4E3A8 | cDNA FLJ53963, highly similar to Leukocyte elastase inhibitor | | **x** | **x** | 38.69 | 6.22 |
| F8WA11 | CLIP-associating protein 1 | | **x** | **x** | 162.66 | 8.72 |
| B7ZAL5 | cDNA, FLJ79229, highly similar to Lactotransferrin | | **x** | **x** | 73.10 | 7.78 |
| P02533 | Keratin, type I cytoskeletal 14 | | **x** | **x** | 51.56 | 5.09 |

| | | | | | |
|---|---|---|---|---|---|
| B2R825 | Alpha-1,4 glucan phosphorylase | **x** | **x** | 97.01 | 7.30 |
| A0A087WWT3 | Serum albumin | **x** | **x** | 43.03 | 5.69 |
| B7Z759 | cDNA FLJ61672, highly similar to Proteoglycan-4 (Fragment) | **x** | **x** | 92.09 | 9.44 |
| J3QLC9 | Haptoglobin (Fragment) | **x** | **x** | 39.03 | 5.54 |
| P01877 | Ig alpha-2 chain C region | **x** | **x** | 36.50 | 6.10 |
| A8K739 | cDNA FLJ77339 | **x** | **x** | 24.84 | 5.06 |
| B7Z747 | cDNA FLJ51120, highly similar to Matrix metalloproteinase-9 | **x** | **x** | 64.09 | 6.42 |
| B7Z565 | cDNA FLJ54739, highly similar to Alpha-actinin-1 | **x** | **x** | 94.72 | 5.69 |
| B4DI70 | cDNA FLJ53509, highly similar to Galectin-3-binding protein | **x** | **x** | 44.37 | 5.03 |
| P35527 | Keratin, type I cytoskeletal 9 | **x** | **x** | 62.06 | 5.14 |
| P80188 | Neutrophilgelatinase-associated lipocalin | **x** | **x** | 20.55 | 9.02 |
| P04040 | Catalase | **x** | **x** | 59.62 | 6.95 |

When results from all approaches were combined, 375 different proteins were recognized. In-solution digestion provided 164 proteins that probably interact with amylase: SDS-PAGE, followed by in-gel tryptic digestion, 237 potential partners; native-PAGE, followed by in-gel digestion, 67 possible complex partners. After selecting only proteins that were identified in two or more of the used approaches, results were narrowed down to 66 proteins, where 27 proteins were detected in all three methods, besides amylase itself, and 39 other proteins were concomitantly identified in only two of the used approaches. A total of 13 unique proteins were identified using both in-solution tryptic digestion and SDS-PAGE, followed by in-gel tryptic digestion; 23 proteins were found in both PAGE approaches; and 3 proteins were uniquely found concomitantly in the samples from in-solution tryptic digestion and native gels, followed by in-gel tryptic digestion (Figure 3.3).


**Figure 3.3 Venn diagram distribution of the proteins found in each of the three proteomic approaches used in this study.**

Moreover, the 66 common proteins displayed MW ranging from 4.31 kDa to 3340.16 kDa (Table 3.2), where most of the identified amylase partners (56%) presented a MW below 56 kDa, amylase's MW (Figure 3.4 (a)). The identified proteins were also grouped based on their isoelectric points (pI). Clearly, most of the 66 proteins (67%) presented pI below 6.8. One-third (33%) of the identified amylase-protein partners exhibited basic characteristics ranging in pI above 6.8 (Figure 3.4 (b)).




**(a)** **(b)**

**Figure 3.4 Percentage distribution of the identified proteins according to the biochemical characteristics of salivary amylase (pI 6.8 and 56 kDa). (a) pI distribution of the identified proteins using salivary pI as a comparison value. (b) Molecular weight distribution of the identified proteins using salivary amylase molecular weight as the comparison value.**

Interestingly, the characterization of the 66 identified proteins based on their biological functions indicated that most of the proteins participating in complex with amylase exhibit protective roles towards the maintenance of the host's health. In fact, from the 66 identified proteins, 37 display oral defensive functions: 13 proteins have antimicrobial activities, 9 elements are capable of neutralizing chemical aggressions to the host's tissues, 10 proteins participate in mechanisms that initiate or modulate the host's immune response and inflammatory process, and 10 proteins contribute to the physical protection of the host's tissue and/or wound healing (Table 3.3).

**Table 3.3 Distribution of proteins identified to interact with salivary amylase forming complex based on their biological functions.**

| Biological function | Accession Number | Protein Name |
|---|---|---|
| **Defense response to bacterium, virus and fungus (n=13)** | Q9HC84 | Mucin-5B |
| | B4E1M1 | cDNA FLJ60391, highly similar to Lactoperoxidase |
| | P05109 * | Protein S100-A8 |
| | Q96DR5 | BPI fold-containing family A member 2 |
| | Q8N4F0 | BPI fold-containing family B member 2 |
| | Q9UGM3 | Deleted in malignant brain tumours 1 protein |
| | P01876 | Ig alpha-1 chain C region |
| | P01834 * | Ig kappa chain C region |
| | P01877 * | Ig alpha-2 chain C region |
| | B7Z759 | cDNA FLJ61672, highly similar to Proteoglycan-4 (Fragment) |
| | Q8TAX7 * | Mucin-7 |
| | B7ZAL5 | cDNA, FLJ79229, highly similar to Lactotransferrin |
| | B4DI70 | cDNA FLJ53509, highly similar to Galectin-3-binding protein |
| **Neutralization of chemical aggression (n=9)** | P04080 | Cystatin-B |
| | P01037 | Cystatin-SN |
| | P23280 | Carbonic anhydrase VI |
| | P01036 | Cystatin-S |
| | P09228 | Cystatin-SA |
| | A8K2U0 | Alpha-2-macroglobulin-like protein 1 |
| | B4E3A8 | cDNA FLJ53963, highly similar to Leukocyte elastase inhibitor |
| | P04040 | Catalase |
| | A7Y9J9 * | Mucin 5AC, oligomeric mucus/gel-forming |
| | P05109 * | Protein S100-A8 |

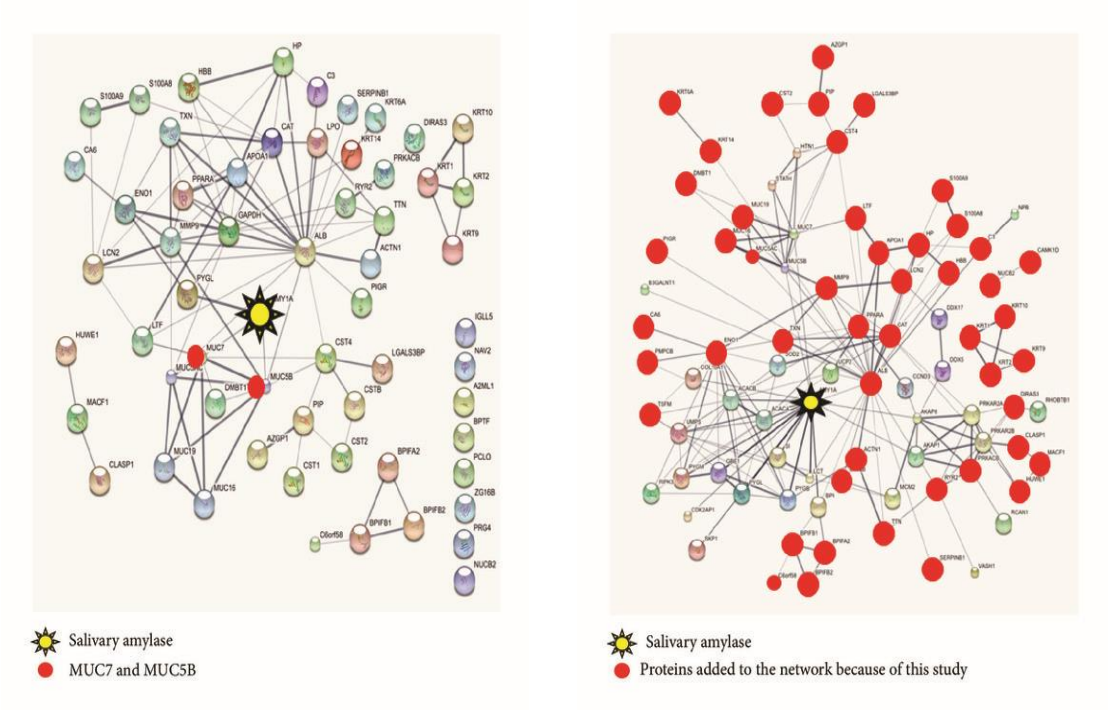
| | | |
|---|---|---|
| **Immune response and regulation of inflammation (n=10)** | P12273 | Prolactin-inducible protein |
| | P01834 * | Ig kappa chain C region |
| | B2R4M6 | Protein S100 |
| | P01024 | Complement C3 |
| | P80188 | Neutrophil gelatinase-associated lipocalin |
| | P01877 * | Ig alpha-2 chain C region |
| | A0A075B6K9 | Ig lambda-2 chain C regions (Fragment) |
| | B7Z747 | cDNA FLJ51120, highly similar to Matrix metalloproteinase-9 |
| | Q6PJF2 | IGK@ protein |
| **Mucosa protection and wound healing (n=10)** | B5ME49 | Mucin-16 |
| | P01833 | Polymeric immunoglobulin receptor |
| | P35908 | Keratin, type II cytoskeletal 2 epidermal |
| | P25311 | Zinc-alpha-2-glycoprotein |
| | Q07869 | Peroxisome proliferator-activated receptor alpha |
| | Q9UPN3 | Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5 |
| | A7Y9J9 * | Mucin 5AC, oligomeric mucus/gel-forming |
| | Q8TAX7 * | Mucin-7 |
| | P02647 | Apolipoprotein A-I |
| | Q9UPN3 | Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5 |
| **Biological functions not directly related to host protection or unknown (n=29)** | G3CIG0 | MUC19 variant 12 |
| | A0A0C4DGN4 | Zymogen granule protein 16 homolog B |
| | A0A024R9Y3 | HECT, UBA and WWE domain containing 1, isoform CRA_a |
| | E7ETD6 | Nucleosome-remodeling factor subunit BPTF |
| | P06733 | Alpha-enolase |
| | A7E2D6 | NAV2 protein |
| | H7BY35 | Ryanodine receptor 2 |
| | Q6P5S2 | Protein LEG1 homolog |

| | | |
|---|---|---|
| | F8WA11 | CLIP-associating protein 1 |
| | P02533 | Keratin, type I cytoskeletal 14 |
| | B2R825 | Alpha-1,4 glucan phosphorylase |
| | B4DVQ0 | cDNA FLJ58286, highly similar to Actin, cytoplasmic 2 |
| | Q0QET7 | Glyceraldehyde-3-phosphate dehydrogenase (Fragment) |
| | J3QLC9 | Haptoglobin (Fragment) |
| | B7Z565 | cDNA FLJ54739, highly similar to Alpha-actinin-1 |
| | C0JYZ2 | Titin |
| | C8C504 | Beta-globin |
| | H6VRF8 | Keratin 1 |
| | P13645 | Keratin, type I cytoskeletal 10 |
| | B1APF8 | cAMP-dependent protein kinase catalytic subunit beta (Fragment) |
| | F6KPG5 | Albumin (Fragment) |
| | E9PKG6 | Nucleobindin-2 |
| | Q9Y6V0 | Protein piccolo |
| | P10599 | Thioredoxin |
| | O95661 | GTP-binding protein Di-Ras3 |
| | B4E1T1 | cDNA FLJ54081, highly similar to Keratin, type II cytoskeletal 5 |
| | A0A087WWT3 | Serum albumin |
| | A8K739 | cDNA FLJ77339 |
| | P35527 | Keratin, type I cytoskeletal 9 |

*Proteins involved in more than one process.

The amylase interactome simulation using the STRING database demonstrated that not all 66 proteins were linked to the protein-amylase network (Figure 3.5 (a)). Three distinct isolated groups of 3 to 4 proteins were formed apart from the network, along with other lonely individual nodes. MUC 7 and MUC 5B were among the identified proteins. When

the newly identified amylase network was merged with the previously documented interactors, only two groups of proteins were not linked to the network (Figure 3.5 (b)). One group contained NUCB2 (nucleobindin-2) and CAMK1D (calcium/calmodulin-dependent protein kinase type 1D), and the other cluster was formed by four keratins.




**(a)** **(b)**

**Figure 3.5 In silico view of the amylase interactome using the STRING database: (a) All proteins identified herein interact with amylase forming complex where represented. (b) An inclusive amylase interactome was constructed merging the proteins identified in this study with all known (in vitro studies) and predicted proteins (in silico database) previously mentioned to form complex with amylase.**

## 3.5 Discussion

A total of 66 proteins that participate in protein complexes with amylase in WS were identified with the application of three different proteomic approaches. Initially, AC was used to enrich amylase along with its partners from the complex saliva solution. The reaction between amylase and starch is an enzymatic reversible mechanism [20], allowing

the recovery of intact amylase after its reaction with the starch. The use of a starch column was previously described as a mean for the depletion of amylase from saliva [31]. However, in this study, we demonstrated that the mentioned method enriches amylase from saliva still complexed with other proteins. This observation was first suggested by the different bands present in the SDS-PAGE when the samples eluted from the starch column were separated by MW (Figure 3.2 (a)). Later, this observation was confirmed by MS analysis of the same amylase-enriched samples where many proteins besides amylase were identified (Table 3.2). Thus, AC starch column alone should not be recommended for the depletion of amylase from saliva, unless a careful dismemberment of protein complexes can be performed earlier in ways that do not interfere with the activity of salivary amylase.

Moreover, the importance of using different methods for the identification of proteins was here demonstrated. Combining all used approaches, a total of 375 unique proteins were identified as potential members of the amylase complex. Interesting to note, SDS-PAGE was the method where the largest number of proteins was identified (237 proteins). From the 66 proteins that were identified in at least two of the used approaches and therefore are more likely to interact with amylase, only 3 proteins were not identified in the approach with SDS-PAGE. This demonstrates that additional sample separation based on the MW of each protein, together with the MS analysis of independent bands from different areas of the gel, prevents highly abundant proteins from masking or hiding low abundant ones, therefore improving the method specificity. On the other hand, while using directly in-solution tryptic digestion uniquely, 24 proteins from our final list of 66 interactors were not identified (Figure 3.3), once again reinforcing the hypothesis of high abundance proteins preventing the identification of low abundance ones unless further separation is performed before MS analysis. Also, 13 of the 66 identified proteins from the amylase complex were not identified in the native-PAGE approach. Since native gels provide a sample separation based on the molecular mobility and charge of the complex, the absence of some of the identified proteins may be a consequence of weaker bindings, thus damaging the stability of some PPIs and preventing all proteins that were originally in the complex from being identified in this method.

Amylase-PPI with histatins (histatin 1 and histatin 5) and with mucins (MUC 5B and MUC 7) were described previously [26,28,29,33]. Mucins (MUC5B and MUC7), a protein family only present in mucous glands such as submandibular and sublingual glands, were here identified among the partners of amylase in WS, confirming previous studies [28,29]. Contrarily, histatins were not identified in this study probably because of their short lifespan in the oral cavity due to protein degradation by endogenous oral proteases [26,34,35]. Protease inhibitors can be used in an attempt to prevent proteolytic degradation. However, in saliva, it has been shown that short-term storage of freshly collected saliva samples on ice is more effective in preventing proteolytic degradation, without interfering with the chemistry of the proteome, than the use of protease inhibitors [30]. Therefore, no protease inhibitors were added to the saliva samples as they could promote chemical alterations on our protein complexes leading to changes in the stability of the complex and to an incorrect identification of the proteins that participate in complexes with amylase.

To distinguish a protein profile among the identified partners of amylase, biochemical characterization was performed according to the calculated pI and MW of the proteins and to their biological functions. Using the prevailing physiological salivary pH as reference (pH 6.8), the identified proteins were divided into two groups: pI below and pI above 6.8. Most of all identified proteins (67%) presented isoelectric points below 6.8 and therefore exhibited negative charge in a solution with pH 6.8. On the other hand, one-third of the identified amylase-protein partners (33%) exhibited more basic characteristics with pI above 6.8, showing positive charges in pH 6.8. Therefore, there appears to be a preference for acidic proteins ($pI < 6.8$) to participate in the identified amylase complex. Knowing that ionic forces and hydrogen bonding, both electrostatic interactions, are involved in the formation of protein complexes, shifts in the net charge of salivary proteins possibly interfere with the nature and abundance of the proteins present in complexes. Differences among the pH of the secretions from the major salivary glands have been described [1,36]. Also, changes in the pH of saliva have been suggested as biomarkers for systemic diseases [37]. In tumors, for example, there seems to be a shift in pH towards being acidic, acting as a favorable factor for tumor cells [37]. The proposition that variations in the salivary pH might

interfere in the formation of salivary complexes suggests a new research and diagnostic avenue combining salivary proteome/interactome and salivary pH.

Since only subjects with overall good systemic and oral health were included in this study, it is implied that all our results were acquired around physiological salivary pH. In this condition, the characterization of the 66 identified proteins based on their biological functions reinforced the possible function of amylase as an important biological carrier. In total, 56% (37 proteins) of the identified partners of amylase exhibited important roles towards the maintenance of oral health. Four main mechanisms were recognized: antimicrobial activities, protection against chemical aggressions, immune response and regulation of inflammation, and physical protection of the mucosa and wound healing. About the debatable participation of amylase in the development of dental caries, this study did not aim to clarify the direct involvement of salivary amylase in the carious process. Contrarily, a new question is here proposed on the potential indirect participation of amylase in the protection against dental caries via functional modulation and/or protection of "anticariogenic" proteins from early proteolytic degradation in the oral cavity. A possible example of such proteins identified in this study is carbonic anhydrase VI. Besides carbonic anhydrase VI's involvement in taste sensation, this isoenzyme maintains the physiological salivary pH by catalyzing the hydration of carbon dioxide (bicarbonate buffer system), assisting in the recovery from acidic, more cariogenic, salivary challenges [17]. Carbonic anhydrase VI can also penetrate in the biofilm to facilitate the neutralization of acids secreted by the bacteria [14]. Carbonic anhydrase VI was identified among the proteins that participate in salivary complex with amylase. However, the direct binding of amylase and carbonic anhydrase VI and the possible consequences of such interaction are yet to be investigated. Other proteins identified in this study were cystatins B, SN, S, and SA. Cystatins are proteins that inhibit cysteine proteases secreted by the host, bacteria, and viruses [38]. Cystatins SA and SN are particularly involved in the control of the proteolytic events *in vivo* such as periodontal tissue destruction [39]. The presence of cystatins B, SN, S, and SA in salivary complexes with amylase suggests that amylase may contribute indirectly against periodontal diseases.

Open proteomics/interactomics databases have been developed to assist in the study of PPIs and to accelerate discoveries in the field. Using the STRING database, a simulation of the amylase interactome with the identified partners of amylase was constructed. Out of the 66 members of the amylase complex listed in this study, only two proteins had been previously reported in the literature to present direct interactions with amylase; they were MUC5B [28] and MUC7 [29]. No direct binding between amylase and any of the other 64 proteins identified herein has been described up till now. Therefore, additional studies are needed to determine if any of these other proteins bind directly to amylase forming the first shell of the protein complex, as well as the exact position of each of the identified members of the amylase complex in the protein-protein network. Furthermore, the creation of a second amylase-protein network merging the newly identified amylase-protein network with the previously known and predicted amylase interactors demonstrated that using in silico approach based on molecular affinity prediction and prior *in vivo* and *in vitro* experiments, most of the 66 proteins identified herein fill the gap in the amylase interactome present in WS.

It is important to highlight that the proteins identified herein in complexes with amylase, the most abundant salivary protein, were detected using three different proteomic approaches, with nine replicates, using saliva from three subjects, collected on three different dates (Supplemental Table 3.1). On the other hand, it is likely that each salivary protein has a different binding affinity with amylase. In fact, it is well known that changes in the salivary flow rate, a person's overall health, and emotional state can promote qualitative and quantitative variations in the salivary proteome [40–44] and, consequently, in the amylase interactome. Future studies need to address the amylase interactome in different physiological/pathological conditions.

In summary, this study pioneered the exploration of the vast salivary interactome. It is important to remember that some of the proteins identified herein may interact with amylase indirectly, having one or more proteins as mediators of such interactions. Unfortunately, very little is known about the dynamics of these interactions. Transient protein complexes are less likely to be identified than permanent protein complexes. Additional studies are needed to confirm how the proteins listed in this manuscript interact

with each other and with amylase. Amylase's ability to protect such partners from proteolytic degradation and/or modulate their biological functions while in the complex is yet to be studied comprehensively.

## 3.6 Conclusion

The large number of amylase complex partners identified herein reinforces the hypothesis that the real role of amylase in the oral cavity might not be related to carbohydrate digestion. Instead, amylase's most important role may be associated with protein transport and possible protection and functional modulation of its partners. In an era of more personalized and targeted medicine, this study opens the hypothesis for a novel therapeutic avenue where amylase can offer information for the development of an ideal carrier for functionally important peptides/proteins towards the prevention of oral diseases. Moreover, the salivary interactome may function as a foundation for the development of more efficient artificial saliva and/or mouthwashes and provide more reliable models to design drugs directed to amylase or dependent on its function.

## 3.7 Acknowledgments

This research was supported by the Canadian Institutes of Health Research (CIHR Grant no. 106657) and the Canada Foundation for Innovation-Leaders Opportunity Fund (CFI-LOF Grant no. 25116). Karla Tonelli Bicalho Crosara holds a Queen Elizabeth II Graduate Scholarship in Science and Technology (QEII-GSST).

## 3.8 Supplementary Materials

**Supplemental Table 3.1 Detailed information regarding the identification of the 27 proteins that were common in all three used proteomic approaches.**

| Accession Number | Protein name | Used proteomic approach | Number of replicates identified | Unique peptides (average) | Score (average) | % Coverage (average) |
|---|---|---|---|---|---|---|
| C0JYZ2 | Titin | In-solution | 8 | 8.63 | 41.76 | 0.68 |
| | | In-gel (SDS-PAGE) | 6 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 5 | 6.8 | 113.36 | 0.38 |

| | | | | | | |
|---|---|---|---|---|---|---|
| B4E1M1 | cDNA FLJ60391, highly similar to Lactoperoxidase | In-solution | 2 | 2 | 10.28 | 4.59 |
| | | In-gel (SDS-PAGE) | 4 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 4 | 4.5 | 32.26 | 8.12 |
| Q9HC84 | Mucin-5B | In-solution | 4 | 6 | 31.22 | 1.63 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 3 | 3.67 | 14.59 | 1.9 |
| P04080 | Cystatin-B | In-solution | 8 | 2.38 | 19.83 | 32.53 |
| | | In-gel (SDS-PAGE) | 8 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 2 | 2.5 | 18.82 | 35.21 |
| B4DVQ0 | cDNA FLJ58286, highly similar to Actin, cytoplasmic 2 | In-solution | 2 | 5 | 77.4 | 19.82 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 4 | 4.75 | 46.9 | 27.7 |
| P01037 | Cystatin-SN | In-solution | 2 | 1.4 | 82.19 | 32.62 |
| | | In-gel (SDS-PAGE) | 8 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 4 | 2.5 | 46.67 | 32.62 |
| Q6PJF2 | IGK@ protein | In-solution | 2 | 5.5 | 57.87 | 33.41 |
| | | In-gel (SDS-PAGE) | 5 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 2 | 5 | 117.65 | 28.94 |
| Q0QET7 | Glyceraldehyde-3-phosphate dehydrogenase (Fragment) | In-solution | 2 | 2.5 | 11.46 | 17.18 |
| | | In-gel (SDS-PAGE) | 6 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 4 | 2 | 15.98 | 12.67 |
| A0A075B6K9 | Ig lambda-2 chain C regions (Fragment) | In-solution | 4 | 2.5 | 18.89 | 36.8 |
| | | In-gel (SDS-PAGE) | 4 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 4 | 2.75 | 27.83 | 37.03 |
| P05109 | Protein S100-A8 | In-solution | 2 | 2 | 18.55 | 19.35 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 2 | 2.5 | 17.82 | 24.73 |
| P12273 | Prolactin-inducible protein | In-solution | 9 | 4.56 | 54.77 | 31.05 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | In-gel (SDS-PAGE) | 8 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 7 | 2.86 | 28.44 | 25.15 |
| Q96DR5 | BPI fold-containing family A member 2 | In-solution | 8 | 6.13 | 42.84 | 21.79 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 3 | 4.67 | 30.08 | 19.68 |
| A0A0C4DGN4 | Zymogen granule protein 16 homolog B | In-solution | 9 | 4.22 | 61.57 | 40.14 |
| | | In-gel (SDS-PAGE) | 8 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 6 | 2.83 | 28 | 22 |
| Q9UGM3 | Deleted in malignant brain tumors 1 protein | In-solution | 6 | 2.17 | 13.7 | 7.89 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 6 | 2.83 | 19.83 | 9.39 |
| P01833 | Polymeric immunoglobulin receptor | In-solution | 9 | 3.33 | 27.07 | 6.81 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 6 | 3.17 | 21.22 | 5.08 |
| P01876 | Ig alpha-1 chain C region | In-solution | 8 | 3.75 | 30.29 | 12.85 |
| | | In-gel (SDS-PAGE) | 7 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 5 | 3.6 | 24.01 | 15.33 |
| P23280 | Carbonic anhydrase VI | In-solution | 9 | 6 | 102.8 | 26.37 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 8 | 3.88 | 72.81 | 14.65 |
| C8C504 | Beta-globin | In-solution | 3 | 2 | 12.69 | 17.01 |
| | | In-gel (SDS-PAGE) | 5 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 3 | 3 | 21.96 | 35.63 |
| A7Y9J9 | Mucin 5AC, oligomeric mucus/gel-forming | In-solution | 7 | 4.29 | 24 | 1.91 |
| | | In-gel (SDS-PAGE) | 2 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 4 | 5 | 30.91 | 1.62 |
| P01834 | Ig kappa chain C region | In-solution | 5 | 3.8 | 50.95 | 45.01 |
| | | In-gel (SDS-PAGE) | 6 | N/A | N/A | N/A |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | In-gel (Native-PAGE) | 4 | 3.5 | 45.66 | 49.06 |
| H6VRF8 | Keratin 1 | In-solution | 7 | 7.29 | 67.35 | 13.73 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 8 | 11.13 | 118.26 | 21.91 |
| P13645 | Keratin, type I cytoskeletal 10 | In-solution | 4 | 6.25 | 47.49 | 13.44 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 8 | 5.88 | 50.54 | 14.01 |
| P01036 | Cystatin-S | In-solution | 9 | 2.67 | 86.01 | 30.5 |
| | | In-gel (SDS-PAGE) | 8 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 5 | 1.8 | 35.47 | 20.57 |
| B2R4M6 | Protein S100 | In-solution | 6 | 4.5 | 103.97 | 47.66 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 8 | 3.88 | 52.34 | 43.86 |
| P35908 | Keratin, type II cytoskeletal 2 epidermal | In-solution | 2 | 4.5 | 41.9 | 12.6 |
| | | In-gel (SDS-PAGE) | 9 | N/A | N/A | N/A |
| | | In-gel (Native-PAGE) | 6 | 4.67 | 55.87 | 13.96 |

N/A (Not Applicable)– Refers to the samples from the In-gel (SDS-PAGE) approach where each gel lane was divided into 6 parts for individual digestion and MS/MS analysis. Averaging the number of unique peptides, score and % coverage for the samples submitted to this approach is not applicable since the relative protein abundance cannot be compared among the analyzed areas from the same lane.

## 3.9 References

1. Edgar, W. M. Saliva: Its secretion, composition and functions. *Br Dent J* **172**, 305–312 (1992).

2. Nieuw Amerongen, A. v. & Veerman, E. C. I. Saliva - The defender of the oral cavity. *Oral Dis* **8**, 12–22 (2002).

3. Siqueira, W. L., Custodio, W. & McDonald, E. E. New insights into the composition and functions of the acquired enamel pellicle. *J Dent Res* **91**, 1110–1118 (2012).

4. Gibbons, R. J., Hay, D. I. & Schlesinger, D. H. Delineation of a segment of adsorbed salivary acidic proline-rich proteins which promotes adhesion of Streptococcus gordonii to apatitic surfaces. *Infect Immun* **59**, (1991).

5. Gibbons, R. J. & Hay, D. I. Human salivary acidic proline-rich proteins and statherin promote the attachment of Actinomyces viscosus LY7 to apatitic surfaces. *Infect Immun* **56**, (1988).

6. Cannon, R. D., Nand, A. K. & Jenkinson, H. F. Adherence of Candida albicans to human salivary components adsorbed to hydroxylapatite. *Microbiology (N Y)* **141**, 213–219 (1995).

7. Scannapieco, F. A., Torres, G. & Levine, M. J. Salivary α-Amylase: Role in Dental Plaque and Caries Formation. **4**, 301–307 (2016).

8. Amano, A. *et al.* Salivary receptors for recombinant fimbrillin of Porphyromonas gingivalis. *Infect Immun* **62**, (1994).

9. Rosan, B. & Lamont, R. J. Dental plaque formation. *Microbes Infect* **2**, 1599–1607 (2000).

10. Siqueira, W. L., Helmerhorst, E. J., Zhang, W., Salih, E. & Oppenheim, F. G. Acquired enamel pellicle and its potential role in oral diagnostics. *Ann N Y Acad Sci* **1098**, 504–509 (2007).

11. Roa, N. S., Chaves, M., Gómez, M. & Jaramillo, L. M. Association of salivary proteins with dental caries in a Colombian population. *Acta Odontol Latinoam* **21**, (2008).

12. Levine, M. Salivary proteins may be useful for determining caries susceptibility. *Journal of Evidence-Based Dental Practice* **13**, 91–93 (2013).

13. Martins, C., Buczynski, A. K., Maia, L. C., Siqueira, W. L. & Castro, G. F. B. D. A. Salivary proteins as a biomarker for dental caries-A systematic review. *J Dent* **41**, 2–8 (2013).

14. Kimoto, M., Kishino, M., Yura, Y. & Ogawa, Y. A role of salivary carbonic anhydrase VI in dental plaque. *Arch Oral Biol* **51**, 117–122 (2006).

15. Frasseto, F. *et al.* Relationship among salivary carbonic anhydrase vi activity and flow rate, biofilm ph and caries in primary dentition. *Caries Res* **46**, 194–200 (2012).

16. Scannapieco, F.A. & Levine, M.J. Salivary mucins and dental plaque formation. *Cariology for the Nineties (1993)*

17. Bardow, A., Moe, D., Nyvad, B. & Nauntofte, B. The buffer capacity and buffer systems of human whole saliva measured without loss of CO2. *Arch Oral Biol* **45**, 1–12 (2000).

18. Borghi, G. N. *et al.* Relationship among α amylase and carbonic anhydrase VI in saliva, visible biofilm, and early childhood caries: a longitudinal study. *Int J Paediatr Dent* **27**, 174–182 (2017).

19. Parkkila, S., Parkkila, A. K., Vierjoki, T., Stahlberg, T. & Rajaniemi, H. Competitive time-resolved immunofluorometric assay for quantifying carbonic anhydrase VI in saliva. *Clin Chem* **39**, (1993).

20. Azzopardi, E., Lloyd, C., Teixeira, S. R., Conlan, R. S. & Whitaker, I. S. Clinical applications of amylase: Novel perspectives. *Surgery (United States)* **160**, 26–37 (2016).

21. Rosenblum, J. L., Irwin, C. L. & Alpers, D. H. Starch and glucose oligosaccharides protect salivary-type amylase activity at acid pH. *Am J Physiol Gastrointest Liver Physiol* **254**, (1988).

22. Kandra, L., Zajácz, Á., Remenyik, J. & Gyémánt, G. Kinetic investigation of a new inhibitor for human salivary α-amylase. *Biochem Biophys Res Commun* **334**, 824–828 (2005).

23. Choi, S. *et al.* Identification of Porphyromonas gingivalis lipopolysaccharide-binding proteins in human saliva. *Mol Immunol* **48**, 2207–2213 (2011).

24. Baik, J. E. *et al.* Alpha-amylase is a human salivary protein with affinity to lipopolysaccharide of Aggregatibacter actinomycetemcomitans. *Mol Oral Microbiol* **28**, 142–153 (2013).

25. Ochiai, A. *et al.* α-Amylase is a potential growth inhibitor of Porphyromonas gingivalis, a periodontal pathogenic bacterium. *J Periodontal Res* **49**, 62–68 (2014).

26. Siqueira, W. L., Lee, Y. H., Xiao, Y., Held, K. & Wong, W. Identification and characterization of histatin 1 salivary complexes by using mass spectrometry. *Proteomics* **12**, 3426–3435 (2012).

27. Moffa, E. B. *et al.* Histatin 5 inhibits adhesion of C. albicans to Reconstructed Human Oral Epithelium. *Front Microbiol* **6**, (2015).

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

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

30. Thomadaki, K. *et al.* Whole-saliva proteolysis and its impact on salivary diagnostics. *J Dent Res* **90**, 1325–1330 (2011).

31. Deutsch, O. *et al.* An approach to remove alpha amylase for proteomic analysis of low abundance biomarkers in human saliva. *Electrophoresis* **29**, 4150–4157 (2008).

32. Crosara, K. T. B., Moffa, E. B., Xiao, Y. & Siqueira, W. L. Merging in-silico and in vitro salivary protein complex partners using the STRING database: A tutorial. *J Proteomics* **171**, 87–94 (2018).

33. Moffa, E. B. *et al.* In vitro identification of histatin 5 salivary complexes. *PLoS One* **10**, (2015).

34. Helmerhorst, E. J., Alagl, A. S., Siqueira, W. L. & Oppenheim, F. G. Oral fluid proteolytic effects on histatin 5 structure and function. *Arch Oral Biol* **51**, 1061–1070 (2006).

35. McDonald, E. E., Goldberg, H. A., Tabbara, N., Mendes, F. M. & Siqueira, W. L. Histatin 1 resists proteolytic degradation when adsorbed to hydroxyapatite. *J Dent Res* **90**, 268–272 (2011).

36. Tenovuo, J. O. Human Saliva: Clinical Chemistry and Microbiology. *Human Saliva, vol 1, CRC press.* (2021).

37. Ramya, A. S., Uppala, D., Majumdar, S., Surekha, C. & Deepak, K. G. K. Are salivary amylase and pH - Prognostic indicators of cancers? *J Oral Biol Craniofac Res* **5**, 81–85 (2015).

38. Bobek, L. A. & Levine, M. J. Cystatins - Inhibitors of cysteine proteinases. *Critical Reviews in Oral Biology and Medicine* **3**, 307–332 (1992).

39. Baron, A. C., DeCarlo, A. A. & Featherstone, J. D. B. Functional aspects of the human salivary cystatins in the oral environment. *Oral Dis* **5**, 234–240 (1999).

40. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human submandibular saliva. *Arch Oral Biol* **19**, 887–895 (1974).

41. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human parotid saliva. *Arch Oral Biol* **14**, 277–294 (1969).

42. Dawes, C. Effects of Diet on Salivary Secretion and Composition. *J Dent Res* **49**, 1263–1272 (1970).

43. Lee, Y. H. & Wong, D. T. Saliva: An emerging biofluid for early detection of diseases. *American Journal of Dentistry* vol. 22 (2009).

44. Bosch, J. A. *et al.* Psychological stress as a determinant of protein levels and salivary- induced aggregation of Streptococcus gordonii in human whole saliva. *Psychosom Med* **58**, 374–382 (1996).

# Chapter 4

# 4 Effects of Intensity and Duration of Stimulation on the Proteome of the Secretion from the Human Parotid Glands

## 4.1 Abstract

The concentrations of total protein and many electrolytes in the parotid gland (PG) secretion are affected by the duration (time) and the intensity (flow) of stimulation, under continuous secretion. How these factors affect the PG secretion proteome is unknown. Variations in the organic composition of the PG secretion were here investigated through total protein measurements and bottom-up mass spectrometry proteomic analysis of PG saliva collected at two constant flow rates (0.25 and 1.00 ml/min), for 30 consecutive minutes. Total protein measurements were done for every minute for each participant, while pooled saliva of selected time points (1, 5, 10, 20, and 30 minutes) were analyzed for proteome determination. Proteins identified in three pools for specific time and flow were considered. Total protein concentration varied within and between flow rates, confirming the effect of flow ($F_{1,538} = 34.143$, $p < 0.001$) and duration ($F_{29,534} = 2.012$, $p < 0.003$) of stimulation. More proteins were identified at 0.25 ml/min than at 1.00 ml/min at all time points. Combining both flows, 169 non-redundant proteins were identified: 79 of them only at 0.25 ml/min, 40 only at 1.00 ml/min, and 50 at both flow rates. Stimulation intensity strongly affected 119 proteins, at least 44 were affected by both factors, and 4 by neither, suggesting possible protein-specific secretory mechanisms. Forty-two novel PG proteins were identified. Biological functional variations between time points and flow rates were also observed in response to the proteomic differences. Effects of intensity and duration of stimulation should be considered when designing saliva collection protocols aimed at biomarker discovery and validation.

## 4.2 Introduction

Increasing interest has been shown in the past decade in the use of saliva for disease diagnosis and monitoring. This complex oral fluid is predominantly formed by the secretion from the major (parotid, submandibular and sublingual) and numerous minor salivary glands, and complemented by the gingival-crevicular fluid, contributions from the oral microbiome, and desquamated cells and fluids from the oral mucosa [1]. This combination of exocrine and non-exocrine components present in the oral cavity is termed WS. The extremely rich composition of WS, in addition to its painless, easy, and fast collection process, makes it an attractive diagnostic option compared to other biological samples. Moreover, the concentration of certain proteins is higher in the secretion from salivary glands than that in WS or blood [2]. In addition to containing many proteins present in plasma and/or tears, WS contains unique proteins broadening its diagnostic possibilities [3]. Moreover, all sources contributing to the formation of WS participate with their distinct proteomes, with the parotid glands (PG) secreting the highest percentage of unique proteins (15.3%), compared to submandibular (SMG)/sublingual (SLG) (8.5%), minor (5.4%) glands, and gingival-crevicular fluid (8%) [4].

Different salivary glands produce specific types of secretion according to the nature of their acinar cells. While PGs are formed by serous cells and produce a watery amylase-rich fluid [5], SMG and SLG contain mixed cell types (serous and mucous) and secrete a more viscous mucin-rich fluid [5]. In the recently discovered tubarial glands, mucous and serous cell types were identified, with predominance for mucous secretion [6]. The secretions from the minor salivary glands vary depending on their location. Palatine and retromolar glands secrete mucin-rich saliva; buccal and labial glands have mixed cell types, with a predominance of mucous cells; and lingual glands hold serous cells and produce a watery, lipase-rich fluid [5]. Because the composition of the saliva secreted by different salivary glands is not the same, many studies have investigated the proteome of each salivary gland separately [3,7–19]. However, the relative contribution of individual glands to WS volume changes impressively due to stimulation. Although secretions from SMG, PG, minor salivary glands, and SLG, represent approximately 65, 20, 10, and 5%, respectively of the total volume of the unstimulated WS, the contribution of the PG increases to up to 50% of the

total WS volume during stimulated flow [1,3,5,20], while the contribution of the SMG to WS decreases to 35% of the total WS volume with stimulation, and the secretions from the SLG and minor glands remain stable regardless of stimulation [5]. Therefore, variations in the volume secreted by each gland type due to changes in the stimulation can impact the final WS composition, with PG being the main contributor during stimulated secretion.

The rate at which saliva is formed in the mouth varies between individuals, with about 10% of the population presenting an unstimulated flow rate of $\leq 0.1$ ml/min, and a mean overall population value of 0.3 to 0.4 ml/min [21]. Salivary secretion continues during the whole day, with a substantial increase in the flow rate during eating and drinking, and virtually zero during sleep [21]. Many factors can influence saliva's volume, flow, and composition [5]. For instance, the circadian rhythm affects flow rate, protein secretion, and salivary sodium and chloride concentrations, particularly for unstimulated saliva [5,22], but shows less influence over stimulated glandular saliva secretion [23]. Additionally, the level of hydration of the body [5,24–27], gland size [5,28,29], diseases [5,30], use of medication [5,31–33], sex [34,35], environmental temperature [36], emotional stress [37–40], the intensity of exercise [41–44], diet [12,45,46], intensity and duration of stimulation [23,47,48], type of stimuli [5,49–54] and time from last meal [55] are some factors that may cause variations in the WS and gland secretions. The effect of aging on the salivary flow rate is not clear, with studies suggesting stable [5,25,33,56–60] or declining salivary gland function with aging [5,61,62]. However, the function of the PG, which contribute most to chewing-stimulated secretion, remains stable despite aging in healthy, non-medicated individuals [5,56], highlighting the dependability on the secretion from PG for diagnosis or monitoring of diseases.

The influence of intensity (reflected by flow rate) and duration (reflected by time) of stimulation on the total concentration of proteins and many of the main electrolytes in the human WS, and in the PG and SMG/SLG secretions, have been studied [23,47–49,63,64]. In fact, through a negative-feedback technique, Dawes observed that when the flow rate of the PG secretion was maintained constant for 15 min at 0.25, 0.50, and 1.00 ml/min, the pH and the concentrations of bicarbonate, chloride, mono, and dihydrogen orthophosphate, calcium, and proteins varied with time throughout the 15 min of stimulation [48]. Interestingly, the protein concentration fell at the beginning of the stimulation for all tested

flow rates, but increased with the duration of stimulation, particularly at the highest flow rate [48], without reaching a maximum in the observed period. This fact demonstrates constant changes in the inorganic and organic composition of the PG secretion depending on the flow and duration of the stimulus. For SMG/SLG secretion, on the other hand, the concentration of total protein increased immediately after stimulation, rose to a maximum, and remained at this level, with the magnitude depending on the intensity, but not on the duration of stimulation [23]. Curiously, the concentration of total protein in WS upon stimulation was quite different from the patterns observed in the glandular secretions. Total protein concentration in WS decreased immediately after stimulation, and it was not affected by the duration (15 min) nor intensity (10 or 60 strokes/min on parafilm) of stimulation [65]. Additionally, no changes were observed in the levels of salivary mucin MG2 in WS stimulated nor unstimulated, regardless of the duration or the intensity of mastication. However, the levels of salivary mucin MG1 varied upon stimulation but were not affected by the duration or intensity of stimulation [65], suggesting different secretion mechanisms for MG2 and MG1 in WS. Similarly, the secretory rates of some antimicrobial proteins (lysozyme, lactoferrin, salivary peroxidase, and sIgA) in the PG secretion were affected by stimulation, while the rates of others were not, suggesting a contrast between active and passive secretion, and supporting differences in the secretory mechanisms of salivary proteins [66–68]. Such variations in the secretion rates of some antimicrobial proteins (lysozyme, lactoferrin, salivary peroxidase, and sIgA), and the total protein concentration of stimulated PG secretion due to different flow rates [66–68], reinforce the difficulty to compare the proteome from persons with distinctive flow rates [66,69].

Since a correct characterization of the proteome of WS, and the secretion from each salivary gland, is critical for the successful utilization of such fluids for the diagnosis and monitoring of local and systemic diseases, we investigated possible variations in the proteome of the secretion from the human PG, under continuously stimulated flow rates of 0.25 and 1.00 ml/min, for 30 subsequent minutes. Because human PGs are the main contributors to the total WS volume under stimulation, our observations are extremely important factors to be considered when designing saliva collection protocols.

## 4.3 Materials and Methods

### 4.3.1 Ethics Approval for Human Participants

This study was approved by the Research Human Ethics Board of the University of Western Ontario (review number 16181E).

### 4.3.2 Subjects Description and Collection of Parotid Secretion Under Continuous Flow

This study included 4 healthy volunteers, ranging in age from 28 to 40 years, with both sexes represented (2 females and 2 males). All volunteers exhibited good oral health and good overall systemic health. Saliva was collected from the duct of the PG (Stensen's duct), with the assistance of a Lashley cup device [70], which was washed and autoclaved the day before each collection. New plastic tubes were autoclaved and attached to the cup for each collection. The collection happened between 10:00 AM and 11:00 AM. Subjects were asked to perform morning oral hygiene and refrain from eating or drinking 2 hours before the collection. All collections were done between the spring and summer seasons. All subjects had their saliva collected at least three times (biological triplicate) with a one-week interval in between collections. Only one flow rate was collected per day.

### 4.3.3 Maintenance of a Continuous Flow Rate

Negative-feedback technique, developed by Dawes [48], was used to maintain two constant stimulated flow rates (0.25 and 1.00 ml/min). Volunteers were trained to secrete parotid saliva in the two desired flow rates. Samples were collected for 30 consecutive minutes, under continuous flow rates of 0.25 ml/min or 1.00 ml/min, and samples related to each minute were placed in separate tubes. Gustatory stimulation was given by sour candy, with no added sugar. Thirty 1.5 ml microcentrifuge tubes (Axygen, Union City, CA, USA) were labelled 1 through 30, had the corresponding volume (0.25 or 1.00 ml) drawn with a marker, and were placed on ice in front of each volunteer. Subjects were instructed to consume the amount of sour candy needed to maintain the desired flow rates (0.25 or 1.00 ml/min) for 30 consecutive minutes, with minimal tongue movement to prevent the displacement of the collection device. A chronometer was positioned in front of each

volunteer so they could know if the flow rate was adequate. Participants could practice until they were comfortable manipulating their flow rate. During all collections, the labelled tubes were constantly kept on ice inside a Styrofoam box. At every one-minute mark, the collection tube was switched. The volume of saliva collected in each tube was carefully measured and annotated.

### 4.3.4 Total Protein Quantification

Bicinchoninic acid assay (BCA) (Pierce Biotechnology, Rockford, USA) was used to measure the total protein concentration. This colorimetric method uses bovine serum albumin as the protein standard for the calibration curve (0 to 2000 μg/ml). After the addition of the working reagent and incubation at 37°C for 30 minutes, the absorbance was measured with a spectrometer (iMArk Microplate Reader, Bio-Rad Inc., Hercules, USA). This method was used for both individual and pooled samples as described below.

*Individual samples.*

The total protein concentration was determined for all participants individually, for every one minute, for the total 30 minutes, immediately after collection.

*Pooled samples.*

The salivary pools were prepared with samples from the participants who collected saliva at the same date and time (subjects 1, 2, and 3). Equal volumes of fresh saliva from each participant were taken from their samples collected at 1, 5, 10, 20, and 30 minutes after the beginning of the stimulation (T1, T5, T10, T20, and T30), and pooled according to the recorded time. Total protein concentration was determined for each pooled sample using the BCA assay, and volumes equivalent to 20 μg of protein from each pool were placed in separate tubes and dried using a rotatory evaporator (Vacufuge Plus, Eppendorf, Hamburg, Germany) for future mass-spectrometric analysis.

#### 4.3.4.1 Statistical Analysis

All statistical analyses were done using Statistical Package for the Social Sciences (SPSS, IBM Corp. Armonk, NY: IBM Corp.) version 24. The data for the total protein concentration of all the individual samples collected from T1 through T30, for both flow rates, were analyzed using a Linear Mixed Effects Model. The data were treated as a

blocked, 2-factor factorial, the fixed effects being: (1) time at 30 levels (30 minutes), (2) flow rate at 2 levels (0.25 ml/min and 1.0 ml/min), and the subjects and collection dates as the blocks (random effects), using Maximum Likelihood and Variance Components model. Significance was tested for time and flow rate, and the interaction between the two fixed factors (Flow*Time), as well as for the random effects (subject and day of collection). The F-test using Satterthwaite's approximation of the degree of freedom [71] was used to determine the significance of flow and time. Wald-based Z test was used to determine the significance of subjects and collection dates. In all analyses, only *p*-values lower than 0.05 were considered significant.

### 4.3.5 Proteomic Analysis

#### 4.3.5.1 In-solution Digestion

The aliquots containing 20 μg of protein from the pooled samples collected at T1, T5, T10, T20, and T30, from the three collection dates, at flow rates of 0.25 and 1.00 ml/min, were denatured and reduced by the addition of 50 μl of 4 M urea, 10 mM DTT, and 50 mM ammonium bicarbonate, pH 7.8. Samples were incubated in this solution for 1 hour at room temperature before they were diluted with the addition of 150 μl of 50 mM $NH_4$ $HCO_3$, pH 7.8. Next, tryptic digestion was carried out by the addition of 2 % w/w sequencing grade trypsin (Promega, Madison, WI, USA), and incubation at 37 °C for 16 hours. After that, samples were dried and submitted to the desalting procedure using C-18 Zip-Tips (Millipore, Billerica, USA), before mass spectrometry (MS).

#### 4.3.5.2 Mass-spectrometric Analysis

Samples were re-suspended in 97.5% distilled water/ 2.4% ACN/ 0.1% formic acid and subjected to RP nLC-ESI-MS/MS. An LTQ-Velos spectrometer (Thermo Fisher Scientific, San Jose, USA) was used. Liquid chromatography (LC) aligned with a C-18 column of capillary-fused silica (100 mm by 75 μm column, 3 μm spherical beads, and 100 Å pore size) was used linked to the mass spectrometer through electrospray ionization (ESI). The survey scan was set in the range (m/z) between 390 and 2000. LC was developed over 85 minutes, by a solvent gradient of 0 to 100% of solvent B (97.5% ACN and 0.1% formic

acid), with a flow rate of 200 nl/min e maximum pressure of 280 bar. The voltage of the electrospray and the ionic transfer capillary temperature were 1.8 kV and 250 ºC, respectively. The MS/MS spectra were matched with the human proteome databank (SwissProt and TrEMBL, Swiss Institute of Bioinformatics, Geneva, Switzerland, http://ca.expasy.org/sprot) using SEQUEST algorithm in Proteome Discover 3.1 software (Thermo Scientific, USA). The results were filtered using XCorr and included a false discovery of 1%. Two wrong cleavages were allowed: carbamide methylated cysteine, and phosphorylated serine, threonine, and tyrosine. Oxidized methionine was considered a dynamic modification. All identified proteins were sequenced according to the UniProt database (https://www.uniprot.org).

#### 4.3.5.3 Bioinformatics Characterization of the Proteins Identified with High Confidence

The proteins that were identified in all three collections (biological triplicate), for each specific time point and flow rate, were considered to show high confidence, and they were used for the subsequent analyses. The identified proteins' calculated isoelectric point (pI) and molecular weight (MW) were used. Characterization of the high-confidence proteome of each time point was done for the two flow rates based on the number of identified proteins, and their average MW and pI. Using the physiologic salivary pH as a reference (pH 6.8), the identified proteins were grouped according to their pI (pI below or above 6.8). The MW of amylase (56 kDa), the most abundant protein in the PG secretion, was assumed for our MW cut-off (MW below and above 56 kDa). The individual pI and MW were calculated after the removal of the signal peptide given by the UniProt database. Otherwise, pI and MW informed in the MS report were adopted. Comparison between the proteins identified with high confidence for each time point was conducted using an interactive tool for comparing lists with Venn diagrams, Venny 2.0 (https://bioinfogp.cnb.csic.es). Venn diagrams were constructed to assess proteomic variations within and between the two studied flow rates. Qualitative variations in the secreted proteome relative to the duration (time) and intensity (flow) of stimulation were investigated by comparing the proteome of immediately subsequent time points within each flow rate and the proteome of a specific time point between the different flows. The glandular origin of the identified proteins was

investigated using the Human Salivary Proteome Wiki (https://www.salivaryproteome.org), Bgee gene expression database (https://bgee.org), and the Human Proteome Atlas (https://www.proteinatlas.org), and compared against the frequency at which the proteins were identified in our study.

#### 4.3.5.4 Functional Analysis and Clinical Implications

The identified proteins were also classified based on their Molecular Function, Biological Processes, and Pathways using data from PANTHER (https://www.pantherdb.org), assessed between December 2021 and February 2022. Proteins involved in biological pathways related to cancers, cardiovascular diseases, and neurological diseases were identified due to their potential clinical applications.

## 4.4 Results

After Ethics Committee approval and written informed consent were obtained, four participants were recruited for the study. Three participants had their saliva collected on the same days, while one participant collected saliva on separate dates, and was excluded from the saliva pools used for proteome determination.

### 4.4.1 Maintenance of a Continuous Flow Rate

All volunteers achieved continuous secretion of parotid saliva at the higher flow rate of 1.00 ml/min for 30 minutes; however, only two volunteers maintained continuous secretion at the lower flow rate of 0.25 ml/min, for the same period. A descriptive analysis showing the achievement of a continuous flow rate by the subjects throughout the three collections is found in Table 4.1.

The average volume for all samples collected at the flow rate of 1.00 ml/min was between 1.02 and 0.96 ml, with a standard deviation of 0.02 to 0.11, and median=1. The participants that achieved the 0.25 ml/min flow rate maintained between 0.25 and 0.26 ml, with a standard deviation between 0.01 and 0.03, and median=0.25.

**Table 4.1 Descriptive analysis showing the achievement of continuous flow rates for 30 minutes by the participants throughout the three collections. Four participants maintained the flow rate of 1.00 ml/min, median=1. Two participants maintained the flow rate of 0.25 ml/min, median=0.25. Results are shown in ml.**

| Participants | 1st collection | 2nd collection | 3rd collection |
|---|---|---|---|
| 1 | Average 1.05 | Average 1.004 | Average 1.04 |
| | SD 0.07 | SD 0.04 | SD 0.06 |
| | Median 1 | Median 1 | Median 1 |
| 2 | Average 1.00 | Average 1.00 | Average 1.02 |
| | SD 0.07 | SD 0.05 | SD 0.02 |
| | Median 1 | Median 1 | Median 1 |
| 3 | Average 1.00 | Average 1.00 | Average 1.00 |
| | SD 0.02 | SD 0.02 | SD 0.03 |
| | Median 1 | Median 1 | Median 1 |
| 4 | Average 0.96 | Average 1.00 | Average 1.00 |
| | SD 0.11 | SD 0.04 | SD 0.10 |
| | Median 1 | Median 1 | Median 1 |
| **Participants** | **1st collection** | **2nd collection** | **3rd collection** |
| 1 | Average 0.26 | Average 0.26 | Average 0.25 |
| | SD 0.03 | SD 0.02 | SD 0.02 |
| | Median 0.25 | Median 0.25 | Median 0.25 |
| 2 | Average 0.26 | Average 0.26 | Average 0.25 |
| | SD 0.017 | SD 0.014 | SD 0.01 |
| | Median 0.25 | Median 0.25 | Median 0.25 |

### 4.4.2 Total Protein Concentration

Total protein concentration fell rapidly at the beginning of the stimulation at both flow rates (Figure 4.1). The minimum average concentration was reached at 7 minutes into the collection at 0.25 ml/min (1078 μg/ml). The lowest average concentration of 1.00 ml/min was reached at 2 minutes into collection (1117 μg/ml). For both flow rates, the concentration rose after the minimum was reached, and the concentration at 1.00 ml/min surpassed the concentration of 0.25 ml/min just 5 minutes after the beginning of the collection. After the minimum was reached for both flows, the total protein concentration increased, with a more pronounced increase at 1.00 ml/min in all subsequent time points. No plateau was reached in the tested flow rates until the end of the 30-minute observation.

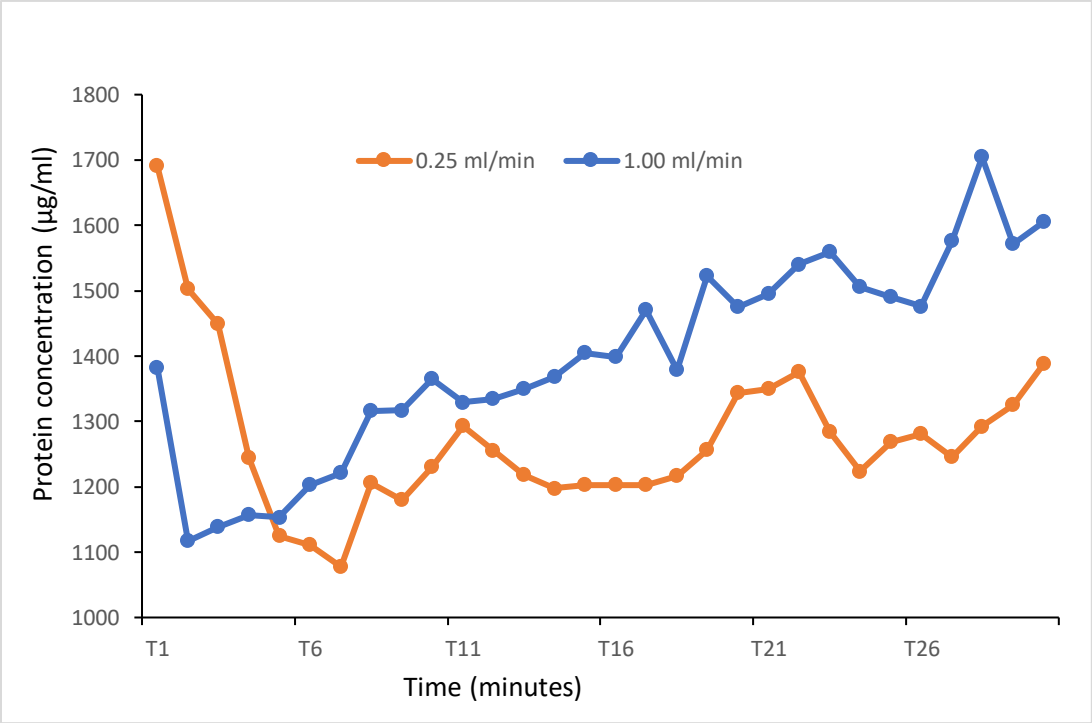



**Figure 4.1 Average total protein concentration measured from the PG ductal secretion, over a period of 30 minutes of stimulation, with achievement of continuous flow rates of 0.25 (shown in orange) and 1.00 ml/min (shown in blue). Each orange point represents the average of 6 measurements (3 different collections from 2 participants). Each blue point represents the average of 12 measurements (3 separate collections from 4 participants).**

The F-test, using Satterthwaite's approximation of the degree of freedom, indicates that the intensity of stimulation had a significant effect on the total protein concentration (Flow, $F_{1,538} = 34.143$, $p < 0.001$), and so did the duration of stimulation (Time, $F_{29,534} = 2.012$, $p < 0.003$) (Table 4.2). No significant differences were observed between subjects ($p = 0.175$) or between the three collections ($p = 0.278$) (Table 4.3).

**Table 4.2 Results of the Type III Tests of Fixed Effects using F-test, with Satterthwaite's approximation of the degree of freedom, indicating that intensity (Flow, $F_{1,538} = 34.143$, $p < 0.001$) and duration (Time, $F_{29,534} = 2.012$, $p < 0.003$) of stimulation had significant effects on the total protein concentration of the secretion from the PG.**

| Source | Numerator df | Denominator df | F | Sig. |
|---|---|---|---|---|
| **Intercept** | 1 | 5.871 | 169.921 | .000 |
| **Flow** | 1 | 537.664 | 34.143 | .000 |
| **Time** | 29 | 533.989 | 2.012 | .002 |
| **Flow*Time** | 29 | 533.989 | 1.599 | .026 |

Dependent Variable: Protein concentration.

**Table 4.3 Results of the Estimates of Covariance Parameters for Random Effects using Wald-based Z test indicating that there were no significant differences between the total protein concentration of the PG saliva collected from different subjects (Subject, p = 0.175) or on different dates (Collection, p = 0.278).**

| Parameter | Estimate | Std. Error | Wald Z | Sig. |
|---|---|---|---|---|
| **Residual** | 83970.5712 | 5138.96893 | 16.340 | .000 |
| **Subject Variance** | 25638.0492 | 18905.3457 | 1.356 | .175 |
| **Collection Variance** | 10312.7798 | 9501.16624 | 1.085 | .278 |

Dependent Variable: Protein Concentration

### 4.4.3 Bioinformatics Characterization of the Identified Proteins

#### 4.4.3.1 All Identified Proteins

The proteins identified with high confidence were used to characterize the selected time points and flow rates based on the number of identified proteins, and their MW and pI. A total of 129 unique proteins were identified with high confidence for the flow rate of 0.25 ml/min, and 90 unique proteins at 1.00 ml/min. The number of proteins identified at different time points varied. At 0.25 ml/min, 49, 47, 39, 37, and 58 proteins were identified in T1, T5, T10, T20, and T30, respectively (Figure 4.2). At 1.00 ml/min, 41, 36, 26, 36, and 22 proteins were identified in T1, T5, T10, T20, and T30, respectively.

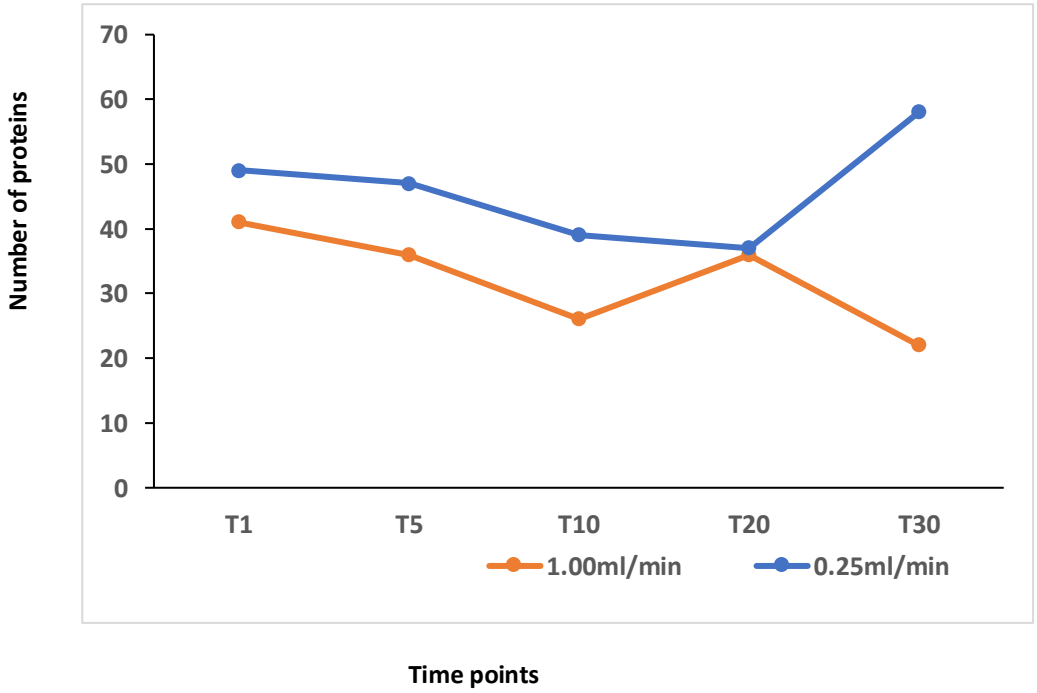
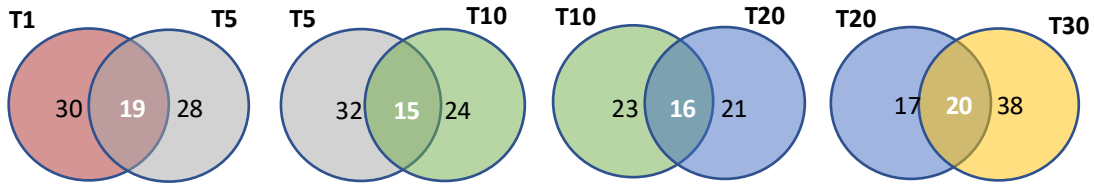



**Figure 4.2 Number of proteins identified in all 3 collections in each specific time point, according to the studied flow rates (1.00 ml/min and 0.25 ml/min).**

Within flow rate analysis showed that the average number of proteins common between two subsequent time points was 17.5 (+/- 2.38) at 0.25 ml/min, and 13.75 (+/- 1.26) proteins at 1.00 ml/min (Figure 4.3). Between flow rates analysis showed an overlap of 13.2 (+/- 1.92) proteins among the same time points (Figure 4.4).

**4.3 (A) Flow rate: 0.25 ml/min**




**4.3 (B) Flow rate: 1.00 ml/min**

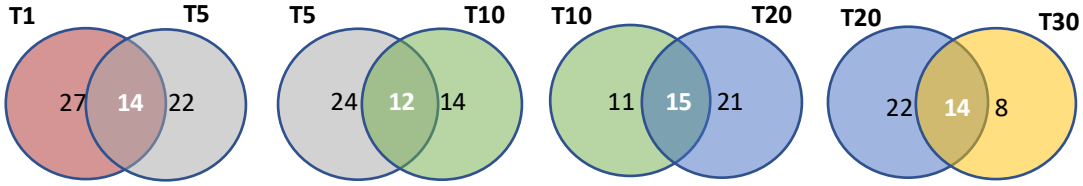



**Figure 4.3 Within flow rates overlapping of proteins identified in subsequent time points at 0.25 ml/min (top, 4.3 (A)), and at 1.00 ml/min (bottom, 4.3 (B)). Mean number of proteins overlapping +/- SD at 0.25 ml/min, 17.5 +/- 2.38, and at 1.00 ml/min, 13.75 +/- 1.26.**

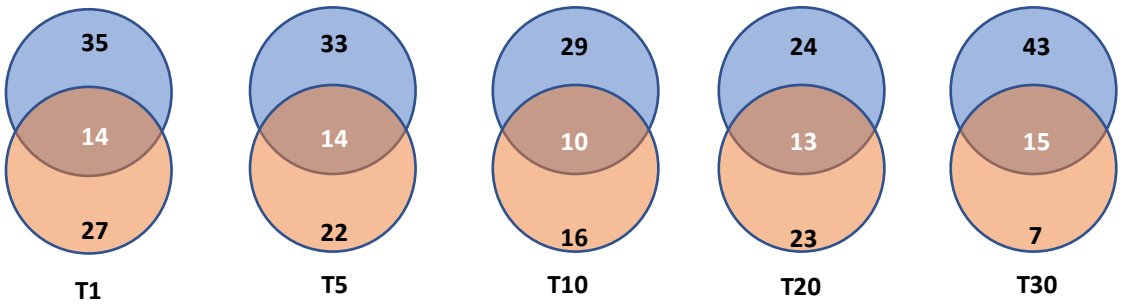



**Figure 4.4 Between flow rates overlapping of proteins identified in the same time point of different flows, 0.25 ml/min (top, blue) and 1.00 ml/min (bottom, orange). The average number of proteins overlapping +/- SD is 13.2 +/- 1.92.**

The distribution of the high-confidence proteins within each flow rate showed that most proteins were specific to individual time points, while some proteins were common to two or more time points (Figure 4.5). A similar number of proteins were common to all time points for both flow rates (8 for 1.00 ml/min, and 7 for 0.25 ml/min); however, only 4 of them were identified in all time points for both flow rates. They were alpha-amylase 1A, carbonic anhydrase 6, mucin-16, and submaxillary gland androgen-regulated protein 3B. The other proteins that were identified in all time points at 0.25 ml/min were mucin-12, mucin-19, and perilipin-4. Other proteins that were found in all time points at 1.00 ml/min were extracellular matrix organizing protein FRAS1, pancreatic adenocarcinoma upregulated factor, prolactin-inducible protein, and titin.

**Figure 4.5 (A)**

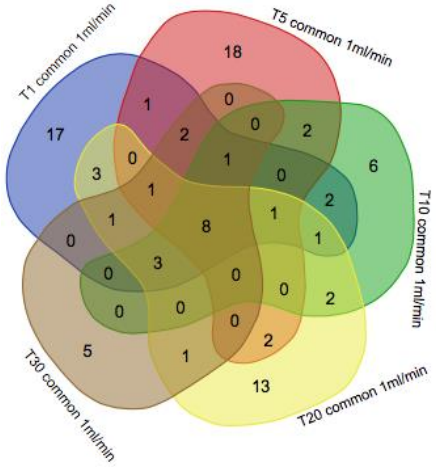



**Figure 4.5 (B)**

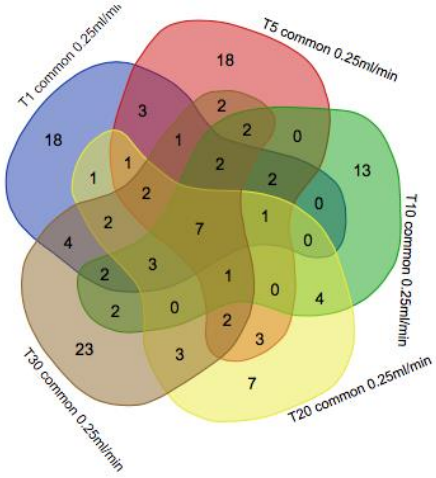



**Figure 4.5 Distribution of proteins identified in all 3 collections, in each specific time point (T1, T5, T10, T20, and T30), according to the studied flow rates of 1mL/min (Fig. 4.5 (A)) and 0.25 mL/min (Fig. 4.5 (B)).**

The average MW of the identified proteins was very similar within and between the two flow rates. The percentage of proteins with MW > 56 kDa was higher than those with MW < 56 kDa in at time points and in both flow rates (Figure 4.6).

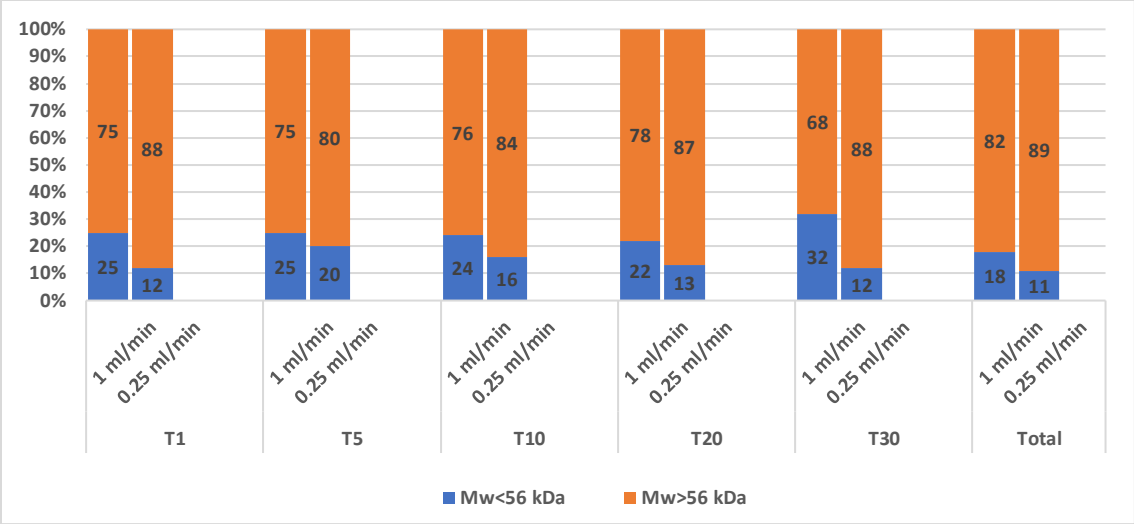



**Figure 4.6 Percentage of proteins identified in all 3 collections, in each specific time point, for the studied flow rates, with molecular weights below or above 56 kDa.**

The average pI was also very similar between and within flow rates. The percentage of proteins identified with pI<6.8 in all time points, for both flows, was higher than those with pI > 6.8 (Figure 4.7).

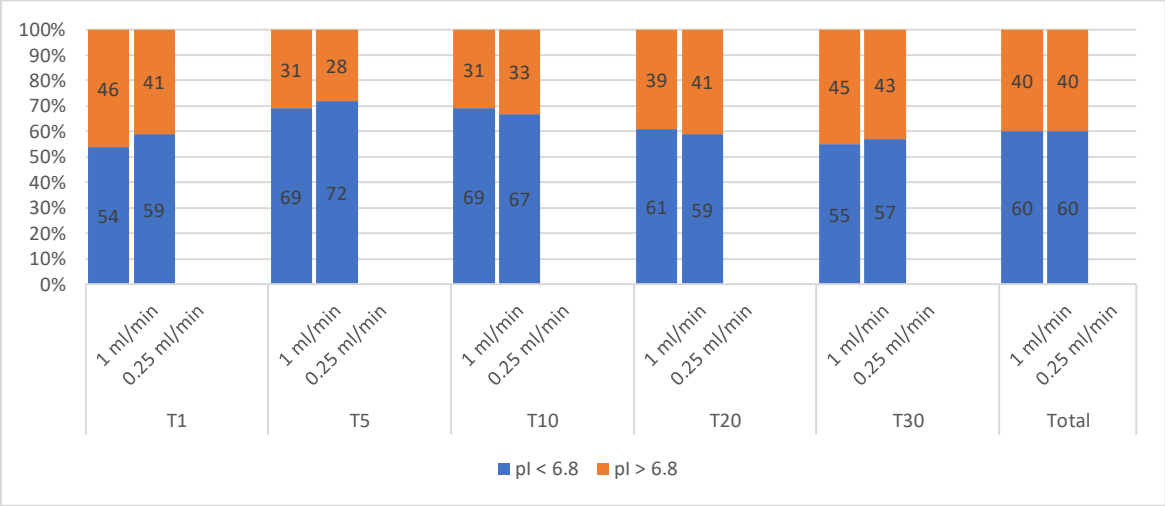


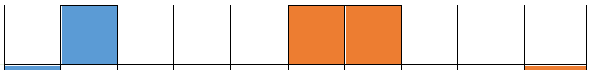
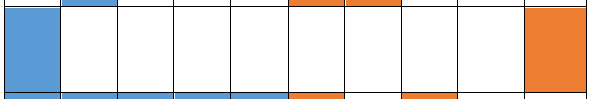
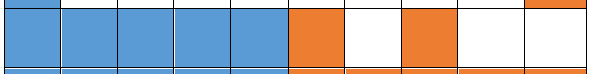
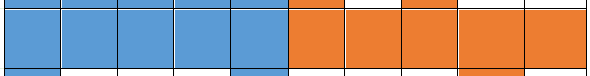
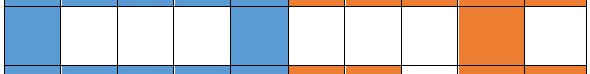
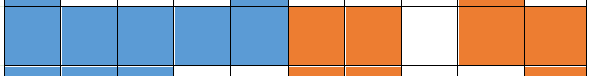
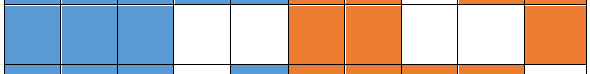
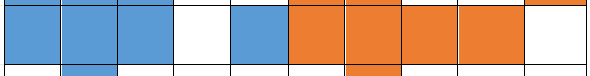
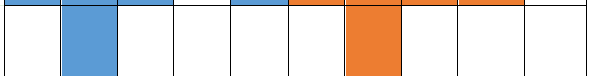
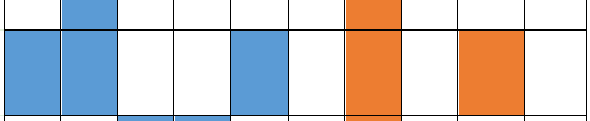
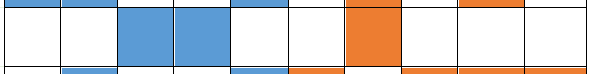
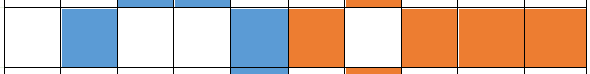
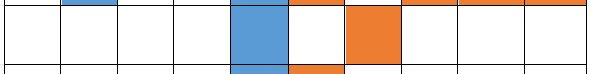
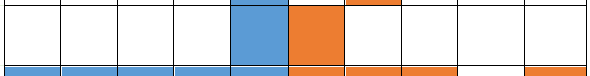
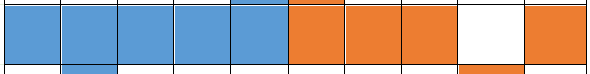
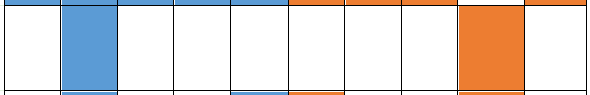
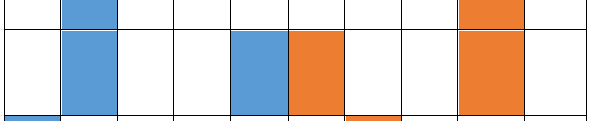
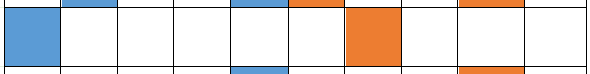
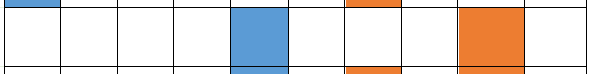
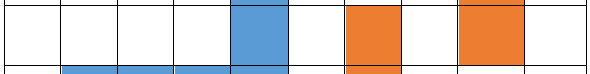
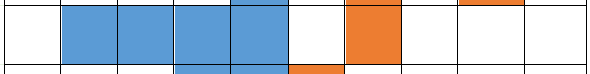
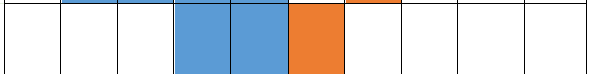
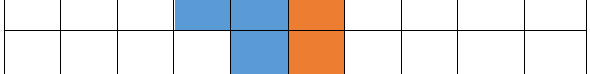
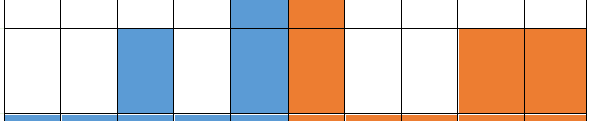

**Figure 4.7 Percentage of proteins identified in all 3 collections, in each specific time point, for the studied flow rates, with isoelectric points below or above 6.8.**

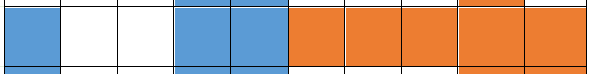
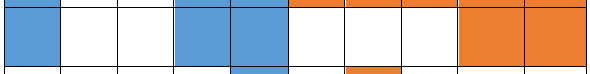
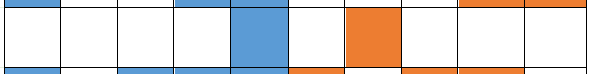
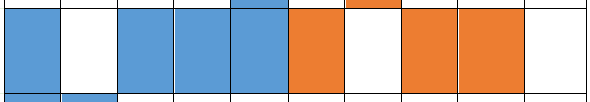
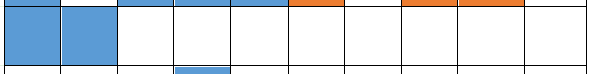
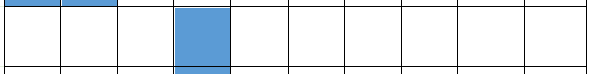
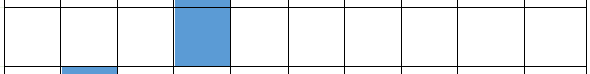
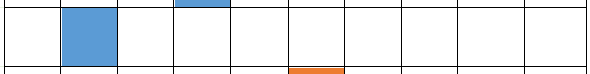
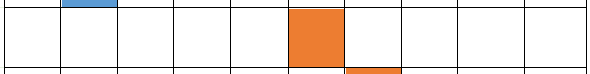
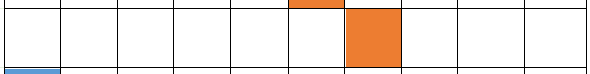
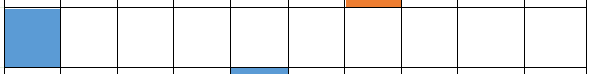
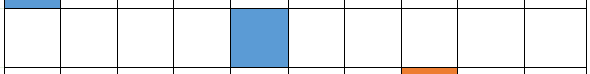
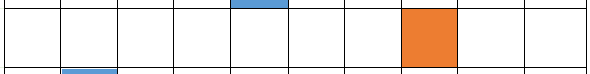
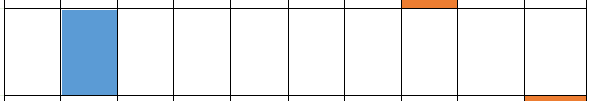
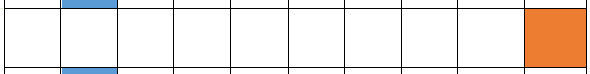
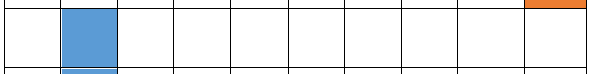
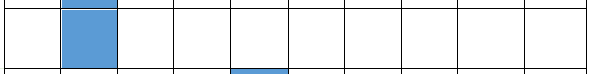
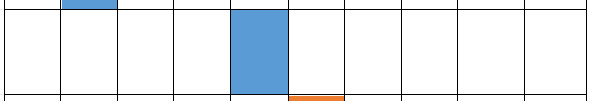
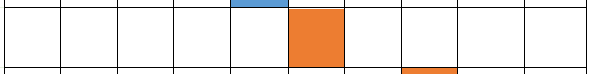
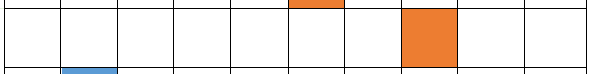
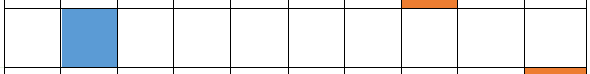
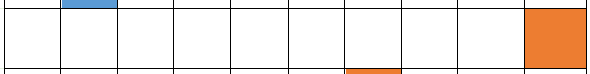
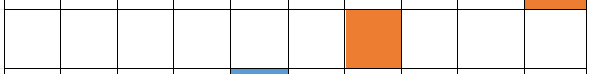
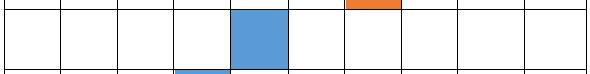
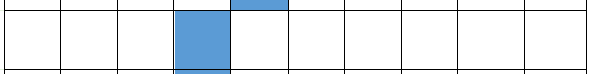
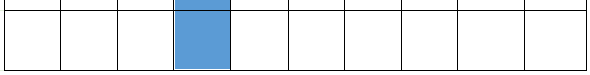
### 4.4.3.2 Proteins that were Identified in Both Flows

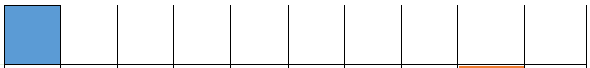
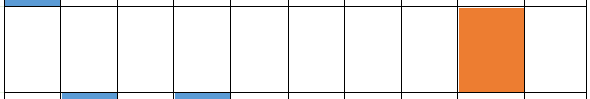
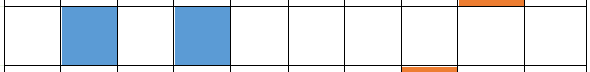
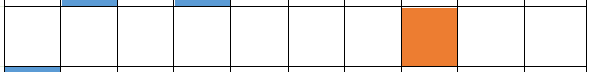
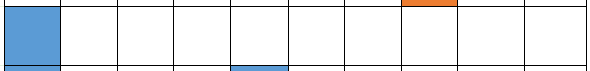
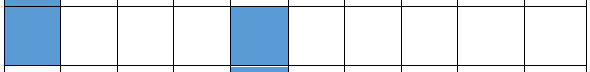
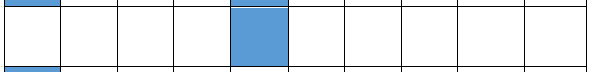
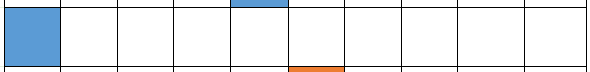
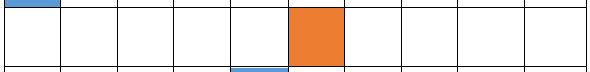
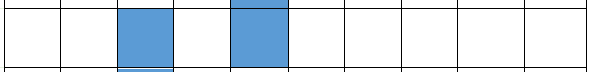
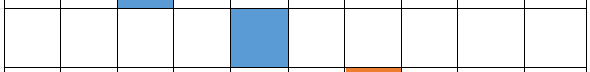
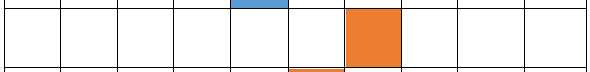
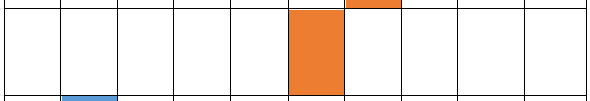
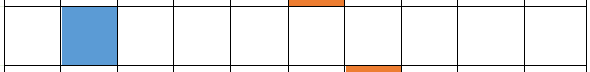
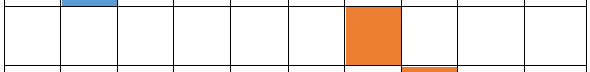
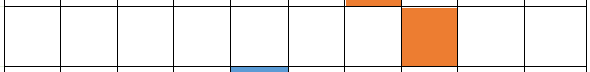
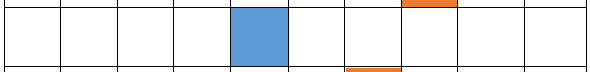
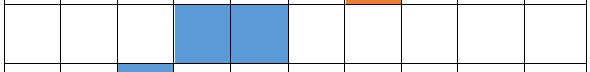
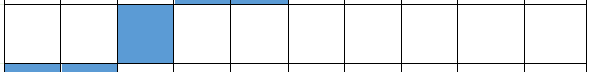
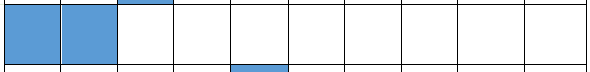
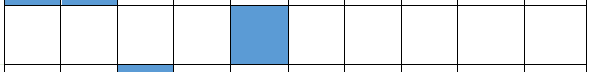
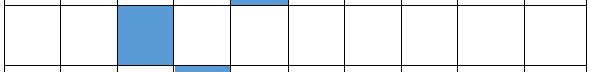
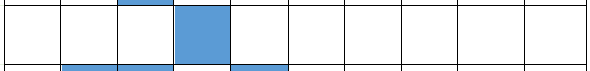
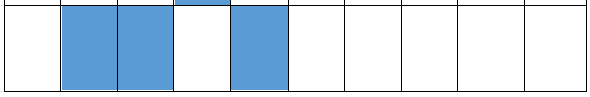
From all 169 proteins identified in this study, only 50 proteins were identified in both flows. Therefore, 79 proteins were identified with high confidence only at 0.25 ml/min, and 40 proteins only at 1.00 ml/min (Table 4.4).

**Table 4.4 Proteins identified in all 3 collection dates, for each specific time point, by the maintained flow rate.**

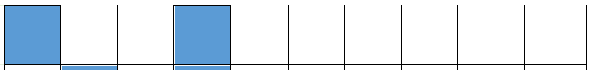
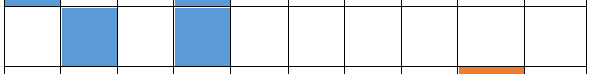
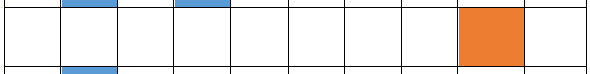
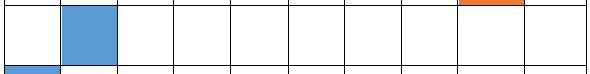
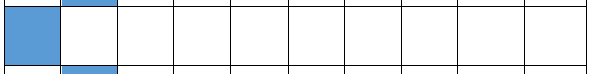
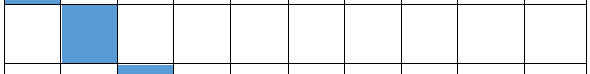
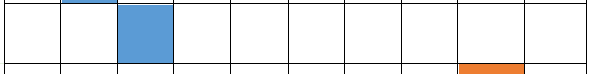
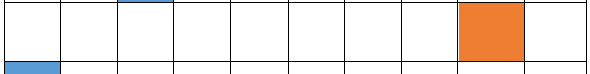
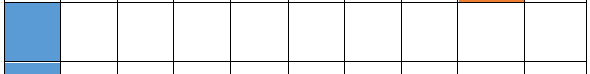
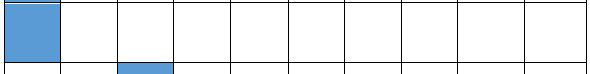
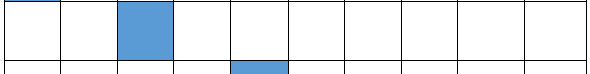
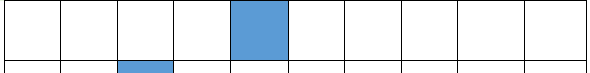
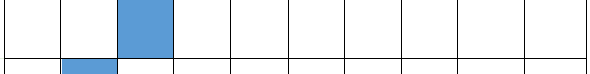
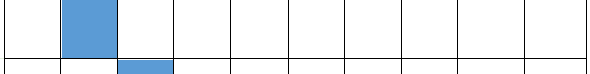
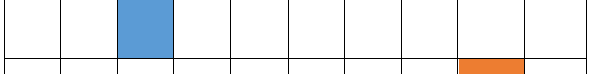
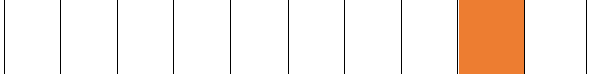
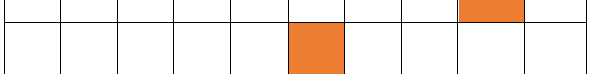
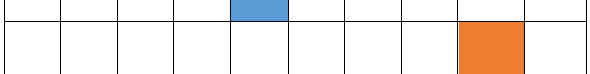
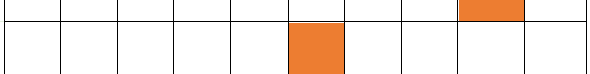
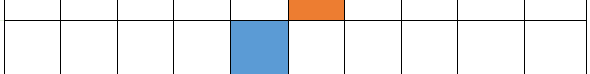
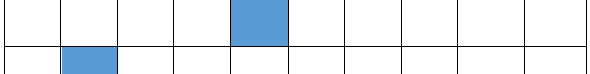
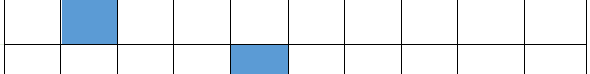
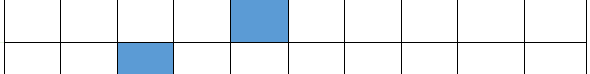
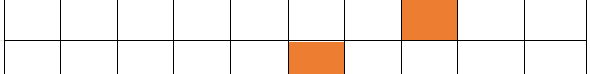
| | | 0.25 mL/min | | | | | 1.00 mL/min | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Accession** | **Protein names** | **T1** | **T5** | **T10** | **T20** | **T30** | **T1** | **T5** | **T10** | **T20** | **T30** |
| Q8IZF6 | Adhesion G-protein coupled receptor G4 | | | | | | | | | | |
| P04745 | Alpha-amylase 1A | | | | | | | | | | |
| Q8TCU4 | Alstrom syndrome protein 1 | | | | | | | | | | |
| Q12955 | Ankyrin-3 | | | | | | | | | | |
| P46013 | Antigen KI-67 | | | | | | | | | | |
| P02812 | Basic salivary proline-rich protein 2 | | | | | | | | | | |
| P23280 | Carbonic anhydrase 6 | | | | | | | | | | |
| Q9H799 | Ciliogenesis and planar polarity effector 1 | | | | | | | | | | |
| Q5QJE6 | Deoxynucleotidyltransferase terminal-interacting protein 2 | | | | | | | | | | |
| O75592 | E3 ubiquitin-protein ligase MYCBP2 | | | | | | | | | | |
| Q7Z6Z7 | E3 ubiquitin-protein ligase HUWE1 | | | | | | | | | | |
| Q5T4S7 | E3 ubiquitin-protein ligase UBR4 | | | | | | | | | | |
| Q86XX4 | Extracellular matrix organizing protein FRAS1 | | | | | | | | | | |
| Q5D862 | Filaggrin-2 | | | | | | | | | | |
| Q8NDA2 | Hemicentin-2 | | | | | | | | | | |
| O14607 | Histone demethylase UTY | | | | | | | | | | |
| Q03164 | Histone-lysine N-methyltransferase MLL | | | | | | | | | | |
| O60381 | HMG box-containing protein 1 | | | | | | | | | | |
| P13646 | Keratin, type I cytoskeletal 13 | | | | | | | | | | |
| P04264 | Keratin, type II cytoskeletal 1 | | | | | | | | | | |

| | | |
|---|---|---|
| Q14676 | Mediator of DNA damage checkpoint protein 1 |  |
| O95819 | Mitogen-activated protein kinase kinase kinase kinase 4 |  |
| Q9UKN1 | Mucin-12 |  |
| Q8WXI7 | Mucin-16 |  |
| Q685J3 | Mucin-17 |  |
| Q7Z5P9 | Mucin-19 |  |
| Q99102 | Mucin-4 |  |
| P98088 | Mucin-5AC |  |
| Q16718 | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5 |  |
| E9PAV3 | Nascent polypeptide-associated complex subunit alpha, muscle-specific form |  |
| O14513 | Nck-associated protein 5 |  |
| Q8IVL0 | Neuron navigator 3 |  |
| Q9Y618 | Nuclear receptor corepressor 2 |  |
| Q12830 | Nucleosome-remodeling factor subunit BPTF |  |
| Q96Q06 | Perilipin-4 |  |
| P37231 | Peroxisome proliferator-activated receptor gamma |  |
| Q53GL0 | Pleckstrin homology domain-containing family H member 2 |  |
| Q8WYP5 | Protein ELYS |  |
| Q5T8A7 | Protein phosphatase 1 regulatory subunit 26 |  |
| Q58EX2 | Protein sidekick-2 |  |
| Q8NET4 | Retrotransposon Gag-like protein 9 |  |
| O43166 | Signal-induced proliferation-associated 1-like protein 1 |  |
| Q9H7N4 | Splicing factor, arginine/serine-rich 19 |  |
| Q15772 | Striated muscle preferentially-expressed protein kinase |  |
| P02814 | Submaxillary gland androgen-regulated protein 3B |  |

| | | |
|---|---|---|
| O15061 | Synemin |  |
| Q8WZ42 | Titin |  |
| Q5T1R4 | Transcription factor HIVEP3 |  |
| Q2LD37 | Transmembrane protein KIAA1109 |  |
| Q15878 | Voltage-dependent R-type calcium channel subunit alpha |  |
| Q6P1M3 | LLGL scribble cell polarity complex component 2 |  |
| Q7RTP6 | [F-actin]-monooxygenase MICAL3 |  |
| P25054 | Adenomatous polyposis coli protein |  |
| O14514 | Adhesion G protein-coupled receptor B1 |  |
| Q8WXG9 | Adhesion G-protein coupled receptor V1 |  |
| Q01484 | Ankyrin-2 |  |
| P04114 | Apolipoprotein B-100 |  |
| O95831 | Apoptosis-inducing factor 1, mitochondrial |  |
| Q12797 | Aspartyl/asparaginyl beta-hydroxylase |  |
| Q8NFC6 | Biorientation of chromosomes in cell division protein 1-like |  |
| Q96DR5 | BPI fold-containing family A member 2 |  |
| Q7Z589 | BRCA2-interacting transcriptional repressor EMSY |  |
| Q9UBW5 | Bridging integrator 2 |  |
| Q9NSI6 | Bromodomain and WD repeat-containing protein 1 |  |
| P30622 | CAP-Gly domain-containing linker protein 1 |  |
| Q711Q0 | Cardiac-enriched FHL2-interacting protein |  |
| Q8N3K9 | Cardiomyopathy-associated protein 5 |  |
| Q8NAH9 | cDNA FLJ35331 fis, clone PROST2014659 |  |
| Q4G0X9 | Coiled-coil domain-containing protein 40 |  |
| P38432 | Coilin |  |
| Q8IZC6 | Collagen alpha-1(XXVII) chain |  |
| A8TX70 | Collagen alpha-5(VI) chain |  |

| | | |
|---|---|---|
| O60494 | Cubilin |  |
| Q8IXT1 | DNA damage-induced apoptosis suppressor protein |  |
| Q9UFH2 | Dynein heavy chain 17, axonemal |  |
| Q8TD57 | Dynein heavy chain 3, axonemal |  |
| Q8TE73 | Dynein heavy chain 5, axonemal |  |
| Q03001 | Dystonin |  |
| P49792 | E3 SUMO-protein ligase RanBP2 |  |
| Q9ULT8 | E3 ubiquitin-protein ligase HECTD1 |  |
| Q8NCN4 | E3 ubiquitin-protein ligase RNF169 |  |
| Q92616 | eIF-2-alpha kinase activator GCN1 |  |
| Q15723 | ETS-related transcription factor Elf-2 |  |
| Q8IYD8 | Fanconi anemia group M protein |  |
| P20930 | Filaggrin |  |
| Q2KHR3 | Glutamine and serine-rich protein 1 |  |
| Q5VVW2 | GTPase-activating Rap/Ran-GAP domain-like protein 3 |  |
| Q96RW7 | Hemicentin-1 |  |
| P15516 | Histatin-3 |  |
| Q9NR48 | Histone-lysine N-methyltransferase ASH1L |  |
| Q9H9B1 | Histone-lysine N-methyltransferase EHMT1 |  |
| P01889 | HLA class I histocompatibility antigen, B alpha chain |  |
| Q86YZ3 | Hornerin |  |
| P01834 | Immunoglobulin kappa constant |  |
| P15814 | Immunoglobulin lambda-like polypeptide 1 |  |
| Q9H792 | Inactive tyrosine-protein kinase PEAK1 |  |
| Q99665 | Interleukin-12 receptor subunit beta-2 |  |
| P22079 | Lactoperoxidase |  |
| P98164 | Low-density lipoprotein receptor-related protein 2 |  |

| | | |
|---|---|---|
| A6NES4 | Maestro heat-like repeat-containing protein family member 2A | |
| Q9NR99 | Matrix-remodeling-associated protein 5 | |
| Q9UPN3 | Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5 | |
| Q02505 | Mucin-3 | |
| Q9HC84 | Mucin-5B | |
| Q6W4X9 | Mucin-6 | |
| Q5VWP3 | Muscular LMNA-interacting protein | |
| Q6T4R5 | Nance-Horan syndrome protein | |
| P20929 | Nebulin | |
| Q92859 | Neogenin | |
| Q8NEY1 | Neuron navigator 1 | |
| Q96HA1 | Nuclear envelope pore membrane protein POM 121 | |
| Q6P4R8 | Nuclear factor related to kappa-B-binding protein | |
| P49790 | Nuclear pore complex protein Nup153 | |
| P48552 | Nuclear receptor-interacting protein 1 | |
| P80303 | Nucleobindin-2 | |
| C3PTT6 | Pancreatic adenocarcinoma upregulated factor | |
| O15018 | PDZ domain-containing protein 2 | |
| P55201 | Peregrin | |
| Q9NTG1 | Polycystic kidney disease and receptor for egg jelly-related protein | |
| P01833 | Polymeric immunoglobulin receptor | |
| Q9HCQ5 | Polypeptide N-acetylgalactosaminyltransferase 9 | |
| Q4VNC1 | Probable cation-transporting ATPase 13A4 | |
| Q9Y4D8 | Probable E3 ubiquitin-protein ligase HECTD4 | |
| P12273 | Prolactin-inducible protein | |
| Q07954 | Prolow-density lipoprotein receptor-related protein 1 | |

| | | |
|---|---|---|
| Q9UPA5 | Protein bassoon |  |
| Q5TBA9 | Protein furry homolog |  |
| Q9ULU4 | Protein kinase C-binding protein 1 |  |
| Q5VXU9 | Protein shortage in chiasmata 1 ortholog |  |
| O15027 | Protein transport protein Sec16A |  |
| Q9P2D8 | Protein unc-79 homolog |  |
| Q96QU1 | Protocadherin-15 |  |
| Q6V1P9 | Protocadherin-23 |  |
| A6NNC1 | Putative POM121-like protein 1-like |  |
| Q6WKZ4 | Rab11 family-interacting protein 1 |  |
| P46695 | Radiation-inducible immediate-early gene IEX-1 |  |
| Q9C0H5 | Rho GTPase-activating protein 39 |  |
| P55199 | RNA polymerase II elongation factor ELL |  |
| Q6ZP01 | RNA-binding protein 44 |  |
| Q92736 | Ryanodine receptor 2 |  |
| Q9C0I3 | Serine-rich coiled-coil domain-containing protein 1 |  |
| Q9UQ35 | Serine/arginine repetitive matrix protein 2 |  |
| P42345 | Serine/threonine-protein kinase mTOR |  |
| Q9Y6X0 | SET-binding protein |  |
| Q8N228 | Sex comb on midleg-like protein 4 |  |
| Q9NQ36 | Signal peptide, CUB and EGF-like domain-containing protein 2 |  |
| Q9Y6M7 | Sodium bicarbonate cotransporter 3 |  |
| Q9P2P6 | StAR-related lipid transfer protein 9 |  |
| P02808 | Statherin |  |
| P82094 | TATA element modulatory factor |  |
| Q9NT68 | Teneurin-2 |  |
| Q5TAX3 | Terminal uridylyltransferase 4 |  |
| Q9BXT5 | Testis expressed 15 |  |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| P31629 | Transcription factor HIVEP2 | | | | | | | | | | |
| O60675 | Transcription factor MafK | | | | | | | | | | |
| Q33E94 | Transcription factor RFX4 | | | | | | | | | | |
| Q8NEM7 | Transcription factor SPT20 homolog | | | | | | | | | | |
| O75962 | Triple functional domain protein | | | | | | | | | | |
| Q9UPU5 | Ubiquitin carboxyl-terminal hydrolase 24 | | | | | | | | | | |
| Q9UF83 | Uncharacterized protein DKFZp434B061 | | | | | | | | | | |
| Q92628 | Uncharacterized protein KIAA0232 | | | | | | | | | | |
| P46939 | Utrophin | | | | | | | | | | |
| O60281 | Zinc finger protein 292 | | | | | | | | | | |
| P15822 | Zinc finger protein 40 | | | | | | | | | | |
| Q96JG9 | Zinc finger protein 469 | | | | | | | | | | |
| Q9BZE0 | Zinc finger protein GLIS2 | | | | | | | | | | |
| O43149 | Zinc finger ZZ-type and EF-hand domain-containing protein 1 | | | | | | | | | | |
| P60852 | Zona pellucida sperm-binding protein 1 | | | | | | | | | | |
| Q96DA0 | Zymogen granule protein 16 homolog B | | | | | | | | | | |

For the proteins identified in both flows (n=50), variations were observed related to the time points in which they were found (Table 4.4), and to the frequency that they were identified in each flow rate, as shown in the heat map (Table 4.5); however, no pattern of secretion was determined. According to the data available in the HSPW on March of 2021, 21 of the 169 proteins identified in our study had been previously documented in the secretion from the PG, as shown in Table 4.5. According to the Bgee gene expression database, 42% (n=71) of all identified proteins had been previously described in PG, and 125 proteins, 78% of all identified proteins, were described as PG proteins in the Human Protein Atlas. A total of 42 proteins identified in our study were not described as PG proteins in any of the three databases.

**Table 4.5. Comparison of the frequency that the high-confidence proteins were found in the two studied flow rates (0.25 and 1.00 ml/min) with their annotated glandular origin according to three databases: the HSPW (data collected in March 2021), the Bgee (data collected in June 2022), and the HPA (data collected in June 2022).**

| Accession | Protein names | HSPW | Bgee | HPA | 0.25 mL/min | 1.00 mL/min |
|---|---|---|---|---|---|---|
| Q8IZF6 | Adhesion G-protein coupled receptor G4 | no | no | no | 2 | 1 |
| P04745 | Alpha-amylase 1A | Parotid, WS, SM, SL | Parotid, Minor, SM, SL | Parotid, Minor, SM, SL | 5 | 5 |
| Q8TCU4 | Alstrom syndrome protein 1 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 3 | 2 |
| Q12955 | Ankyrin-3 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 4 | 4 |
| P46013 | Antigen KI-67 | no | Minor | Minor, SM, SL | 2 | 2 |
| P02812 | Basic salivary proline-rich protein 2 | Parotid, WS, SM, SL | Minor, SM, SL | Parotid, Minor, SM, SL | 3 | 4 |
| P23280 | Carbonic anhydrase 6 | Parotid, WS, SM, SL | Parotid, Minor | Parotid, Minor, SM, SL | 5 | 5 |
| Q9H799 | Ciliogenesis and planar polarity effector 1 | WS | Minor | Parotid, Minor, SM, SL | 4 | 1 |
| Q5QJE6 | Deoxynucleotidyltransferase terminal-interacting protein 2 | no | Parotid, Minor | Parotid, Minor, SM, SL | 3 | 1 |
| Q7Z6Z7 | E3 ubiquitin-protein ligase HUWE1 | WS | Parotid | Parotid, Minor, SM, SL | 1 | 1 |
| O75592 | E3 ubiquitin-protein ligase MYCBP2 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 2 |
| Q5T4S7 | E3 ubiquitin-protein ligase UBR4 | WS | Minor | Parotid, Minor, SM, SL | 3 | 1 |
| Q86XX4 | Extracellular matrix organizing protein FRAS1 | WS | Minor | Parotid, Minor, SM, SL | 3 | 5 |
| Q5D862 | Filaggrin-2 | no | Minor | Minor | 4 | 1 |
| Q8NDA2 | Hemicentin-2 | WS | Minor, SM, SL | Minor, SM, SL | 4 | 1 |
| O14607 | Histone demethylase UTY | no | Minor | Parotid, Minor, SM, SL | 1 | 2 |
| Q03164 | Histone-lysine N-methyltransferase MLL | no | Parotid, Minor | Parotid, Minor, SM, SL | 4 | 2 |
| O60381 | HMG box-containing protein 1 | no | Minor | Parotid, Minor, SM, SL | 1 | 1 |
| P13646 | Keratin, type I cytoskeletal 13 | Parotid, WS, SM, SL | Minor | Parotid, Minor, SM, SL | 3 | 1 |
| P04264 | Keratin, type II cytoskeletal 1 | Parotid, WS, SM, SL | Minor | Parotid, Minor | 1 | 3 |
| Q14676 | Mediator of DNA damage checkpoint protein 1 | no | Minor | Parotid, Minor, SM, SL | 1 | 2 |
| O95819 | Mitogen-activated protein kinase kinase kinase kinase 4 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 1 |
| Q9UKN1 | Mucin-12 | no | Minor | no | 5 | 2 |
| Q8WXI7 | Mucin-16 | no | Minor | Minor | 5 | 5 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Q685J3 | Mucin-17 | no | no | no | 2 | 1 |
| Q7Z5P9 | Mucin-19 | WS | Minor | SM, SL | 5 | 4 |
| Q99102 | Mucin-4 | no | Minor | Parotid, Minor, SM, SL | 3 | 3 |
| P98088 | Mucin-5AC | Parotid, WS, SM, SL | Minor, SM, SL | Parotid, Minor, SL | 4 | 4 |
| Q16718 | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5 | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 1 |
| E9PAV3 | Nascent polypeptide-associated complex subunit alpha, muscle-specific form | WS, SM, SL | Parotid, Minor | Parotid, Minor, SM, SL | 3 | 2 |
| O14513 | Nck-associated protein 5 | no | Minor, SM, SL | Parotid, Minor, SM, SL | 2 | 1 |
| Q8IVL0 | Neuron navigator 3 | no | Minor, SM, SL | Parotid, Minor, SM, SL | 2 | 4 |
| Q9Y618 | Nuclear receptor corepressor 2 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 1 |
| Q12830 | Nucleosome-remodeling factor subunit BPTF | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 1 |
| Q96Q06 | Perilipin-4 | no | Parotid, Minor | Parotid, Minor, SM, SL | 5 | 4 |
| P37231 | Peroxisome proliferator-activated receptor gamma | WS | Minor | Parotid, Minor, SM, SL | 1 | 1 |
| Q53GL0 | Pleckstrin homology domain-containing family H member 2 | no | Minor | Parotid, Minor, SM, SL | 2 | 2 |
| Q8WYP5 | Protein ELYS | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 1 |
| Q5T8A7 | Protein phosphatase 1 regulatory subunit 26 | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 1 |
| Q58EX2 | Protein sidekick-2 | no | Minor, SM, SL | Minor, SM, SL | 1 | 2 |
| Q8NET4 | Retrotransposon Gag-like protein 9 | WS | no | no | 4 | 1 |
| O43166 | Signal-induced proliferation-associated 1-like protein 1 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 2 | 1 |
| Q9H7N4 | Splicing factor, arginine/serine-rich 19 | WS | Minor | Parotid, Minor, SM, SL | 1 | 1 |
| Q15772 | Striated muscle preferentially-expressed protein kinase | no | Parotid, Minor | Parotid, Minor, SM, SL | 2 | 3 |
| P02814 | Submaxillary gland androgen-regulated protein 3B | Parotid, WS, SM, SL | Parotid, Minor | Parotid, Minor, SM, SL | 5 | 5 |
| O15061 | Synemin | no | Parotid, Minor | Parotid, Minor, SM, SL | 2 | 1 |
| Q8WZ42 | Titin | Parotid, WS, SM, SL | Minor | Parotid, Minor, SM, SL | 3 | 5 |
| Q5T1R4 | Transcription factor HIVEP3 | no | Minor | Parotid, Minor, SM, SL | 3 | 2 |
| Q2LD37 | Transmembrane protein KIAA1109 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 1 |
| Q15878 | Voltage-dependent R-type calcium channel subunit alpha | no | Minor, SM, SL | no | 4 | 3 |
| Q6P1M3 | LLGL scribble cell polarity complex component 2 | no | Parotid, Minor | Parotid, Minor, SM, SL | 2 | 0 |

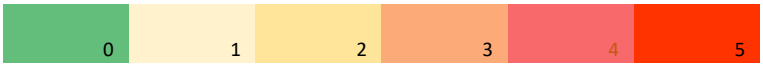
| | | | | | | |
|---|---|---|---|---|---|---|
| Q7RTP6 | [F-actin]-monooxygenase MICAL3 | WS | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P25054 | Adenomatous polyposis coli protein | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| O14514 | Adhesion G protein-coupled receptor B1 | no | Minor, SM, SL | Parotid, Minor | 1 | 0 |
| Q8WXG9 | Adhesion G-protein coupled receptor V1 | WS, SM, SL | Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q01484 | Ankyrin-2 | no | Minor | Parotid, Minor, SM, SL | 0 | 1 |
| P04114 | Apolipoprotein B-100 | WS | Minor | no | 1 | 0 |
| O95831 | Apoptosis-inducing factor 1, mitochondrial | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q12797 | Aspartyl/asparaginyl beta-hydroxylase | Parotid, WS, SM, SL | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q8NFC6 | Biorientation of chromosomes in cell division protein 1-like | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q96DR5 | BPI fold-containing family A member 2 | Parotid, WS, SM, SL | Parotid, Minor | Parotid, SM, SL | 0 | 1 |
| Q7Z589 | BRCA2-interacting transcriptional repressor EMSY | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9UBW5 | Bridging integrator 2 | WS | Minor, SM, SL | Parotid, Minor, SM, SL | 1 | 0 |
| Q9NSI6 | Bromodomain and WD repeat-containing protein 1 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P30622 | CAP-Gly domain-containing linker protein 1 | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q711Q0 | Cardiac-enriched FHL2-interacting protein | no | Minor, SM, SL | Parotid, Minor, SM, SL | 0 | 1 |
| Q8N3K9 | Cardiomyopathy-associated protein 5 | WS | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q8NAH9 | cDNA FLJ35331 fis, clone PROST2014659 | no | no | no | 0 | 1 |
| Q4G0X9 | Coiled-coil domain-containing protein 40 | WS | Minor | Minor, SM, SL | 0 | 1 |
| P38432 | Coilin | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q8IZC6 | Collagen alpha-1(XXVII) chain | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| A8TX70 | Collagen alpha-5(VI) chain | WS | Minor, SM, SL | no | 1 | 0 |
| O60494 | Cubilin | WS | Minor | Minor | 1 | 0 |
| Q8IXT1 | DNA damage-induced apoptosis suppressor protein | no | Minor | Minor, SM | 0 | 1 |
| Q9UFH2 | Dynein heavy chain 17, axonemal | WS | Minor | no | 2 | 0 |
| Q8TD57 | Dynein heavy chain 3, axonemal | WS | Minor, SM, SL | no | 0 | 1 |
| Q8TE73 | Dynein heavy chain 5, axonemal | WS | Minor | Minor, SL | 1 | 0 |
| Q03001 | Dystonin | WS | Parotid, Minor | Parotid, Minor, SM, SL | 2 | 0 |
| P49792 | E3 SUMO-protein ligase RanBP2 | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9ULT8 | E3 ubiquitin-protein ligase HECTD1 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Q8NCN4 | E3 ubiquitin-protein ligase RNF169 | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q92616 | eIF-2-alpha kinase activator GCN1 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q15723 | ETS-related transcription factor Elf-2 | no | Minor | Parotid, Minor, SM, SL | 2 | 0 |
| Q8IYD8 | Fanconi anemia group M protein | WS | Minor | Minor, SM, SL | 1 | 0 |
| P20930 | Filaggrin | Parotid, WS, SM, SL | Minor | Minor | 1 | 0 |
| Q2KHR3 | Glutamine and serine-rich protein 1 | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q5VVW2 | GTPase-activating Rap/Ran-GAP domain-like protein 3 | no | Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q96RW7 | Hemicentin-1 | WS | Minor | Minor, SM, SL | 1 | 0 |
| P15516 | Histatin-3 | Parotid, WS, SM, SL | Minor, SM, SL | Parotid, Minor, SM, SL | 0 | 1 |
| Q9NR48 | Histone-lysine N-methyltransferase ASH1L | WS | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q9H9B1 | Histone-lysine N-methyltransferase EHMT1 | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P01889 | HLA class I histocompatibility antigen, B alpha chain | no | Minor, SM, SL | Parotid, Minor, SM, SL | 0 | 1 |
| Q86YZ3 | Hornerin | Parotid, SM, SL | Minor, SM, SL | Minor | 2 | 0 |
| P01834 | Immunoglobulin kappa constant | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P15814 | Immunoglobulin lambda-like polypeptide 1 | no | Minor, SM, SL | no | 2 | 0 |
| Q9H792 | Inactive tyrosine-protein kinase PEAK1 | no | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q99665 | Interleukin-12 receptor subunit beta-2 | no | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P22079 | Lactoperoxidase | Parotid, WS, SM, SL | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P98164 | Low-density lipoprotein receptor-related protein 2 | WS | Parotid | Parotid | 3 | 0 |
| A6NES4 | Maestro heat-like repeat-containing protein family member 2A | WS | SM, SL | no | 1 | 0 |
| Q9NR99 | Matrix-remodeling-associated protein 5 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9UPN3 | Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 2 |
| Q02505 | Mucin-3 | no | Minor, SM, SL | Parotid, Minor, SM | 0 | 1 |
| Q9HC84 | Mucin-5B | Parotid, WS, SM, SL | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q6W4X9 | Mucin-6 | no | Minor, SM, SL | Minor, SM, SL | 0 | 1 |
| Q5VWP3 | Muscular LMNA-interacting protein | no | Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q6T4R5 | Nance-Horan syndrome protein | WS | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P20929 | Nebulin | WS | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q92859 | Neogenin | WS | Parotid, Minor, SM, SL | Parotid, Minor, SM, SL | 0 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Q8NEY1 | Neuron navigator 1 | no | Minor, SM, SL | Parotid, Minor, SM, SL | 2 | 0 |
| Q96HA1 | Nuclear envelope pore membrane protein POM 121 | no | Minor | Parotid, Minor, SM, SL | 2 | 0 |
| Q6P4R8 | Nuclear factor related to kappa-B-binding protein | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P49790 | Nuclear pore complex protein Nup153 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P48552 | Nuclear receptor-interacting protein 1 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P80303 | Nucleobindin-2 | Parotid, WS, SM, SL | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| C3PTT6 | Pancreatic adenocarcinoma upregulated factor | no | no | no | 0 | 5 |
| O15018 | PDZ domain-containing protein 2 | WS | Parotid, Minor | Parotid, Minor, SM | 1 | 0 |
| P55201 | Peregrin | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9NTG1 | Polycystic kidney disease and receptor for egg jelly-related protein | no | no | Minor | 0 | 1 |
| P01833 | Polymeric immunoglobulin receptor | Parotid, WS, SM, SL | Parotid, Minor | Parotid, Minor, SM, SL | 4 | 0 |
| Q9HCQ5 | Polypeptide N-acetylgalactosaminyltransferase 9 | no | Minor, SM, SL | Minor | 1 | 0 |
| Q4VNC1 | Probable cation-transporting ATPase 13A4 | no | Minor | Minor, SM, SL | 1 | 0 |
| Q9Y4D8 | Probable E3 ubiquitin-protein ligase HECTD4 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P12273 | Prolactin-inducible protein | Parotid, WS, SM, SL | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 5 |
| Q07954 | Prolow-density lipoprotein receptor-related protein 1 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 2 | 0 |
| Q9UPA5 | Protein bassoon | WS | no | Parotid, Minor, SL | 2 | 0 |
| Q5TBA9 | Protein furry homolog | WS | Parotid, Minor | Parotid, Minor, SM, SL | 2 | 0 |
| Q9ULU4 | Protein kinase C-binding protein 1 | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q5VXU9 | Protein shortage in chiasmata 1 ortholog | no | Minor | no | 1 | 0 |
| O15027 | Protein transport protein Sec16A | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9P2D8 | Protein unc-79 homolog | no | no | no | 1 | 0 |
| Q96QU1 | Protocadherin-15 | no | SM, SL | Parotid, SM | 1 | 0 |
| Q6V1P9 | Protocadherin-23 | no | Minor, SM, SL | Minor | 0 | 1 |
| A6NNC1 | Putative POM121-like protein 1-like | no | Minor | no | 1 | 0 |
| Q6WKZ4 | Rab11 family-interacting protein 1 | Parotid, WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P46695 | Radiation-inducible immediate-early gene IEX-1 | no | Minor, SM, SL | Parotid, Minor, SM, SL | 1 | 0 |
| Q9C0H5 | Rho GTPase-activating protein 39 | no | Minor | Minor | 1 | 0 |
| P55199 | RNA polymerase II elongation factor ELL | no | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q6ZP01 | RNA-binding protein 44 | no | Minor | Minor, SM, SL | 1 | 0 |
| Q92736 | Ryanodine receptor 2 | WS | Minor | Minor, SM | 1 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Q9C0I3 | Serine-rich coiled-coil domain-containing protein 1 | WS | Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q9UQ35 | Serine/arginine repetitive matrix protein 2 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| P42345 | Serine/threonine-protein kinase mTOR | no | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9Y6X0 | SET-binding protein | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q8N228 | Sex comb on midleg-like protein 4 | no | Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q9NQ36 | Signal peptide, CUB and EGF-like domain-containing protein 2 | no | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9Y6M7 | Sodium bicarbonate cotransporter 3 | no | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9P2P6 | StAR-related lipid transfer protein 9 | WS | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P02808 | Statherin | Parotid, WS, SM, SL | Minor, SM, SL | Parotid, Minor, SM, SL | 1 | 0 |
| P82094 | TATA element modulatory factor | WS | Parotid, Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q9NT68 | Teneurin-2 | no | Minor, SM, SL | Minor | 1 | 0 |
| Q5TAX3 | Terminal uridylyltransferase 4 | no | Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q9BXT5 | Testis expressed 15 | no | no | no | 0 | 1 |
| P31629 | Transcription factor HIVEP2 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 2 |
| O60675 | Transcription factor MafK | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q33E94 | Transcription factor RFX4 | no | Minor, SM, SL | no | 1 | 0 |
| Q8NEM7 | Transcription factor SPT20 homolog | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| O75962 | Triple functional domain protein | WS | Minor | Parotid, Minor, SM, SL | 2 | 0 |
| Q9UPU5 | Ubiquitin carboxyl-terminal hydrolase 24 | WS | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| Q9UF83 | Uncharacterized protein DKFZp434B061 | no | no | no | 1 | 0 |
| Q92628 | Uncharacterized protein KIAA0232 | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| P46939 | Utrophin | WS | Parotid, Minor | Parotid, Minor, SM, SL | 2 | 0 |
| O60281 | Zinc finger protein 292 | WS | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| P15822 | Zinc finger protein 40 | WS | Minor | Parotid, Minor, SM, SL | 1 | 0 |
| Q96JG9 | Zinc finger protein 469 | WS | Minor, SM, SL | Minor, SM | 1 | 0 |
| Q9BZE0 | Zinc finger protein GLIS2 | no | Minor, SM, SL | Parotid, Minor, SM, SL | 0 | 1 |
| O43149 | Zinc finger ZZ-type and EF-hand domain-containing protein 1 | no | Parotid, Minor | Parotid, Minor, SM, SL | 0 | 1 |
| P60852 | Zona pellucida sperm-binding protein 1 | no | Minor, SM, SL | no | 1 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| Q96DA0 | Zymogen granule protein 16 homolog B | Parotid, WS, SM, SL | Minor, SM, SL | Parotid, Minor, SM, SL | 3 0 |

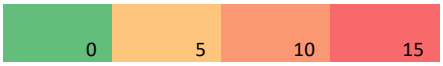



### 4.4.4 Functional analysis and clinical implications

Functional analysis using the PANTHER database showed that most identified molecular functions (binding (GO:0005488), catalytic activity (GO:0003824), molecular function regulator (GO:0098772), molecular transducer activity (GO:0060089), structural molecule activity (GO:0005198), and transporter activity (GO:0005215)) were predominant in the 0.25 ml/min flow (Table 4.6), except for the molecular adaptor activity (GO:0006009).

**Table 4.6 Functional distribution of proteins identified with high confidence, according to the proteins' Molecular Function, and the respective flow rate and time point of identification, based on the number of hits in the PANTHER database.**

| Flow rates | 0.25 ml/min | | | | | 1.00 ml/min | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Molecular function** | **T1** | **T5** | **T10** | **T20** | **T30** | **T1** | **T5** | **T10** | **T20** | **T30** |
| binding (GO:0005488) | 13 | 8 | 7 | 6 | 14 | 9 | 6 | 5 | 7 | 2 |
| catalytic activity (GO:0003824) | 6 | 6 | 2 | 4 | 8 | 4 | 5 | 6 | 4 | 4 |
| molecular adaptor activity (GO:0006009) | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| molecular function regulator (GO:0098772) | 5 | 2 | 2 | 1 | 4 | 2 | 2 | 1 | 3 | 1 |
| molecular transducer activity (GO:0060089) | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| structural molecule activity (GO:0005198) | 1 | 1 | 0 | 4 | 2 | 0 | 0 | 0 | 1 | 0 |
| transporter activity (GO:0005215) | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

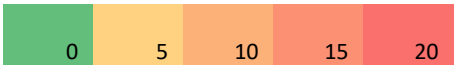



For their biological processes, proteins related to biological phase (GO:0044848), interaction between organisms (GO:0044419), locomotion (GO:0040011), reproduction (GO:0000003), and reproductive process (GO:0022414) were only identified at 0.25 ml/min (Table 4.7). In addition, more proteins related to biological regulation

(GO:0065007) and cellular process (GO:0009987) were found at 0.25 ml/min, mostly in T1 and T30.

**Table 4.7. Functional distribution of proteins identified with high confidence, according to proteins' Biological Processes, and the respective flow rate and time point of identification, based on the number of hits in the PANTHER database.**

| Flow rates | 0.25 ml/min | | | | | 1.00 ml/min | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Biological Process** | **T1** | **T5** | **T10** | **T20** | **T30** | **T1** | **T5** | **T10** | **T20** | **T30** |
| Biological adhesion (GO:0022610) | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| Biological phase (GO:0044848) | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Biological regulation (GO:0065007) | 12 | 8 | 7 | 7 | 12 | 8 | 7 | 5 | 7 | 4 |
| Cellular process (GO:0009987) | 21 | 15 | 11 | 13 | 19 | 9 | 11 | 8 | 9 | 3 |
| Developmental process (GO:0032502) | 5 | 4 | 3 | 4 | 9 | 1 | 1 | 1 | 2 | 3 |
| Immune system process (GO:0002376) | 1 | 1 | 1 | 0 | 0 | 1 | 2 | 1 | 1 | 1 |
| interspecies interaction between organisms (GO:0044419) | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Localization (GO:0051179) | 11 | 7 | 4 | 5 | 5 | 2 | 2 | 2 | 3 | 1 |
| Locomotion (GO:0040011) | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Metabolic process (GO:0008152) | 10 | 5 | 4 | 3 | 8 | 7 | 6 | 4 | 5 | 3 |
| Multicellular organismal process (GO:0032501) | 4 | 3 | 2 | 5 | 8 | 1 | 1 | 1 | 2 | 3 |
| Reproduction (GO:0000003) | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Reproductive process (GO:0022414) | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Response to stimulus (GO:0050896) | 5 | 4 | 2 | 2 | 4 | 1 | 2 | 1 | 2 | 1 |
| Signaling (GO:0023052) | 5 | 4 | 3 | 3 | 4 | 1 | 1 | 1 | 2 | 1 |

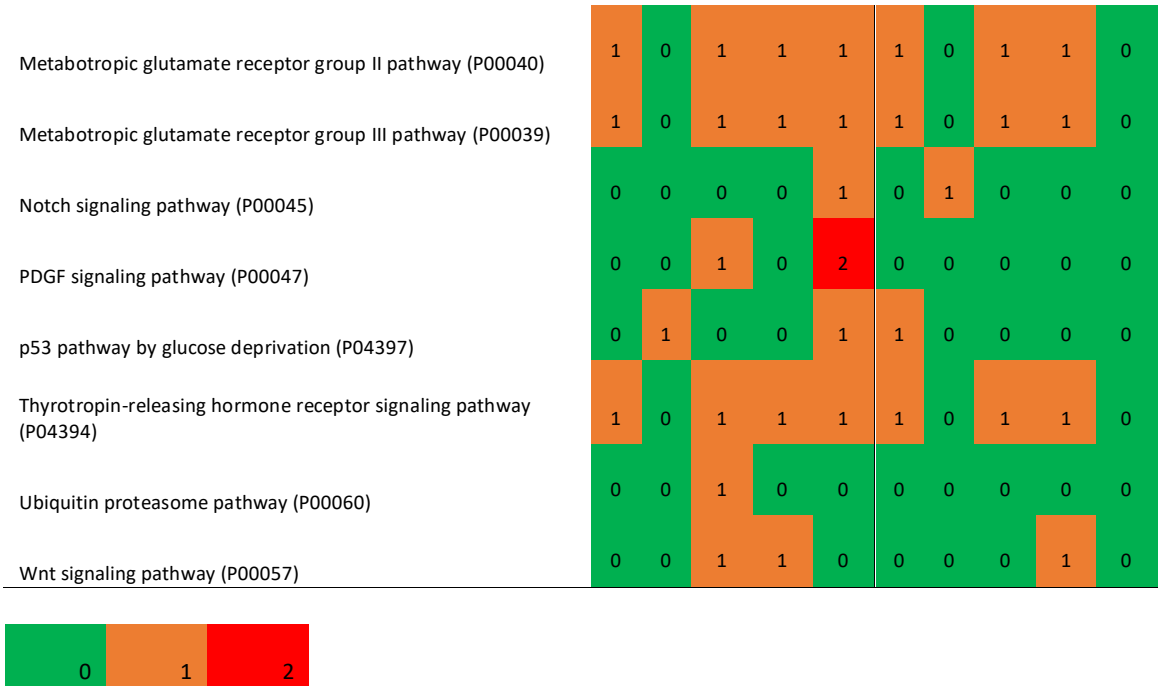



Proteins involved in 21 different Pathways were identified (Table 4.8). There were six pathways with similar distribution in both flows: Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway (P00027), Ionotropic glutamate receptor pathway (P00037), Metabotropic glutamate receptor group II pathway (P00040), Metabotropic glutamate receptor group III pathway (P00039), Thyrotropin-releasing

hormone receptor signaling pathway (P04394), and Wnt signaling pathway (P00057). Proteins related to 8 pathways were identified only at 0.25 ml/min: Alzheimer disease-presenilin pathway (P00004), Angiogenesis (P00005), Beta1 adrenergic receptor signaling pathway (P04377), Beta2 adrenergic receptor signaling pathway (P04378), Hypoxia response via HIF activation (P00030), Interleukin signaling pathway (P00036), PDGF signaling pathway (P00047), and Ubiquitin proteasome pathway (P00060). All other pathways were identified in both flows but at different time points.

**Table 4.8. Functional distribution of proteins identified with high confidence, according to the proteins' Pathway, and the respective flow rate and time point of identification, based on the number of hits in the PANTHER database.**

| | 0.25 ml/min | | | | | 1.00 ml/min | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Pathway | T1 | T5 | T10 | T20 | T30 | T1 | T5 | T10 | T20 | T30 |
| Alzheimer's disease-presenilin pathway (P00004) | 0 | 1 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Angiogenesis (P00005) | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Apoptosis signaling pathway (P00006) | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Beta1 adrenergic receptor signaling pathway (P04377) | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Beta2 adrenergic receptor signaling pathway (P04378) | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CCKR signaling pathway (P06959) | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Cadherin signaling pathway (P00012) | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Gonadotropin-releasing hormone receptor pathway (P06664) | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway (P00027) | 1 | 0 | 1 | 1 | 1 | 2 | 0 | 1 | 1 | 1 |
| Huntington's disease (P00029) | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| Hypoxia response via HIF activation (P00030) | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Interleukin signaling pathway (P00036) | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Ionotropic glutamate receptor pathway (P00037) | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |




The pathways found in our study related to cancers, cardiovascular diseases, and neurological diseases were listed in Table 4.9, along with the identified proteins and the corresponding time point and flow rate of identification.

**Table 4.9. Description of Pathways identified which might be related to cancers, cardiovascular diseases, and neurological diseases, their related proteins, and their respective flow and time point of identification.**

| Disease | Pathway | Panther Id | Protein name | Flow (ml/min) | Time (min) |
|---|---|---|---|---|---|
| **Cancer** | Angiogenesis (P00005) | PTHR12607: SF11 | Adenomatous polyposis coli protein | 0.25 | T20 |
| | Apoptosis signaling pathway (P00006) | PTHR43557: SF4 | Apoptosis-inducing factor 1, mitochondrial | 0.25 | T30 |
| | | PTHR48015: SF2 | Mitogen-activated protein kinase kinase kinase kinase 4 | 0.25 | T1 |
| | | | | 1.00 | T30 |
| | CCKR signaling map (P06959) | PTHR46399: SF7 | Ryanodine receptor 2 | 0.25 | T20 |
| | | PTHR24082: SF488 | Peroxisome proliferator-activated receptor gamma | 0.25 | T10 |
| | | | | 1.00 | T20 |
| | | PTHR16915: SF0 | Radiation-inducible immediate-early gene IEX-1 | 0.25 | T20 |

| | | | | | |
|---|---|---|---|---|---|
| | Cadherin signaling pathway (P00012) | PTHR24028: SF11 | Protocadherin-15 | 0.25 | T20 |
| | Hypoxia response via HIF activation (P00030) | PTHR11139: SF9 | Serine/threonine-protein kinase mTOR | 0.25 | T30 |
| | Interleukin signaling pathway (P00036) | PTHR23036: SF79 | Interleukin-12 receptor subunit beta-2 | 0.25 | T20 |
| | | PTHR11139: SF9 | Serine/threonine-protein kinase mTOR | 0.25 | T30 |
| | Notch signaling pathway (P00045) | PTHR13992: SF21 | Nuclear receptor corepressor 2 | 0.25 | T30 |
| | | | | 1.00 | T5 |
| | Wnt signaling pathway (P00057) | PTHR12607: SF11 | Adenomatous polyposis coli protein | | |
| | | PTHR24028: SF11 | Protocadherin-15 | 0.25 | T20 |
| **Cardiovascular Disease** | Angiogenesis (P00005) | PTHR12607: SF11 | Adenomatous polyposis coli protein | 0.25 | T20 |
| | Hypoxia response via HIF activation (P00030) | PTHR11139: SF9 | Serine/threonine-protein kinase mTOR | 0.25 | T30 |
| | Beta1 adrenergic receptor signaling pathway (P04377) | PTHR 46399: SF7 | Ryanodine receptor 2 | 0.25 | T20 |
| | Beta2 adrenergic receptor signaling pathway (P04378) | PTHR46399: SF7 | Ryanodine receptor 2 | 0.25 | T20 |
| | Notch signaling pathway (P00045) | PTHR13992: SF21 | Nuclear receptor corepressor 2 | 0.25 | T30 |
| | | | | 1.00 | T5 |
| **Alzheimer's disease** | Alzheimer disease-presenilin pathway (P00004) | PTHR24270: SF23 | Prolow-density lipoprotein receptor-related protein 1 | 0.25 | T10, T20 |
| | | PTHR22722: SF11 | Low-density lipoprotein receptor-related protein 2 | 0.25 | T5, T10, T30 |
| | Apoptosis signaling pathway (P00006) | PTHR43557: SF4 | Apoptosis-inducing factor 1, mitochondrial | 0.25 | T30 |
| | Notch signaling pathway (P00045) | PTHR13992: SF21 | Nuclear receptor corepressor 2 | 0.25 | T30 |
| | | | | 1.00 | T5 |
| | Cadherin signaling pathway (P00012) | PTHR24028: SF11 | Protocadherin-15 | 0.25 | T20 |
| | Ionotropic glutamate receptor pathway | PTHR45628: SF5 | Voltage-dependent R-type calcium channel subunit alpha-1E | 0.25 | T1, 10, 20, 30 |

| | | | | | |
|---|---|---|---|---|---|
| | (P00037) | | | 1.00 | T1, 10, 20 |
| | Ubiquitin proteasome pathway (P00060) | PTHR11254: SF67 | E3 ubiquitin-protein ligase HUWE1 | 0.25 | T10 |
| | | | | 1.00 | T1 |
| **Other** | Huntington's disease (P00029) | PTHR10676: SF344 | Dynein axonemal heavy chain 5 | 0.25 | T1 |
| | | PTHR13992: SF21 | Nuclear receptor corepressor 2 | 0.25 | T30 |
| | | | | 1.00 | T5 |
| | | PTHR10676: SF242 | Dynein axonemal heavy chain 3 | 1.00 | T10 |

## 4.5 Discussion

Our study demonstrates that not only the total protein concentration but also the proteome of the secretion from the PG varies over time under continuous flow rates, with effects of duration and intensity of stimulation present on most of the identified proteome.

Many factors can contribute to intra and inter-subject variability in salivary flow rates making it difficult to define a "normal" or "maximal" salivary flow rate [72–76]. Two of our volunteers were not able to maintain the 0.25 ml/min flow, despite repeated attempts, and they were excluded from the study. Dawes [48], in his study, also reported issues collecting PG secretion at 0.25 ml/min. Some of the subjects in his study repeated the collection procedure up to four times at 0.25 ml/min until they could maintain the constancy of flow rate, an issue not reported at higher flow rates. Additionally, intense gustatory stimulation (10% citric acid), for 30 minutes, provided large variations between minimal and maximum stimulated PG flow rates [76]. Dawes and Ong [73] reported inter-subject variability in the mean unstimulated parotid flow, as well as regarding the variations in the parotid flow in response to the circadian rhythms, which were reduced by maintaining a constant parotid flow of 1.00 ml/min [73].

Strong influences of intensity and duration of stimulation in the concentrations of total protein and many electrolytes in the secretion from the parotid glands under continuous flows have been reported [48]. Our results for the total protein concentration demonstrated significant differences within and between flow rates, dependent on the duration and the

intensity of the stimulation, respectively. Our results agree with the literature [48] demonstrating that, when a continuous flow rate is maintained, the total protein concentration of the secretion from the PG falls at the beginning of the stimulation, and it rises after the minimum is reached. Also, the fact that the minimum protein concentration was achieved in the second minute of collection (T2) at 1.00 ml/min, and only after 7 minutes of stimulated collection (T7) at 0.25 ml/min, suggests that not only the PG requires a period to respond to the stimulus and increase the protein secretion, but this response period is likely more dependent on the intensity than on the duration of the stimulus. Moreover, since the total volume secreted when the minimum protein concentration was recorded for 0.25 ml/min (1.75 ml) was similar to that at 1.00 ml/min (2.00 ml), a volume-based response mechanism for PG protein secretion should not be dismissed.

After observing that the total protein concentration secreted by the PG under different continuously stimulated flow rates for 15 consecutive minutes rose linearly [48], Dawes questioned whether the human PG contained a large storage of protein available for secretion or if the protein synthesis was sufficiently fast to replace continuously the protein secreted [48]. Our study extended the total protein concentration measurements, under continuous flows, for 30 consecutive minutes, but we were still not able to determine the mechanism by which the PG manages to continuously increase the protein output, which was higher at the more intense flow (1.00 ml/min) than at the less intense flow (0.25 ml/min). After observing high PG protein secretion for 55 minutes, Dawes [51] reported a gradual increase in PG total protein concentration during the first 15-20 min of stimulation, with an almost constant flow rate (0.96-1.06 ml/min), followed by a relative constant protein concentration, with a small tendency to decrease after secretion of 50-60 ml from one gland. In our study, even with the extended period of 30 minutes, the total protein concentration continued to rise until the end, suggesting that the protein secretory capacity of the PG was not exhausted. Nonetheless, the volume collected from one gland in our study did not exceed 30 ml.

Dawes also questioned if the proportion of different salivary proteins varied after prolonged stimulation of the salivary glands; however, he did not find significant differences between samples from the PG and SMG collected at the beginning and the end

of stimulation [51]. Our results demonstrate that the proteomic profile of the PG secretion varies over time, under continuous flow rates. Analysis of the proteins identified with high confidence in the salivary pools of selected time points, resulted in more proteins identified at the lower flow rate (0.25 ml/min), than at the higher flow (1.00 ml/min) in all examined time points, suggesting that the diversity of the PG secretion proteome reduces with the increase in the intensity of stimulation (higher flow rate). Although the average MW and pI of the proteins identified in the two flow rates were very similar, the percentage of proteins with MW below 56 kDa identified in all time points was higher at 1.00 ml/min, than at 0.25 ml/min, possibly because the proteins secreted at higher flow rates are either smaller in nature or they are more prone to proteolysis due to the higher secretion rates. On the other hand, the percentage of proteins with pI below 6.8 was greater than that above 6.8 in all time points, for both flows, with a slight increase in the percentage of acidic proteins secreted at 5 and 10 minutes of collection (T5 and T10) for both flows. The predominance of acidic proteins (majority pI between 4 and 8, and many of them with pI 4-5) was reported by Denny and collaborators [3] in salivary proteins collected from human major salivary glands; however, no differentiation was made in their study between each of the salivary glands.

Interestingly, only 50 proteins out of the 169 unique proteins identified in our study were common to both flow rates. The fact that 79 proteins were identified only at 0.25 ml/min, and 40 proteins were found only at 1.00 ml/min, suggests a strong dependence of the proteome on the intensity of stimulation (flow rate). Additionally, the distribution of the proteins identified with high confidence within each flow rate shows that most proteins were unique to a specific time point, indicating the effect of the duration of stimulation (time) on the proteome of the PG secretion. Interestingly, only 4 proteins were identified in all time points for both flow rates, suggesting that they are secreted constantly by the PG and that their secretion might not be affected by the duration or intensity of stimulation. On the other hand, mucin-12, mucin-19, and perilipin-4 were identified at every time point only at 0.25 ml/min; while extracellular matrix organizing protein FRAS1, pancreatic adenocarcinoma upregulated factor, prolactin-inducible protein, and titin were identified at every time point only at 1.00 ml/min, suggesting that their secretion depends on the

intensity of stimulation. Strong dependence of the proteome on the intensity of stimulation was shown by proteins that were not found at 1.00 ml/min, but were identified at 0.25 ml/min at 2 time points (Hornerin), 3 time points (Zymogen granule protein 16 homolog B), or 4 time points (Polymeric immunoglobulin receptor); and by the identification of Prolactin-inducible protein at every time point at 1.00 ml/min, but not found at 0.25 ml/min.

Furthermore, the expression of the 50 proteins common to both flows did not exhibit a common secretion pattern within or between flows. First, if intensity were the main factor influencing the proteome, the same protein would be identified in earlier time points in the more intense flow (1.00 ml/min), than in the less intense flow (0.25 ml/min). However, this predicted pattern was only true for 3 of the 50 common proteins, they were Basic salivary proline-rich protein-2, Pleckstrin homology domain-containing family H member 2, and Protein phosphatase 1 regulatory subunit 26. Second, if neither intensity nor duration affected the proteome, the expression of proteins common to both flows should also be at all time points. Interestingly, only 4 of the 50 common proteins were identified at the same time points for both flows, they were Alpha-amylase 1A, Carbonic anhydrase 6, Mucin-16, Submaxillary gland androgen-regulated protein 3B.. Finally, if effects of both intensity and duration of stimulation are accepted, the proteome of the secretion from the PG under continuous flow should present no specific secretion pattern since the expression of the proteins would be aleatory. The random distribution of the other 44 proteins that were common to both flows suggests that their secretion depends on both the intensity and the duration of stimulation. Thus, it appears that both intensity and duration of stimulation can affect most of the proteome from the PG; however, the secretion of some PG proteins might be affected differently by only one of the tested factors or by none of them. What appears to be a protein-specific secretion mechanism was suggested before [66–68]. Non-uniform variations in the secretion rates for lysozyme, lactoferrin, salivary peroxidase, and sIgA in stimulated parotid saliva suggested differences in the secretory mechanism of such antimicrobial PG proteins possibly due to different levels of PG activity and based on their site of origin, acinar cells, or intercalated duct cells [66–68].

Additionally, our intra and inter-flow rate comparisons of the PG proteome identified with high confidence in subsequent time points of one same flow, and between same time points of different flows, reinforces the effects of duration and intensity of stimulation on the secreted proteome. The effects of duration of stimulation are demonstrated by the large number of proteins that were not common to subsequent time points of one same flow; while effects of intensity of stimulation are observed by the proteins that were not common between the same time points of different flows. On the other hand, the proteins identified in the proteome of overlapped subsequent time points, of the same flow, suggest that their secretion suffered little or no influence of time (duration) in the observed period. Similarly, the overlapping of the proteome identified in the two flows, for one same time point, suggests little or no influence of the intensity of stimulation on the secretory mechanism of such proteins in that period. Therefore, our analysis reinforces a protein-specific secretory mechanism, where the secretion of different proteins may be affected differently by the duration and intensity of stimulation when the flow rate is maintained constant.

Our use of the recently created HSPW database indicated that approximately 12% (n=21) of the proteins identified in our study had been previously reported in the PG secretion. According to the gene expression database Bgee, approximately 42% (n=71) of our proteins originated in the PG, while the Human Protein Atlas database, which includes both gene and protein expression, indicated that approximately 78% (n=125) of our identified proteome had been identified in the PG or its secretion. Interestingly, around 25% (n=42) of the proteins identified in this study were not listed in any of the searched databases as expressed in the PG or its secretion. Therefore, to the best of our knowledge, this study identified 42 novel proteins in the PG secretion. The stratified collection adopted in our study, with two fixed flow rates, fixed time, and consequently fixed volume within each flow, allowed the expansion of the PG secretion proteome while highlighting the importance of collection protocol selection in proteome identification.

Moreover, 33 of the 42 potentially novel PG proteins were described in the minor salivary glands by at least one of the three databases that were consulted for this study. Therefore, participation of the secretion from the minor salivary glands in our final analyzed proteome should be further investigated as it could be inherent to the collection technique employed

in our study; however, our study cannot confirm that. Because the Lashley cup device extends a few millimeters around the parotid duct, some participation of the secretion from the minor salivary glands or oral mucosa is possible. Although it would be interesting to have the glandular secretion collected via cannulation to prevent any contamination, this method would be more complex and painful than the standardized collection using the Lashley cup device.

Furthermore, the analysis of the proteomic profile presented in this study was conducted with the strictest selection criteria based on the proteins identified in all biological triplicates (3 collections). Lowering the threshold to extend the analysis to include proteins found in samples identified in duplicates would certainly broaden our proteome discovery, but it would also reduce our confidence level, which is fundamental in investigative studies like this one.

The functional analysis of the proteins identified in each time point, and different flow rates, suggests that most molecular functions, biological processes, and related pathways of the parotid saliva also vary due to flow and time, and that the multi-functionality or functional redundancy of salivary proteins is unable to compensate for the observed proteomic variations.

Binding (GO:0005488) was the most identified molecular function at both flow rates. This molecular function term refers to the selective non-covalent interaction of a molecule with one or more specific sites on another molecule (QuickGo, https://www.ebi.ac.uk/QuickGO/term/GO:0005488 ). Many different interactions are included under the Binding annotation term, including protein and hydroxyapatite binding, and positive and negative regulation of interactions. More hits for Binding were observed in T1 and T30 at 0.25 ml/min, and at T1 and T20 at 1.00 ml/min, suggesting that proteins related to this molecular function are readily available immediately after stimulation (T1), maybe even stored or secreted in the absence of stimulation, and that their secretion is probably also affected by the duration of stimulation, as observed in their increased identification in later time points for both flows.

Among the identified biological processes, the Cellular process (GO:0009987) showed the largest number of hits in both flows. This biological process term refers to processes carried out at the cellular level, not necessarily restricted to a single cell, like cell communication and other cell cycle processes (QuickGO, https://www.ebi.ac.uk/QuickGO/term/GO:0009987 ). The trend observed for Binding, in the molecular function section, was also observed for the Cellular process at 0.25 ml/min, with most proteins identified at T1 and T30. This trend was not observed at 1.00 ml/min; however, the number of hits for Cellular process at 1.00 ml/min remained somewhat constant for the first 4 time points and fell to about 1/3 at T30, suggesting a change toward the end of the collection at 1.00ml/min that is possibly time-dependent and might suggest secretion limitation due to gland exhaustion.

Proteins can pass from the blood into the saliva via passive diffusion, active transport, or microfiltration [16,34,77,78], thus assuring the usefulness of saliva in the detection of biomarkers for systemic diseases. In our study, effects of flow and time were also observed on proteins involved in pathways related to cancers, cardiovascular diseases, and neurological diseases (Alzheimer's and Huntington's). While more than one pathway could be identified for the mentioned diseases, some related proteins were found in only one of the studied flows, and at specific time points; thus, alerting for the effects of time and flow on collection protocols aimed at disease biomarker discovery and identification.

Biomarkers for different cancers have been described in saliva, including oral [79–86], breast [87,88], pancreatic [89], ovarian [90], melanoma [91], gastric [92], and lung cancers [93]. A total of 8 pathways related to cancers were identified in our study, with the possible involvement of 10 different proteins. To the best of our knowledge, none of the 10 cancer-related proteins identified in our study have been investigated as possible salivary biomarkers for this disease [94].

Circulating biomarkers can provide an alternative diagnostic and monitoring assessment for cardiovascular disease risk and injury directly from blood tests, and, possibly, from saliva samples [95]. The use of point-of-care testing devices for heart failure analyzing both blood and saliva samples has been reported [96]. Our analysis identified 5 pathways relate to

heart physiology, with the identification of 4 unique proteins with recognized participation in cardiovascular disease, but no previous classification as salivary biomarkers for cardiovascular diseases [95].

Many biomarker candidates for Alzheimer's Disease (AD), the most prevalent neurodegenerative disease in the elderly [97], have been investigated [77,97–100]. Salivary lactoferrin has been shown to be negatively correlated with the severity of AD, and well correlated with cerebrospinal fluid biomarkers, promising to be an effective alternative to detect preclinical stages [77,99,100]. On the other hand, levels of statherin, histatin 1, cystatins, and S100s were elevated in the saliva of AD patients [97]. It is currently unknown how the concentrations of amyloid-β, tau, and other biomarkers vary between different salivary glands and WS, between stimulated and unstimulated secretion, or if their presence in the saliva is simply from blood [77]. A comprehensive analysis of changes in the salivary proteome due to physiological aging versus aging with AD will certainly provide strong arguments for the validation of salivary biomarkers for this disease [97]. A total of 7 proteins identified in our study participate in 6 pathways related to AD and might represent new avenues to be investigated in future studies about AD salivary biomarkers.

Since a correct characterization of the proteome of WS, and the secretion from each salivary gland, is critical for the successful utilization of such fluids for the diagnosis and monitoring of local and systemic diseases, our observations are extremely important factors to be considered when designing saliva collection protocols. Additionally, because the human PGs are the main contributors to the total WS volume under stimulation [1,5,73,101–104], one can infer that WS proteome is also dependent on the two factors tested in our study (flow rate and time); however, additional investigation is required to confirm this hypothesis.

## 4.6 Conclusions

Our results confirm that the PG requires a short interval after the start of the stimulation to increase its protein secretion. Therefore, samples collected at the first few minutes of stimulation, especially at higher flow rates (1.00 ml/min), may provide very different protein concentrations if compared to later minutes of the collection. This should be

considered when developing collection protocols based on the measurements of the total protein concentration of the PG secretion. Our results suggest that the PG can increase continuously the total concentration of proteins in its secretion for 30 minutes, even at high flow rates, without reaching exhaustion. Additionally, our results indicate that the PG secretion proteome is affected by both duration and intensity of the stimulation; however, not all proteins secreted by the PG present the same secretory mechanism. A protein-specific secretion was found for most of the 169 proteins identified in this study (n=165). More than two-thirds (n=119) of all proteins identified were strongly dependent on the intensity of stimulation, given by the two flow rates. Contrarily, very few proteins (n=4) did not demonstrate dependence on the intensity or duration of stimulation, being constantly secreted at both flows. This is the first study to demonstrate that the proteome of the PG secretion is affected by both intensity and duration of stimulation. As a result, the functional characteristics of the secreted PG saliva were also different within and between flow rates, with proteins related to different molecular functions, biological processes, and physiological pathways more likely to be identified at certain time points after the beginning of the stimulation, and at certain flow rates. Finally, this study highlights the importance of better understanding the physiologic changes in the proteomic composition of the PG secretion for the development of more robust evidence-based saliva collection protocols aimed at biomarker discovery and identification.

## 4.7 References

1. Edgar, W. M. Saliva: Its secretion, composition and functions. *Br Dent J* **172**, (1992).

2. Nagler, R. M., Hershkovich, O., Lischinsky, S., Diamond, E. & Reznick, A. Z. Saliva analysis in the clinical setting: Revisiting an underused diagnostic tool. *Journal of Investigative Medicine* **50**, 214–225 (2002).

3. Denny, P. *et al.* The proteomes of human parotid and submandibular/sublingual gland salivas collected as the ductal secretions. *J Proteome Res* **7**, (2008).

4. Rosa, N. *et al.* From the salivary proteome to the OralOme: Comprehensive molecular oral biology. *Arch Oral Biol* **57**, (2012).

5. Pedersen, A. M. L., Sørensen, C. E., Proctor, G. B., Carpenter, G. H. & Ekström, J. Salivary secretion in health and disease. *Journal of Oral Rehabilitation* vol. 45 (2018).

6. Valstar, M. H. *et al.* The tubarial salivary glands: A potential new organ at risk for radiotherapy. *Radiotherapy and Oncology* **154**, (2021).

7. Siqueira, W. L., Salih, E., Wan, D. L., Helmerhorst, E. J. & Oppenheim, F. G. Proteome of human minor salivary gland secretion. *J Dent Res* **87**, (2008).

8. Veerman, E. C. I., van den Keybus, P. A. M., Vissink, A. & Nieuw Amerongen, A. v. Human glandular salivas: Their separate collection and analysis. *Eur J Oral Sci* **104**, (1996).

9. Hardt, M. *et al.* Toward defining the human parotid gland salivary proteome and peptidome: Identification and characterization using 2D SDS-PAGE, ultrafiltration, HPLC, and mass spectrometry. *Biochemistry* **44**, (2005).

10. Hardt, M. *et al.* Assessing the effects of diurnal variation on the composition of human parotid saliva: Quantitative analysis of native peptides using iTRAQ reagents. *Anal Chem* **77**, (2005).

11. Ryu, O. H., Atkinson, J. C., Hoehn, G. T., Illei, G. G. & Hart, T. C. Identification of parotid salivary biomarkers in Sjögren's syndrome by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry and two-dimensional difference gel electrophoresis. *Rheumatology* **45**, (2006).

12. Neyraud, E., Sayd, T., Morzel, M. & Dransfield, E. Proteomic analysis of human whole and parotid salivas following stimulation by different tastes. *J Proteome Res* **5**, (2006).

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

14. Ramachandran, P. *et al.* Comparison of N-linked glycoproteins in human whole saliva, parotid, submandibular, and sublingual glandular secretions identified using hydrazide chemistry and mass spectrometry. *Clin Proteomics* **4**, (2008).

15. Gonzalez-Begne, M. *et al.* Proteomic analysis of human parotid gland exosomes by multidimensional protein identification technology (MudPIT). *J Proteome Res* **8**, (2009).

16. Jasim, H., Olausson, P., Hedenberg-Magnusson, B., Ernberg, M. & Ghafouri, B. The proteomic profile of whole and glandular saliva in healthy pain-free subjects. *Sci Rep* **6**, (2016).

17. Schulte, F., Hasturk, H. & Hardt, M. Mapping Relative Differences in Human Salivary Gland Secretions by Dried Saliva Spot Sampling and nanoLC–MS/MS. *Proteomics* **19**, (2019).

18. Esteves, C. V. *et al.* Diagnostic potential of saliva proteome analysis: A review and guide to clinical practice. *Braz Oral Res* **33**, (2019).

19. Saitou, M. *et al.* Functional Specialization of Human Salivary Glands and Origins of Proteins Intrinsic to Human Saliva. *Cell Rep* **33**, (2020).

20. Humphrey & Williamson, R. T. A review of saliva Normal composition, flow, and function. Humphrey, Williamson. 2001. Journal of Prosthetic Dentistry.pdf. *J Prosthet Dent* **85**, (2001).

21. Dawes, C. & Wong, D. T. W. Role of Saliva and Salivary Diagnostics in the Advancement of Oral Health. *J Dent Res* **98**, (2019).

22. Dawes, C. Circadian rhythms in human salivary flow rate and composition. *J Physiol* **220**, (1972).

23. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human submandibular saliva. *Arch Oral Biol* **19**, (1974).

24. Dawes, C. Physiological Factors Affecting Salivary Flow Rate, Oral Sugar Clearance, and the Sensation of Dry Mouth in Man. *J Dent Res* **66**, (1987).

25. Ship, J. A. & Fischer, D. J. The relationship between dehydration and parotid salivary gland function in young and older healthy adults. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* **52**, (1997).

26. Fortes, M. B., Diment, B. C., di Felice, U. & Walsh, N. P. Dehydration decreases saliva antimicrobial proteins important for mucosal immunity. *Applied Physiology, Nutrition and Metabolism* **37**, (2012).

27. Walsh, N. P., Montague, J. C., Callow, N. & Rowlands, A. v. Saliva flow rate, total protein concentration and osmolality as potential markers of whole body hydration status during progressive acute dehydration in humans. *Arch Oral Biol* **49**, (2004).

28. Inoue, H. *et al.* Gender difference in unstimulated whole saliva flow rate and salivary gland sizes. *Arch Oral Biol* **51**, (2006).

29. Ono, K. *et al.* Relationship of chewing-stimulated whole saliva flow rate and salivary gland size. *Arch Oral Biol* **52**, (2007).

30. Proctor, G. B. & Shaalan, A. M. Disease-Induced Changes in Salivary Gland Function and the Composition of Saliva. *Journal of Dental Research* vol. 100 (2021).

31. Villa, A. *et al.* World Workshop on Oral Medicine VI: a systematic review of medication-induced salivary gland dysfunction. *Oral Diseases* vol. 22 (2016).

32. Wolff, A. *et al.* A Guide to Medications Inducing Salivary Gland Dysfunction, Xerostomia, and Subjective Sialorrhea: A Systematic Review Sponsored by the World Workshop on Oral Medicine VI. *Drugs in R and D* vol. 17 (2017).

33. Smidt, D., Torpet, L. A., Nauntofte, B., Heegaard, K. M. & Pedersen, A. M. L. Associations between labial and whole salivary flow rates, systemic diseases and medications in a sample of older people. *Community Dent Oral Epidemiol* **38**, (2010).

34. Proctor, G. B. The physiology of salivary secretion. *Periodontology 2000* vol. 70 (2016).

35. Heintze, U., Birkhed, D. & Björn, H. Secretion rate and buffer effect of resting and stimulated whole saliva as a function of age and sex. *Swed Dent J* **7**, (1983).

36. Ligtenberg, A. J. M., Meuffels, M. & Veerman, E. C. I. Effects of environmental temperature on saliva flow rate and secretion of protein, amylase and mucin 5B. *Arch Oral Biol* **109**, (2020).

37. Engeland, C. G. *et al.* Psychological distress and salivary secretory immunity. *Brain Behav Immun* **52**, (2016).

38. Bosch, J. A. *et al.* Psychological stress as a determinant of protein levels and salivary- induced aggregation of Streptococcus gordonii in human whole saliva. *Psychosom Med* **58**, (1996).

39. Phillips, A. C. *et al.* Stressful life events are associated with low secretion rates of immunoglobulin a in saliva in the middle aged and elderly. *Brain Behav Immun* **20**, (2006).

40. Bosch, J. A. The use of saliva markers in psychobiology: Mechanisms and methods. *Monogr Oral Sci* **24**, (2014).

41. Walsh, N. P. *et al.* The effects of high-intensity intermittent exercise on saliva IgA, total protein and α-amylase. *J Sports Sci* **17**, (1999).

42. Ligtenberg, A., Liem, E., Brand, H. & Veerman, E. The Effect of Exercise on Salivary Viscosity. *Diagnostics* **6**, (2016).

43. Ligtenberg, A. J. M., Brand, H. S., van den Keijbus, P. A. M. & Veerman, E. C. I. The effect of physical exercise on salivary secretion of MUC5B, amylase and lysozyme. *Arch Oral Biol* **60**, (2015).

44. Hopkins, M. *et al.* Salivary lubricity (ex vivo) enhances upon moderate exercise: A pilot study. *Arch Oral Biol* **116**, (2020).

45. Dawes, C. Effects of Diet on Salivary Secretion and Composition. *J Dent Res* **49**, (1970).

46. Proctor, G. B. *et al.* The SALAMANDER project: SALivAry bioMarkers of mediterraneAN Diet associated with long-tERm protection against type 2 diabetes. *Nutr Bull* **42**, (2017).

47. Dawes, C. The effect of flow rate and length of stimulation on the protein concentration in human parotid saliva. *Arch Oral Biol* **12**, (1967).

48. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human parotid saliva. *Arch Oral Biol* **14**, (1969).

49. Dawes, C. & Jenkins, G. N. The effects of different stimuli on the composition of saliva in man. *J Physiol* **170**, (1964).

50. Watanabe, S. & Dawes, C. The effects of different foods and concentrations of citric acid on the flow rate of whole saliva in man. *Arch Oral Biol* **33**, (1988).

51. Dawes, C. The composition of human saliva secreted in response to a gustatory stimulus and to pilocarpine. *J Physiol* **183**, (1966).

52. Proctor, G. B. & Carpenter, G. H. Regulation of salivary gland function by autonomic nerves. *Autonomic Neuroscience: Basic and Clinical* vol. 133 (2007).

53. Ekström, J., Khosravani, N., Castagnola, M. & Messana, I. Saliva and the control of its secretion. in *Medical Radiology* vol. 0 (2019).

54. Pedersen, A. M. L., Sørensen, C. E., Dynesen, A. W. & Jensen, S. B. Salivary gland structure and functions and regulation of saliva secretion in health and disease. in *Salivary Glands Anatomy, Functions in Digestion and Role in Disease* (2013).

55. Dawes, C. & Chebib, F. S. The influence of previous stimulation and the day of the week on the concentrations of protein and the main electrolytes in human parotid saliva. *Arch Oral Biol* **17**, (1972).

56. Fischer, D. & Ship, J. A. Effect of age on variability of parotid salivary gland flow rates over time. *Age Ageing* **28**, (1999).

57. Tylenda, C. A., Ship, J. A., Fox, P. C. & Baum, B. J. Evaluation of Submandibular Salivary Flow Rate in Different Age Groups. *J Dent Res* **67**, (1988).

58. Ship, J. A., Nolan, N. E. & Puckett, S. A. Longitudinal analysis of parotid and submandibular salivary flow rates in healthy, different-aged adults. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* **50A**, (1995).

59. Heft, M. W. & Baum, B. J. Basic Biological Sciences Unstimulated and Stimulated Parotid Salivary Flow Rate in Individuals of Different Ages. *J Dent Res* **63**, (1984).

60. Baum, B. J. Clinical Science: Evaluation of Stimulated Parotid Saliva Flow Rate in Different Age Groups. *J Dent Res* **60**, (1981).

61. Pedersen, W. *et al.* Clinical Science Age-dependent Decreases in Human Submandibular Gland Flow Rates as Measured Under Resting and Post-stimulation Conditions. *J Dent Res* **64**, (1985).

62. Smith, C. H. *et al.* Effect of aging on stimulated salivary flow in adults. *J Am Geriatr Soc* **61**, (2013).

63. Thaysen, J. H., Thorn, N. A. & Schwartz, I. L. Excretion of sodium, potassium, chloride and carbon dioxide in human parotid saliva. *Am J Physiol* **178**, (1954).

64. Burgen, A. S. V. & Terroux, K. G. The effect of changes in the rate of flow of the saliva on the concentration of iodide in parotid saliva. *J Physiol* **163**, (1962).

65. Rayment, S. A., Liu, B., Offner, G. D., Oppenheim, F. G. & Troxler, R. F. Immunoquantification of human salivary mucins MG1 and MG2 in stimulated whole saliva: Factors influencing mucin levels. *J Dent Res* **79**, (2000).

66. Rudney, J. D. Implications of a model from olfactory research for the use of secretion rates in salivary studies. *Arch Oral Biol* **35**, (1990).

67. Rudney, J. D. Relationships between human parotid saliva lysozyme lactoferrin, salivary peroxidase and secretory immunoglobulin A in a large sample population. *Arch Oral Biol* **34**, (1989).

68. Rudney, J. D., Kajander, K. C. & Smith, Q. T. Correlations between human salivary levels of lysozyme, lactoferrin, salivary peroxidase and secretory immunoglobulin A with different stimulatory states and over time. *Arch Oral Biol* **30**, (1985).

69. Mandel, I. D. Sialochemistry in diseases and clinical situations affecting salivary glands. *Crit Rev Clin Lab Sci* **12**, (1980).

70. Lashley, K. S. Reflex secretion of the human parotid gland. *J Exp Psychol* **1**, (1916).

71. Satterthwaite, F. E. An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin* **2**, (1946).

72. Dawes, C. Factors influencing salivary flow rate and composition. in *Saliva and Oral Health* vol. 56 (2004).

73. Dawes, C. & Ong, B. Y. Circadian rhythms in the flow rate and proportional contribution of parotid to whole saliva volume in man. *Arch Oral Biol* **18**, (1973).

74. Ship, J. A. & Baum, B. J. Is reduced salivary flow normal in old people? *The Lancet* vol. 336 (1990).

75. Wu, A. J. & Ship, J. A. A characterization of major salivary gland flow rates in the presence of medications and systemic diseases. *Oral Surgery, Oral Medicine, Oral Pathology* **76**, (1993).

76. Wu, A. J., Baum, B. J. & Ship, J. A. Extended stimulated parotid and submandibular secretion in a healthy young and old population. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* **50A**, (1995).

77. Ashton, N. J., Ide, M., Zetterberg, H. & Blennow, K. Salivary Biomarkers for Alzheimer's Disease and Related Disorders. *Neurology and Therapy* vol. 8 (2019).

78. Farah, R. *et al.* Salivary biomarkers for the diagnosis and monitoring of neurological diseases. *Biomedical Journal* vol. 41 (2018).

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

80. Roi, A. *et al.* The challenges of OSCC diagnosis: Salivary cytokines as potential biomarkers. *Journal of Clinical Medicine* vol. 9 (2020).

81. Park, N. J. *et al.* Salivary microRNA: Discovery, characterization, and clinical utility for oral cancer detection. *Clinical Cancer Research* **15**, (2009).

82. St. John, M. A. R. *et al.* Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. *Archives of Otolaryngology - Head and Neck Surgery* **130**, (2004).

83. Li, Y. *et al.* Salivary transcriptome diagnostics for oral cancer detection. *Clinical Cancer Research* **10**, (2004).

84. Nagler, R., Bahar, G., Shpitzer, T. & Feinmesser, R. Concomitant analysis of salivary tumor markers - A new diagnostic tool for oral cancer. *Clinical Cancer Research* **12**, (2006).

85. Warnakulasuriya, S., Soussi, T., Maher, R., Johnson, N. & Tavassoli, M. Expression of p53 in oral squamous cell carcinoma is associated with the presence of IgG and IgA p53 autoantibodies in sera and saliva of the patients. *Journal of Pathology* **192**, (2000).

86. Balan, J. J., Rao, R. S., Premalatha, B. R. & Patil, S. Analysis of tumor marker CA 125 in saliva of normal and oral squamous cell carcinoma patients: A comparative study. *Journal of Contemporary Dental Practice* **13**, (2012).

87. Porto-Mascarenhas, E. C. *et al.* Salivary biomarkers in the diagnosis of breast cancer: A review. *Critical Reviews in Oncology/Hematology* vol. 110 (2017).

88. Koopaie, M., Kolahdooz, S., Fatahzadeh, M. & Manifar, S. Salivary biomarkers in breast cancer diagnosis: A systematic review and diagnostic meta-analysis. *Cancer Medicine* Preprint at https://doi.org/10.1002/cam4.4640 (2022).

89. Sturque, J., Berquet, A., Loison-Robert, L. S., Ahossi, V. & Zwetyenga, N. Interest of studying the saliva metabolome, transcriptome and microbiome in screening for pancreatic cancer. *Journal of Stomatology, Oral and Maxillofacial Surgery* vol. 120 (2019).

90. Lee, Y. H., Kim, J. H., Zhou, H., Kim, B. W. & Wong, D. T. Salivary transcriptomic biomarkers for detection of ovarian cancer: for serous papillary adenocarcinoma. *J Mol Med* **90**, (2012).

91. Gao, K. *et al.* Systemic disease-induced salivary biomarker profiles in mouse models of melanoma and non-small cell lung cancer. *PLoS One* **4**, (2009).

92. Xiao, H. *et al.* Differential Proteomic Analysis of Human Saliva using Tandem Mass Tags Quantification for Gastric Cancer Detection. *Sci Rep* **6**, (2016).

93. Zhang, L. *et al.* Development of transcriptomic biomarker signature in human saliva to detect lung cancer. *Cellular and Molecular Life Sciences* **69**, (2012).

94. Buzalaf, M. A. R. *et al.* Saliva as a diagnostic tool for dental caries, periodontal disease and cancer: is there a need for more biomarkers? *Expert Review of Molecular Diagnostics* vol. 20 (2020).

95. Rammos, A. *et al.* Review salivary biomarkers for diagnosis and therapy monitoring in patients with heart failure. A systematic review. *Diagnostics* vol. 11 (2021).

96. Tripoliti, E. E. *et al.* Point-of-Care Testing Devices for Heart Failure Analyzing Blood and Saliva Samples. *IEEE Rev Biomed Eng* **13**, (2020).

97. Contini, C. *et al.* Corrigendum: Top-Down Proteomics of Human Saliva Highlights Anti-inflammatory, Antioxidant, and Antimicrobial Defense Responses in Alzheimer Disease (Frontiers in Neuroscience, (2021), 15, (668852), 10.3389/fnins.2021.668852). *Frontiers in Neuroscience* vol. 15 (2021).

98. Ashton, N. J. *et al.* An update on blood-based biomarkers for non-Alzheimer neurodegenerative disorders. *Nature Reviews Neurology* vol. 16 (2020).

99. Tavares, E. *et al.* Potential Role of Aminoprocalcitonin in the Pathogenesis of Alzheimer Disease. *American Journal of Pathology* **186**, (2016).

100. Reale, M., Gonzales-Portillo, I. & Borlongan, C. V. Saliva, an easily accessible fluid as diagnostic tool and potent stem cell source for Alzheimer's Disease: Present and future applications. *Brain Research* vol. 1727 (2020).

101. Shannon, T. L. Parotid fluid flow rate as related to whole saliva volume. *Arch Oral Biol* **7**, (1962).

102. Shannon, I. L. & Chauncey, H. H. Hyperhydration and Parotid Flow in Man. *J Dent Res* **46**, (1967).

103. Dawes, C. & Wood, C. M. The contribution of oral minor mucous gland secretions to the volume of whole saliva in man. *Arch Oral Biol* **18**, (1973).

104. Sas, R. & Dawes, C. The intra-oral distribution of unstimulated and chewing-gum-stimulated parotid saliva. *Arch Oral Biol* **42**, (1997).

# Chapter 5

# 5 Conclusion and Future Directions

This Chapter revisits the overall aim of this thesis and summarizes the main results from the previous Chapters. Limitations encountered in each project are also discussed, together with potential solutions to strengthen our studies and ideas for future developments in the field.

## 5.1 Overview and Research Questions

There has been increasing interest in the use of many salivary components for clinical applications related to disease diagnostic and monitoring in the last decade [1]. The fact that the collection of human saliva is painless, and it is cheaper, easier, and faster than the collection of other biological fluids such as blood [2], for example, makes saliva very attractive. The composition of saliva is exceptionally rich. Most of the total volume of the saliva found in the oral cavity (whole saliva) derives from the secretion of the three major salivary glands (parotid, submandibular and sublingual glands) and numerous minor salivary glands (palatine, retromolar, buccal, labial, and lingual glands) [3]. In addition to the exocrine components, whole saliva (WS) contains elements from blood, cells and fluids from the oral mucosa, and many microorganisms and their metabolites [3]. Non-host-derived elements in saliva, such as viruses, are relatively easy to identify since the result is based on the presence or not of the agent's fingerprints [4,5]. Using salivary biomarkers generated by the host requires a full understanding of the normal salivary composition in healthy conditions and their physiological variations [1,6]. Therefore, fully understanding how the composition of the saliva changes in response to different factors is determinant for the elaboration of proper protocols and methods aimed at improving saliva's clinical applications [6]. Saliva's composition is affected by the stability of its components in response to the harsh oral environment [6,7]. Recent studies have suggested that the stability of some salivary proteins might be increased due to their interaction with other proteins [8]. Modulation and/or modification of the original individual function of salivary proteins when participating in heterotypic protein complexes have also been reported [8]. Moreover, changes in the organic and inorganic composition of the secretions from salivary glands in

response to different intensity and duration of stimulation have also been demonstrated and may limit the clinical use of certain salivary components [9–14]. Based on our search in the scientific literature, which focus on salivaomics [15], we conclude that salivary diagnostic should not rely on single biomarkers, but rather on panels of measurements for conclusive information, relaying on more than one single protein for disease diagnosis and possibly combining results from other omics investigations. For that, fully understanding the factors that affect the availability of salivary components, such as the formation of protein-protein complexes and the effect of different stimulations in the proteomic composition of saliva, are determinant building blocks in the search for salivary biomarkers and full clinical application of saliva for disease diagnosis.In attempt to fill knowledge gaps, this thesis aimed to address some problems through the following research questions: Could protein-protein-interaction (PPI) databases assist in the study of protein complexes by merging *in vitro* and *in-silico* partners of histatin 1? What proteins could be found forming complexes with amylase in WS? Is the proteome secreted by parotid gland (PG) affected by the intensity and duration of stimulation? This thesis sought to answer these questions through the following objectives:

1. Demonstrate the usefulness of the STRING database for studying PPIs in the histatin 1-protein network (Chapter 2)
2. Reveal the salivary proteins that interact with amylase forming heterotypic complexes in WS by using different proteomic approaches (Chapter 3)
3. Investigate the effects of intensity and duration of stimulation on the proteome of the secretion from the human PG under continuous flow rates (Chapter 4)

## 5.2 Summary

Results obtained brought about relevant information that might contribute to the pursuit of saliva as an effective diagnostic tool. Future studies can rely on our findings to help ascertain the composition of saliva.

When merging the *in-silico* and *in vitro* salivary protein complex partners using the STRING database [16], the research demonstrated the usefulness of the STRING database (http://string-db.org/) [17] for studying PPIs (Chapter 2). The *in-silico* approach was used to successfully perform a fast simulation of a novel constructed histatin 1-protein network,

including both the known and the predicted interactors, along with *in vitro* complex partners previously identified. Our study makes a significant contribution to the literature since it demonstrates the usefulness of the STRING database to integrate data about known and predicted protein-protein associations from many organisms, including both direct (physical) and indirect (functional) interactions, in an easy-to-use interface.

By revealing the amylase interactome in WS using proteomic approaches [18], we identified salivary proteins that interact with salivary amylase forming heterotypic complexes in WS, characterized the high-confidence interactors, and constructed a simulated novel amylase-protein network with the aid of bioinformatics tools (Chapter 3). In addition, we discussed the advantages and limitations of using three different proteomic approaches for protein identification: direct mass spectrometric analysis after tryptic digestion, and mass spectrometric analysis after sample separation using the two distinctive types of gel electrophoresis (SDS and Native PAGE) and in-gel tryptic digestion.

Last, investigating the effects of intensity and duration of stimulation on the proteome of the secretion from PG, we demonstrated that the proteome of secretion from the PG is affected by intensity and duration of stimulation when continuous flow rates are maintained (Chapter 4). The results from this preliminary study suggest that not only the total protein concentration of the secretion from the PG but also its proteome change due to the intensity of flow rate and the duration of stimulation, and that the secreted proteome may also depend on protein-specific secretory mechanisms. Results obtained reinforce the need for standardized protocols for saliva collection to assure the accuracy and reproducibility necessary for studies aimed at biomarker discovery and validation.

## 5.3 Limitations

In this section, the limitations discussed in Chapters 2 through 4 are summarized. Study-specific limitations are also presented in each Chapter's Discussion section.

Common limitations to biological studies, in general, are the physiological variation that can be found between different subjects and between samples collected from the same individual in different situations, and the limited number of participants [19]. Inter and intra-

subject variability is a common factor in salivary research [10,13,19–25]. To minimize these issues, all saliva collection was performed following a predefined protocol with attention to the time of collection. In addition, samples from different subjects in our studies were pooled to limit the individual variations.

To address general limitations related to the sample population, future works should aim to include a larger number of participants whenever possible. Although the use of pooled saliva samples is an alternative to reduce the effects of intra and inter-subject variability, the inclusion of more subjects in future studies, and the individual analysis of their saliva, can provide a more accurate panel of the salivary proteome and interactome of a population, as well as the identification of discrepancies that might need to be addressed individually [26–29]. This approach will not only provide more biological data but will also increase the power of the statistical analysis.

Since the purpose of this thesis was to gain a better understanding of aspects of the salivary proteome in healthy individuals, no participants with known systemic or oral diseases were included in the work here presented. The results described in this thesis can therefore only be compared to healthy cohorts and are not intended for the discovery of disease biomarkers.

Regarding the general limitations related to technical aspects of the work included in this thesis, employment of additional separation techniques such as gel electrophoresis can considerably improve the identification of less abundant proteins, as shown in the published work presented in Chapter 3 [30] and should be incorporated in the protocols aimed at exploration of the proteome whenever possible. Additionally, the inclusion of mass spectrometry (MS) techniques for quantitative measurements of proteins of interest [31] could provide valuable information about variations in the abundance of certain proteins in response to intensity and duration of stimulation [32–34].

Moreover, all studies included in this thesis were conducted using "shotgun" MS. A general limitation of this method is that it favors the identification of more abundant proteins in detriment of lower-abundance ones [31]. To reduce this issue, additional separation methods

like SDS and native gel electrophoresis, followed by band extraction and in-gel digestion, were used in Chapter 3 [30].

The primary limitation of Chapter 2 is that it is based on data that were analyzed and published at an earlier date. Since the database used for the identification of the proteins is constantly updated, some of the matching used for this publication might have missed newly identified and further characterized proteins. Ideally, the raw MS/MS data could be resubmitted to new protein matching. However, that does not interfere with the purpose of this work which was to demonstrate the use of the STRING web tool and to integrate *in vitro* and *in-silico* data. The STRING database itself is also constantly updated, highlighting the importance of documenting the date of the data acquisition.

To minimize the limitations of Chapter 3 various tests were conducted in our laboratory to determine the final optimal experimental design. However, it is not possible to assure that proteins with weaker binding affinity were not lost in the washing process, or that other proteins with affinity for starch were eluted along with the enriched amylase complex. To improve the accuracy of our results, selection criteria were implemented to include in the list of the amylase interactome only proteins that were identified in two or more of the three used approaches. Moreover, the results of this study do not provide information to determine which complex members interact directly with amylase, and which ones rely on intermediary proteins to participate in the heterotypic complex. Further studies looking at the binding affinity between amylase and some of the proteins identified in the amylase interactome should be able to determine the first shell interactors of this complex and validate our results. Additionally, the salivary pH was not determined at the time of the collection, a factor that can influence the interaction of different proteins in the complex formation. Future studies including healthy and diseased subjects could perform this additional measurement to determine if the salivary pH differs between the two conditions and if the salivary pH influences the amylase interactome. Finally, although the motivation for this study was the suggested new biological function of salivary amylase as a natural biological carrier for the distribution of other salivary proteins throughout the mouth, our results were not able to confirm this hypothesis, but they provided the substrate for future studies that can further investigate this proposition.

A limitation of Chapter 4 was that two participants could not maintain the lower flow rate (0.25 ml/min). Difficulty maintaining a 0.25 ml/min flow rate from the PG was reported in a similar study [11], while the higher flow rate (1.00 ml/min) was easily maintained by all participants. One can hypothesize that the lower flow was probably too close to the physiologic unstimulated flow of the participants that could not maintain secretion at 0.25 ml/min. In future studies investigating the effect of the intensity of stimulation on the PG secretion, flow rates higher than 0.25 ml/min might be advised. Additionally, because only pooled saliva samples were used for proteome identification, our study lacks the proteomic analysis of the saliva from each subject individually. Also, because only five selected time points were used for MS-based protein identification, our study can't provide information about the proteomic profile at each minute individually, contrary to the results for the total protein concentration measurements. As a result, a proper statistical analysis of characteristics of the proteome such as the number of identified proteins in each flow rate and in each time point, for example, could not be performed. In future studies investigating the effect of intensity and duration of stimulation on the PG proteome, it would be ideal to have the samples of each participant analyzed individually with MS for proteome identification instead of only pooled saliva, and to have the proteome identified in all separate 30 minutes instead of only five selected time points. Finally, this study was unable to investigate quantitative variations in the proteome from different time points and different flow rates. Future studies could include additional MS-based approaches to investigate changes in the abundance of certain proteins of interest. Nonetheless, the exploratory study presented in Chapter 4 provides strong information to support future studies to further characterize the changes in the PG proteome due to flow rate and time.

## 5.4 Future Work

Studies have demonstrated the importance of creating biobanks that can be used for biomarker research in different fields [35–38]. Ideally, a human saliva biobank could be created, with attention to standardized collection, handling, and storage protocols agreed upon by multiple research centers. Such biobanks could be assessed for both longitudinal and retrospective studies aimed at disease biomarker discovery. Meanwhile, future studies

may rely on our findings to investigate disease-specific markers by comparing interactome and proteome identified in this thesis with groups representative of specific diseases.

With respect to Chapter 2, the methodology described can be used by researchers of all levels, working on the interactome of any system, organism, or biological sample. The use of public PPI databases should be included in most interactomics studies as a fast and reliable alternative to reduce the data analysis time, provide up-to-date information from manually curated data, and assist in decision-making for subsequent analysis and study designs. Additionally, future works based on computational methods could improve the knowledge regarding the salivary interactome by modeling protein-protein complexes using public webservers (e.g., HADDOCK) [39] and conducting *in-silico* prediction of physical protein interactions using protein-protein prediction databases and prediction software [40].

Considering the methodology applied in Chapter 3, such as in-solution and in-gel digestion, SDS and native PAGE, and MS, we presented the first amylase-protein network in WS and demonstrated variations in the identified proteome inherent to the different approaches used in the study. While it was interesting for our purpose to have a comprehensive identification of *in vitro* partners by applying different analysis techniques, future studies should aim at documenting such variations through standardized protocols. Furthermore, our findings align with the proposed new function of amylase as a biological carrier. In the future, the novel amylase complex partners presented in this study can further be investigated to determine if the identified proteins interact direct or indirectly with amylase. Techniques like surface plasmon resonance (SPR) [41] and surface-enhanced Raman spectroscopy (SERS) [42–45] could be used in subsequent studies to verify if the amylase-protein network partners identified in our study possess direct interaction with amylase, and their binding affinities. Future studies should also aim at investigating if the interaction of amylase with the partners modulates their individual biological functions or generates new biological functions of the complex. This could be done via *in vitro* complex reconstruction and protein-specific biological assays, as demonstrated for the amylase-histatin 1 complex [8]. Also, since our results open new avenues to investigate amylase's protective function and delivery capabilities, the stability of all confirmed direct interactors

should be tested in future studies. Finally, peptidomics studies [46,47] could be done to determine if there are naturally occurring amylase peptides that could exhibit this same protective function and act as protein carriers, thus reducing the cost of possible therapeutic applications of amylase as a delivery system. Naturally occurring peptides of histatin 1 have already been shown to maintain their anti-fungal biological functions [48–52], but the functionality of amylase's proteoforms is still under exploration [53–55].

Moreover, findings presented in Chapter 4 demonstrated the effects of intensity and duration of stimulation on PG secreted proteome, and suggested protein-specific secretory mechanisms in response to the tested factors. Future studies could expand our work by including more subjects in their study design and analyzing the samples individually for proteomic compositions, instead of using only pooled saliva samples. Ideally, MS should also be performed for all individuals and all minutes included in the observation, instead of investigating a few selected time points only. We understand that MS-based analysis is quite expensive and time-consuming. These factors are often limitations in proteomics studies. Nonetheless, our work achieved its main goal of raising awareness about the importance of collection protocols on the final PG proteome. Future studies aimed at biomarker identification in the PG secretion can benefit from our preliminary protein lists as initial guidelines for their study design. Also, future works could compare the PG proteome for individual time points to that of pooled samples of a few subsequent time points, like minutes 1 to 5, for example, to simulate clinical practices. Our fractionated approach of collecting and analyzing the proteome of the PG secretion for specific time points under continuous flow rates was used here as a proof of concept of the influence of intensity and duration of stimulation on the PG proteome, but we understand that clinical applications of the PG proteome will likely use more simple protocols based on a fixed time or fixed volume.

## 5.5 Discussion and Conclusion

Many studies have improved our knowledge about salivary proteins in WS [6,7,23,30,56–60] and in secretions from the salivary glands [61–71], and how they compare to other body fluids [2,72–74]. However, the scientific exploration of the salivary proteome is still far from over. Robust

disease biomarkers can only be appointed in such salivary fluids when physiologic and pathologic variations in their composition can be differentiated undoubtedly. The lack of standardization in the protocols used by different research groups has been appointed as one of the reasons for contradictory results among salivary studies [65,75]. Nonetheless, other factors that affect the secretion, composition, and stability of the final salivary proteome are not yet fully comprehended [1]. Although saliva is a biological fluid with numerous biomarker candidate molecules, a definitive measurement for disease diagnosis is not yet available for most studied conditions. Our perception is that salivary diagnostic should focus not on single biomarkers, but panels of measurements for conclusive information and standardized protocols for saliva collection. Understanding the functional significance and diagnostic capabilities of WS and salivary gland secretions are greatly dependent on our ability to establish its composition. The work presented in this thesis applied different proteomics approaches to investigate two important factors that can impact the final proteome: the interaction among salivary proteins and the effect of intensity and duration of stimulation on the proteomic composition of the secretion from salivary glands. This thesis achieved the aim of enhancing the current understanding of the interactions between human salivary proteins and the physiological variations caused by the intensity and duration of the stimulation on the proteome of the human PG secretion. Our data contribute to the improvement of clinical applications of the salivary proteome and may be used as foundation knowledge for future studies. Our results open new avenues to be investigated regarding the study of PPI, the function of salivary proteins when participating in complexes, and the development of protocols for diagnosis and disease monitoring through saliva. Based on the interesting and novel information presented in this thesis, we conclude that the investigation of the salivary proteome must continue to clarify gaps in the knowledge about physiologic factors like PPIs and effects of stimulation on the secreted proteome, for example, which can affect saliva's function and composition and limit saliva's clinical applications.

## 5.6 References

1. Dawes, C. & Wong, D. T. W. Role of Saliva and Salivary Diagnostics in the Advancement of Oral Health. *J Dent Res* **98**, (2019).

2. Nagler, R. M., Hershkovich, O., Lischinsky, S., Diamond, E. & Reznick, A. Z. Saliva analysis in the clinical setting: Revisiting an underused diagnostic tool. *Journal of Investigative Medicine* **50**, 214–225 (2002).

3. Edgar, W. M. Saliva: Its secretion, composition and functions. *Br Dent J* **172**, (1992).

4. Siqueira, W. L., Moffa, E. B., Mussi, M. C. M. & de Andrade Moreira Machado, M. A. Zika virus infection spread through saliva – a truth or myth? *Braz Oral Res* **30**, (2016).

5. Tauro, L. B. *et al.* Potential use of saliva samples to diagnose Zika virus infection. *Journal of Medical Virology* vol. 89 (2017).

6. Helmerhorst, E. J., Dawes, C. & Oppenheim, F. G. The complexity of oral physiology and its impact on salivary diagnostics. *Oral Dis* **24**, (2018).

7. Helmerhorst, E. J. & Oppenheim, F. G. Saliva: A dynamic proteome. *Journal of Dental Research* vol. 86 (2007).

8. Siqueira, W. L., Lee, Y. H., Xiao, Y., Held, K. & Wong, W. Identification and characterization of histatin 1 salivary complexes by using mass spectrometry. *Proteomics* **12**, (2012).

9. Dawes, C. Factors Influencing Protein Secretion in Human Saliva. in *The environment of the teeth* vol. 3 125-137 (1981).

10. Dawes, C. The effect of flow rate and length of stimulation on the protein concentration in human parotid saliva. *Arch Oral Biol* **12**, (1967).

11. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human parotid saliva. *Arch Oral Biol* **14**, (1969).

12. Dawes, C. Stimulus effects on protein and electrolyte concentrations in parotid saliva. *J Physiol* **346**, (1984).

13. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human submandibular saliva. *Arch Oral Biol* **19**, (1974).

14. Dawes, C. Factors influencing salivary flow rate and composition. in *Saliva and Oral Health* vol. 56 (2004).

15. Wong, D. T. W. Salivaomics. *Journal of the American Dental Association* **143**, (2012).

16. Crosara, K. T. B., Moffa, E. B., Xiao, Y. & Siqueira, W. L. Merging in-silico and in vitro salivary protein complex partners using the STRING database: A tutorial. *J Proteomics* **171**, 87–94 (2018).

17. Szklarczyk, D. *et al.* The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**, (2017).

18. Crosara, K. T. B. *et al.* Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches. *Biomed Res Int* **2018**, (2018).

19. Ruhl, S. The scientific exploration of saliva in the post-proteomic era: From database back to basic function. *Expert Review of Proteomics* vol. 9 (2012).

20. Dawes, C. Physiological Factors Affecting Salivary Flow Rate, Oral Sugar Clearance, and the Sensation of Dry Mouth in Man. *J Dent Res* **66**, (1987).

21. Shannon, T. L. Parotid fluid flow rate as related to whole saliva volume. *Arch Oral Biol* **7**, (1962).

22. Dawes, C. The effects of flow rate and duration of stimulation on the concentrations of protein and the main electrolytes in human parotid saliva. *Arch Oral Biol* **14**, 277–294 (1969).

23. Jasim, H., Olausson, P., Hedenberg-Magnusson, B., Ernberg, M. & Ghafouri, B. The proteomic profile of whole and glandular saliva in healthy pain-free subjects. *Sci Rep* **6**, (2016).

24. Siqueira, W. L. & Dawes, C. The salivary proteome: Challenges and perspectives. *Proteomics - Clinical Applications* vol. 5 (2011).

25. Edgar, W., O'Mullane, D. & Dawes, C. *Saliva and oral health.* (British Dental Association, 2004).

26. Nedelkov, D., Kiernan, U. A., Niederkofler, E. E., Tubbs, K. A. & Nelson, R. W. Population Proteomics. *Molecular & Cellular Proteomics* **5**, (2006).

27. Nedelkov, D. Population proteomics: Addressing protein diversity in humans. *Expert Review of Proteomics* vol. 2 (2005).

28. Biron, D. G. *et al.* Population proteomics: An emerging discipline to study metapopulation ecology. *Proteomics* vol. 6 (2006).

29. Nedelkov, D. Population proteomics: Investigation of protein diversity in human populations. *Proteomics* vol. 8 Preprint at https://doi.org/10.1002/pmic.200700501 (2008).

30. Crosara, K. T. B. *et al.* Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches. *Biomed Res Int* **2018**, (2018).

31. Gillet, L. C., Leitner, A. & Aebersold, R. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annual Review of Analytical Chemistry* vol. 9 (2016).

32. Yates, J. R., Ruse, C. I. & Nakorchevsky, A. Proteomics by mass spectrometry: Approaches, advances, and applications. *Annual Review of Biomedical Engineering* vol. 11 (2009).

33. Angel, T. E. *et al.* Mass spectrometry-based proteomics: Existing capabilities and future directions. *Chem Soc Rev* **41**, (2012).

34. Calderón-Celis, F., Encinar, J. R. & Sanz-Medel, A. Standardization approaches in absolute quantitative proteomics with mass spectrometry. *Mass Spectrometry Reviews* vol. 37 (2018).

35. Nedelkov, D., Kiernan, U. A., Niederkofler, E. E., Tubbs, K. A. & Nelson, R. W. Population proteomics: The concept, attributes, and potential for cancer biomarker research. *Molecular and Cellular Proteomics* vol. 5 (2006).

36. Nedelkov, D., Kiernan, U. A., Niederkofler, E. E., Tubbs, K. A. & Nelson, R. W. Investigating diversity in human plasma proteins. *Proc Natl Acad Sci U S A* **102**, (2005).

37. Ryu, S. Y. *et al.* Detecting differential protein expression in large-scale population proteomics. *Bioinformatics* **30**, (2014).

38. Rees, B. B., Andacht, T., Skripnikova, E. & Crawford, D. L. Population proteomics: Quantitative variation within and among populations in cardiac protein expression. *Mol Biol Evol* **28**, (2011).

39. van Zundert, G. C. P. & Bonvin, A. M. J. J. Modeling protein–protein complexes using the HADDOCK webserver "modeling protein complexes with HADDOCK". *Methods in Molecular Biology* **1137**, (2014).

40. Kotlyar, M. *et al.* In silico prediction of physical protein interactions and characterization of interactome orphans. *Nat Methods* **12**, (2014).

41. Drescher, D. G., Drescher, M. J. & Ramakrishnan, N. A. Surface plasmon resonance (SPR) analysis of binding interactions of proteins in inner-ear sensory epithelia. *Methods in Molecular Biology* **493**, (2009).

42. Chen, C., Liu, W., Tian, S. & Hong, T. Novel surface-enhanced raman spectroscopy techniques for DNA, protein and drug detection. *Sensors (Switzerland)* vol. 19 (2019).

43. Siddhanta, S. & Narayana, C. Surface enhanced Raman spectroscopy of proteins: Implications for drug designing. *Nanomaterials and Nanotechnology* **2**, (2012).

44. Bunaciu, A. A., Aboul-Enein, H. Y. & Hoang, V. D. Raman spectroscopy for protein analysis. *Applied Spectroscopy Reviews* vol. 50 (2015).

45. Tuma, R. Raman spectroscopy of proteins: From peptides to large assemblies. *Journal of Raman Spectroscopy* **36**, (2005).

46. Vitorino, R. Digging Deep into Peptidomics Applied to Body Fluids. *Proteomics* vol. 18 (2018).

47. Trindade, F. *et al.* Salivary peptidomic as a tool to disclose new potential antimicrobial peptides. *J Proteomics* **115**, (2015).

48. Rothstein, D. M., Helmerhorst, E. J., Spacciapoli, P., Oppenheim, F. G. & Friden, P. Histatin-derived peptides: Potential agents to treat localised infections. *Expert Opinion on Emerging Drugs* vol. 7 (2002).

49. Oppenheim, F. G. *et al.* Histatins, a novel family of histidine-rich proteins in human parotid saliva. *J Biol Chem* **263**, (1988).

50. Zambom, C. R., da Fonseca, F. H. & Garrido, S. S. Bio- and nanotechnology as the key for clinical application of salivary peptide histatin: A necessary advance. *Microorganisms* vol. 8 (2020).

51. Khurshid, Z. *et al.* Histatin peptides: Pharmacological functions and their applications in dentistry. *Saudi Pharmaceutical Journal* vol. 25 (2017).

52. Ruissen, A. L. A. *et al.* Effects of histatin 5 and derived peptides on Candida albicans. *Biochemical Journal* **356**, (2001).

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

54. Contreras-Aguilar, M. D. *et al.* Variation of human salivary alpha-amylase proteoforms in three stimulation models. *Clin Oral Investig* **24**, (2020).

55. Contreras-Aguilar, M. D. *et al.* Changes occurring on the activity of salivary alpha-amylase proteoforms in two naturalistic situations using a spectrophotometric assay. *Biology (Basel)* **10**, (2021).

56. Neyraud, E., Sayd, T., Morzel, M. & Dransfield, E. Proteomic analysis of human whole and parotid salivas following stimulation by different tastes. *J Proteome Res* **5**, (2006).

57. Rayment, S. A., Liu, B., Offner, G. D., Oppenheim, F. G. & Troxler, R. F. Immunoquantification of human salivary mucins MG1 and MG2 in stimulated whole saliva: Factors influencing mucin levels. *J Dent Res* **79**, (2000).

58. Hirtz, C. *et al.* Complexity of the human whole saliva proteome. *J Physiol Biochem* **61**, (2005).

59. Xiuli, S., Erdjan, S., Oppenheim, F. G. & Helmerhorst, E. J. Kinetics of histatin proteolysis in whole saliva and the effect on bioactive domains with metal-binding, antifungal, and wound-healing properties. *The FASEB Journal* **23**, (2009).

60. Campese, M., Sun, X., Bosch, J. A., Oppenheim, F. G. & Helmerhorst, E. J. Concentration and fate of histatins and acidic proline-rich proteins in the oral environment. *Arch Oral Biol* **54**, (2009).

61. Denny, P. *et al.* The proteomes of human parotid and submandibular/sublingual gland salivas collected as the ductal secretions. *J Proteome Res* **7**, (2008).

62. Wu, S. *et al.* Quantitative analysis of human salivary gland-derived intact proteome using top-down mass spectrometry. *Proteomics* **14**, (2014).

63. Proctor, G. B. & Carpenter, G. H. The function of salivary proteins and the regulation of their secretion by salivary glands. *Biomedical Reviews* vol. 9 (1998).

64. Liu, B. *et al.* Expression of membrane-associated mucins MUC1 and MUC4 in major human salivary glands. *Journal of Histochemistry and Cytochemistry* **50**, (2002).

65. Saitou, M. *et al.* Functional Specialization of Human Salivary Glands and Origins of Proteins Intrinsic to Human Saliva. *Cell Rep* **33**, (2020).

66. Siqueira, W. L., Salih, E., Wan, D. L., Helmerhorst, E. J. & Oppenheim, F. G. Proteome of human minor salivary gland secretion. *J Dent Res* **87**, (2008).

67. Milne, R. W. & Dawes, C. The Relative Contributions of Different Salivary Glands to the Blood Group Activity of Whole Saliva in Humans. *Vox Sang* **25**, (1973).

68. Pedersen, A. M. L., Sørensen, C. E., Dynesen, A. W. & Jensen, S. B. Salivary gland structure and functions and regulation of saliva secretion in health and disease. in *Salivary Glands Anatomy, Functions in Digestion and Role in Disease* (2013).

69. Proctor, G. B. & Shaalan, A. M. Disease-Induced Changes in Salivary Gland Function and the Composition of Saliva. *Journal of Dental Research* vol. 100 (2021).

70. Schulte, F., Hasturk, H. & Hardt, M. Mapping Relative Differences in Human Salivary Gland Secretions by Dried Saliva Spot Sampling and nanoLC–MS/MS. *Proteomics* **19**, (2019).

71. Hardt, M. *et al.* Toward defining the human parotid gland salivary proteome and peptidome: Identification and characterization using 2D SDS-PAGE, ultrafiltration, HPLC, and mass spectrometry. *Biochemistry* **44**, (2005).

72. Lee, Y. H. & Wong, D. T. Saliva: An emerging biofluid for early detection of diseases. *American Journal of Dentistry* vol. 22 (2009).

73. Spielmann, N. & Wong, D. Saliva: Diagnostics and therapeutic perspectives. *Oral Diseases* vol. 17 (2011).

74. Yan, W. *et al.* Systematic comparison of the human saliva and plasma proteomes. *Proteomics Clin Appl* **3**, (2009).

75. Esteves, C. V. *et al.* Proteomic profile of saliva collected directly from ducts: a systematic review. *Clinical Oral Investigations* vol. 24 (2020).

# 6 Appendices

## 6.1 Appendix A: Health Science Research Ethics Board Approval


**Office of Research Ethics**

The University of Western Ontario

**Use of Human Subjects - Ethics Approval Notice**

**Principal Investigator:** Dr. W.L. Siqueira **Review Level:** Expedited
**Review Number:** 16181E **Revision Number:** 1
**Review Date:** March 24, 2010 **Approved Local # of Participants:** 100
**Protocol Title:** Composition, Structure and Function of Salivary Proteins
**Department and Institution:** Dentistry, University of Western Ontario
**Sponsor:** NSERC-NATURAL SCIENCES ENGINEERING RSRCH COU
**Ethics Approval Date:** April 20, 2010 **Expiry Date:** June 30, 2016
**Documents Reviewed and Approved:** Revised Letter of Information and Consent. Assent Letter.
**Documents Received for Information:**

This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/ICH Good Clinical Practice Practices: Consolidated Guidelines; and the applicable laws and regulations of Ontario has reviewed and granted approval to the above referenced revision(s) or amendment(s) on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.

The ethics approval for this study shall remain valid until the expiry date noted above assuming timely and acceptable responses to the HSREB's periodic requests for surveillance and monitoring information. If you require an updated approval notice prior to that time you must request it using the UWO Updated Approval Request Form.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the HSREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of monitor, telephone number). Expedited review of minor change(s) in ongoing studies will be considered. Subjects must receive a copy of the signed information/consent documentation.

Investigators must promptly also report to the HSREB:

a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to this office for approval.

Members of the HSREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the HSREB.

Chair of HSREB: Dr. Joseph Gilbert
FDA Reg. #: IRB 00000940

*This is an official document. Please retain the original in your files.*

cc: ORE File

UWO HSREB Ethics Approval - Revision
V.2008-07-01 (ApprovalNoticeHSREB_REV)

16181E

Page 1 of 1

Western Research

Research Ethics

**Western University Health Science Research Ethics Board**
**HSREB Annual Continuing Ethics Approval Notice**

**Date**: May 23, 2017
**Principal Investigator**: Dr. Walter Siqueira
**Department & Institution**: Schulich School of Medicine and Dentistry\Dentistry, Western University

**Review Type**: Delegated
**HSREB File Number**: 6251
**Study Title**: Composition, Structure and Function of Salivary Proteins
**Sponsor**: Natural Sciences and Engineering Research Council

**HSREB Renewal Due Date & HSREB Expiry Date**:
Renewal Due -2018/05/31
Expiry Date -2018/06/25

The Western University Health Science Research Ethics Board (HSREB) has reviewed the Continuing Ethics Review (CER) Form and is re-issuing approval for the above noted study.

The Western University HSREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Guideline for Good Clinical Practice (ICH E6 R1), the Ontario Freedom of Information and Protection of Privacy Act (FIPPA, 1990), the Ontario Personal Health Information Protection Act (PHIPA, 2004), Part 4 of the Natural Health Product Regulations, Health Canada Medical Device Regulations and Part C, Division 5, of the Food and Drug Regulations of Health Canada.

Members of the HSREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

[redacted]

Ethics Officer, on behalf of Dr. Joseph Gilbert, HSREB Chair
EO: Erika Basile ___ Grace Kelly ___ Katelyn Harris___ Nicola Morphet ___ Karen Gopaul ✓ Patricia Sargeant

**Western University, Research**, Support Services Bldg., Rm. 5150
London, ON, Canada N6G 1G9 t. 519.661.3036 f. 519.850.2466 www.uwo.ca/research/ethics

## 6.2 Appendix B: Permission for Reproduction of Scientific Articles of my Authorship

**Copyright Agreement for the *Journal of Proteomics* (Chapter 2)**

**Permissions Helpdesk** 
Inbox - Google March 31, 2022 at 12:22 PM 
Re: Copyright Permission Request [220330-020114]
To: Karla Tonelli Bicalho Crosara,
Reply-To: Permissions Helpdesk,
Resent-From: Karla Tonelli Bicalho Crosara

Dear Dr. Karla Crosara,

Thank you for contacting us.

Please note that, as one of the authors of this article, you retain the right to reuse it in your thesis/dissertation. You do not require formal permission to do so. You are permitted to post this Elsevier article online if it is embedded within your thesis.

Please feel free to contact me if you have any queries.

Have a great day ahead.

Kind regards,

[redacted]
Senior Copyrights Coordinator
ELSEVIER | HCM - Health Content Management

Visit Elsevier Permissions

---

**From:** Administrator
**Date:** Wednesday, March 30, 2022 04:08 PM GMT

Dear Customer

Thank you for contacting Elsevier's Permissions Helpdesk.

This is an automated acknowledgement to confirm we have received your query. Ticket number 220330-020114 has been opened on your behalf and we aim to respond within five business days.

Regards,

Permissions Helpdesk

---

**From:** Karla Crosara
**Date:** Wednesday, March 30, 2022 04:08 PM GMT

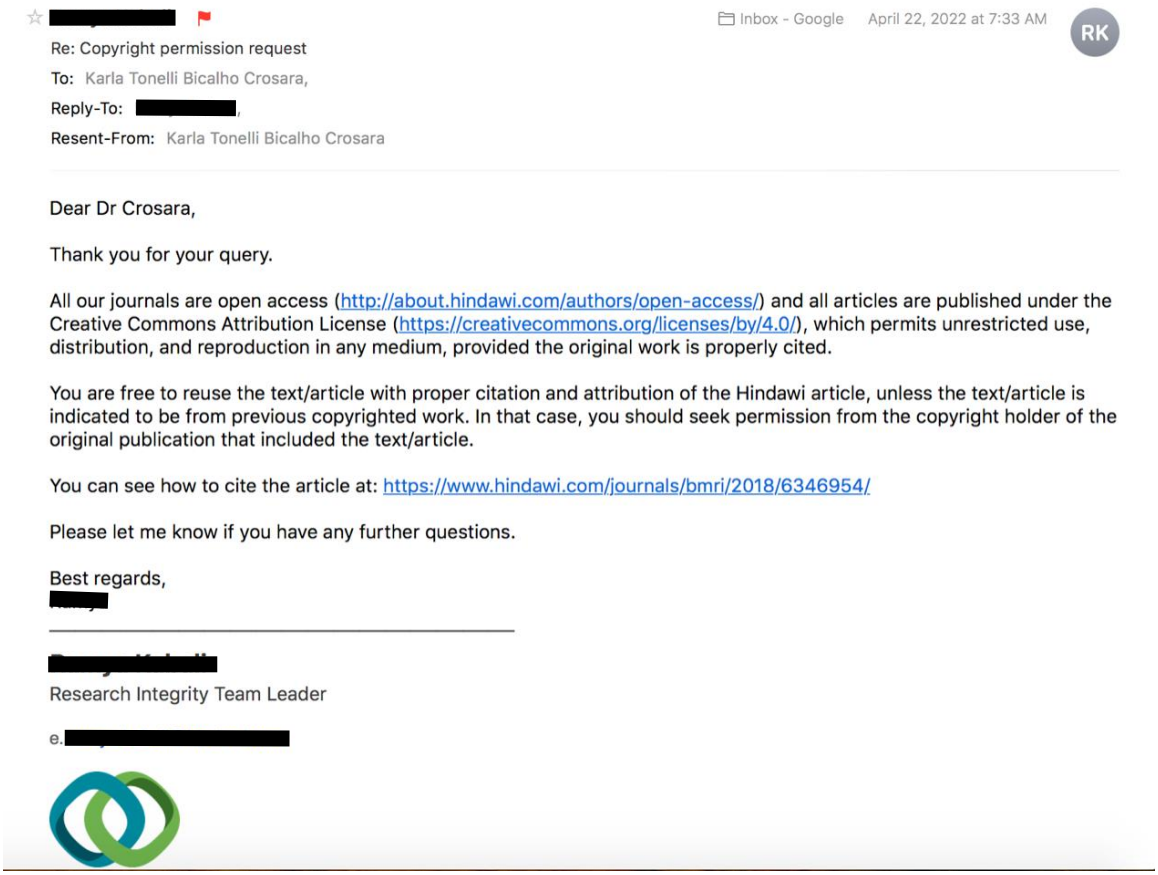
I am the first author in a paper published in this prestigious journal and I would like permission to include this publication as one of the chapters of my PhD Thesis. The publication title is: "Merging in-silico and in vitro salivary protein complex partners using the STRING database: A tutorial".

K.T.B. Crosara, E.B.Moffa, Y. Xiao, and W.L. Siqueira, "Merging in-silico and in vitro salivary protein complex partners using the STRING database: a tutorial", Journal of Proteomics, vol. 171, pp. 87-94, 2017.

Thank you,
Karla Tonelli Bicalho Crosara

---

## Copyright Agreement for the *Biomed Research International* (Chapter 3)



[redacted]

Inbox - Google April 22, 2022 at 7:33 AM

RK

Re: Copyright permission request

To: Karla Tonelli Bicalho Crosara,

Reply-To: [redacted]

Resent-From: Karla Tonelli Bicalho Crosara

Dear Dr Crosara,

Thank you for your query.

All our journals are open access (http://about.hindawi.com/authors/open-access/) and all articles are published under the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

You are free to reuse the text/article with proper citation and attribution of the Hindawi article, unless the text/article is indicated to be from previous copyrighted work. In that case, you should seek permission from the copyright holder of the original publication that included the text/article.

You can see how to cite the article at: https://www.hindawi.com/journals/bmri/2018/6346954/

Please let me know if you have any further questions.

Best regards,

[redacted]

Research Integrity Team Leader

e. [redacted]

Hindawi

Hindawi.com | Twitter | Facebook | LinkedIn | YouTube

On Wed, 6 Apr at 3:14 PM , Karla Tonelli Bicalho Crosara <[redacted]> wrote:

From: Karla Tonelli Bicalho Crosara
**Sent:** Wednesday, March 30, 2022 12:21 PM
**To:** [redacted]
**Subject:** Copyright permission request

I am the first author in a paper published in this prestigious journal and I would like permission to include this publication as one of the chapters of my PhD Thesis. The publication title is: "Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches". Article ID 6346954.

K.T.B. Crosara, D. Zuanazzi, E.B.Moffa, Y. Xiao, M.A.D.A.M. Machado and W.L. Siqueira, "Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches", BioMed Research International, 2018.

Thank you,
Karla Tonelli Bicalho Crosara

## 6.3 Appendix C: Permission for Reproduction of Scientific Articles from Other Authors (Figures and Tables in Chapter 1)

**Copyright Agreement for the *Oral Diseases* (Figure 1.1)**

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS
Sep 28, 2022

---

This Agreement between Western University -- Karla Tonelli Bicalho Crosara ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

| | |
|---|---|
| License Number | 5394251391671 |
| License date | Sep 22, 2022 |
| Licensed Content Publisher | John Wiley and Sons |
| Licensed Content Publication | Oral Diseases |
| Licensed Content Title | Saliva – the defender of the oral cavity |
| Licensed Content Author | ECI Veerman, AV Nieuw Amerongen |
| Licensed Content Date | Sep 5, 2008 |
| Licensed Content Volume | 8 |
| Licensed Content Issue | 1 |
| Licensed Content Pages | 11 |
| Type of use | Dissertation/Thesis |
| Requestor type | University/Academic |
| Format | Print and electronic |
| Portion | Figure/table |
| Number of figures/tables | 1 |
| Will you be translating? | No |
| Title | Expanding the knowledge about the interactions and the secretion of salivary proteins |
| Institution name | Western University |
| Expected presentation date | Nov 2022 |
| Portions | Figure 1 |
| | Western University<br>[redacted] |
| Requestor Location | [redacted]<br>Canada<br>Attn: Western University |
| Publisher Tax ID | EU826007151 |
| Total | 0.00 CAD |

## Copyright Agreement for the *Journal of Oral Rehabilitation* (Table 1.1)

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS
Sep 28, 2022

---

This Agreement between Western University -- Karla Tonelli Bicalho Crosara ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

| | |
|---|---|
| License Number | 5393710098785 |
| License date | Sep 21, 2022 |
| Licensed Content Publisher | John Wiley and Sons |
| Licensed Content Publication | Journal of Oral Rehabilitation |
| Licensed Content Title | Salivary secretion in health and disease |
| Licensed Content Author | J. Ekström, G. H. Carpenter, G. B. Proctor, et al |
| Licensed Content Date | Jun 25, 2018 |
| Licensed Content Volume | 45 |
| Licensed Content Issue | 9 |
| Licensed Content Pages | 17 |
| Type of use | Dissertation/Thesis |
| Requestor type | University/Academic |
| Format | Print and electronic |
| Portion | Figure/table |
| Number of figures/tables | 1 |
| Will you be translating? | No |
| Title | Expanding the knowledge about the interactions and the secretion of salivary proteins |

| | |
|---|---|
| Institution name | Western University |
| Expected presentation date | Nov 2022 |
| Portions | Table 1 |
| | Western University<br>[redacted] |
| Requestor Location | [redacted]<br>Canada<br>Attn: Western University |
| Publisher Tax ID | EU826007151 |
| Total | 0.00 CAD |

## Copyright Agreement for the *Journal of Oral Rehabilitation* (Figure 2.1)

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS
Sep 28, 2022

---

This Agreement between Western University -- Karla Tonelli Bicalho Crosara ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

| | |
|---|---|
| License Number | 5393701432055 |
| License date | Sep 21, 2022 |
| Licensed Content Publisher | John Wiley and Sons |
| Licensed Content Publication | Journal of Oral Rehabilitation |
| Licensed Content Title | Salivary secretion in health and disease |
| Licensed Content Author | J. Ekström, G. H. Carpenter, G. B. Proctor, et al |
| Licensed Content Date | Jun 25, 2018 |
| Licensed Content Volume | 45 |
| Licensed Content Issue | 9 |
| Licensed Content Pages | 17 |
| Type of use | Dissertation/Thesis |
| Requestor type | University/Academic |
| Format | Print and electronic |
| Portion | Figure/table |
| Number of figures/tables | 1 |
| Will you be translating? | No |
| Title | Expanding the knowledge about the interactions and the secretion of salivary proteins |
| Institution name | Western University |
| Expected presentation date | Nov 2022 |
| Portions | Figure 1 |
| Requestor Location | Western University<br>[redacted]<br>[redacted]<br>Canada<br>Attn: Western University |
| Publisher Tax ID | EU826007151 |
| Total | 0.00 CAD |

## Copyright Agreement for the *Periodontology 2000* (Figure 3.1)

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS
Sep 30, 2022

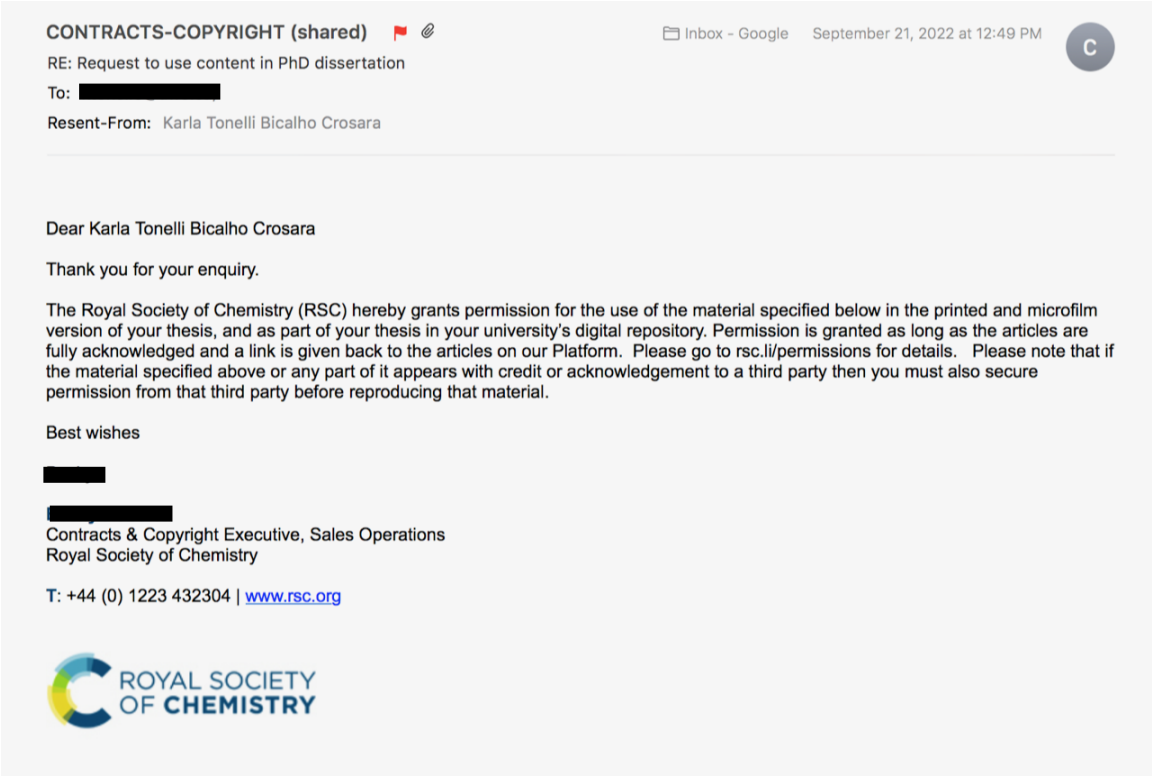
---

This Agreement between Western University -- Karla Tonelli Bicalho Crosara ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

| | |
|---|---|
| License Number | 5398850439505 |
| License date | Sep 30, 2022 |
| Licensed Content Publisher | John Wiley and Sons |
| Licensed Content Publication | Periodontology 2000 |
| Licensed Content Title | The physiology of salivary secretion |
| Licensed Content Author | Gordon B. Proctor |
| Licensed Content Date | Dec 10, 2015 |
| Licensed Content Volume | 70 |
| Licensed Content Issue | 1 |
| Licensed Content Pages | 15 |
| Type of use | Dissertation/Thesis |
| Requestor type | University/Academic |
| Format | Print and electronic |
| Portion | Figure/table |
| Number of figures/tables | 1 |
| Will you be translating? | No |
| Title | Expanding the knowledge about the interactions and the secretion of salivary proteins |

| | |
|---|---|
| Institution name | Western University |
| Expected presentation date | Nov 2022 |
| Portions | Figure 3 |
| | Western University<br>[redacted] |
| Requestor Location | [redacted]<br>Canada<br>Attn: Western University |
| Publisher Tax ID | EU826007151 |
| Total | 0.00 CAD |
| Terms and Conditions | |

## Copyright Agreement for the *Chemical Society Reviews* (Figure 4.1)



**CONTRACTS-COPYRIGHT (shared)** Inbox - Google September 21, 2022 at 12:49 PM

RE: Request to use content in PhD dissertation

To: [redacted]

Resent-From: Karla Tonelli Bicalho Crosara

Dear Karla Tonelli Bicalho Crosara

Thank you for your enquiry.

The Royal Society of Chemistry (RSC) hereby grants permission for the use of the material specified below in the printed and microfilm version of your thesis, and as part of your thesis in your university's digital repository. Permission is granted as long as the articles are fully acknowledged and a link is given back to the articles on our Platform. Please go to rsc.li/permissions for details. Please note that if the material specified above or any part of it appears with credit or acknowledgement to a third party then you must also secure permission from that third party before reproducing that material.

Best wishes

[redacted]

[redacted]
Contracts & Copyright Executive, Sales Operations
Royal Society of Chemistry

T: +44 (0) 1223 432304 | www.rsc.org

ROYAL SOCIETY OF CHEMISTRY

-----Original Message-----
From: [redacted]
Sent: 21 September 2022 16:23
To: CONTRACTS-COPYRIGHT (shared) <[redacted]>
Subject: Request to use content in PhD dissertation

Name: Karla Tonelli Bicalho Crosara
Message: My name is Karla Tonelli Bicalho Crosara, and I am a PhD candidate at the Department of Medical Biophysics of the Western University, in London, Ontario, Canada. I am supervised by Drs. Walter Siqueira, Charles McKenzie and Liliani Vieira, and I will be submitting by doctoral thesis soon which is about the Salivary Proteome.

Most of my research was based on bottom-up mass spectrometry. Therefore, I would like to request your permission to use in the Introduction chapter of my thesis the Figure 1, from the publication titled "Mass spectrometry-based proteomics: existing capabilities and future directions", co-authored by Thomas E. Angel, Uma K. Aryal, Shawna M. Hengel, Erin S. Baker, Ryan T. Kelly, Errol W. Robinson and Richard D. Smith. (Cite: Chem. Soc. Rev., 2012, 41, 3912–3928)

I look forward to your response.

Best regards,
Karla Crosara

## Copyright Agreement for the *Archives of Oral Biology* (Figure 5.1)

ELSEVIER LICENSE
TERMS AND CONDITIONS
Oct 14, 2022

---

This Agreement between Western University -- Karla Tonelli Bicalho Crosara ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

| | |
|---|---|
| License Number | 5406111446957 |
| License date | Oct 11, 2022 |
| Licensed Content Publisher | Elsevier |
| Licensed Content Publication | Archives of Oral Biology |
| Licensed Content Title | From the salivary proteome to the OralOme: Comprehensive molecular oral biology |
| Licensed Content Author | Nuno Rosa Maria José Correia Joel P. Arrais Pedro Lopes José Melo José Luís Oliveira Marlene Barros |
| Licensed Content Date | Jul 1, 2012 |
| Licensed Content Volume | 57 |
| Licensed Content Issue | 7 |
| Licensed Content Pages | 12 |
| Start Page | 853 |
| End Page | 864 |
| Type of Use | reuse in a thesis/dissertation |
| Portion | figures/tables/illustrations |
| Number of figures/tables/illustrations | 1 |
| Format | both print and electronic |
| Are you the author of this Elsevier article? | No |
| Will you be translating? | No |
| Title | Expanding the knowledge about the interactions and the secretion of salivary proteins |
| Institution name | Western University |
| Expected presentation date | Nov 2022 |
| Portions | Figure 1 |
| Requestor Location | Western University<br>[redacted]<br>[redacted]<br>Canada<br>Attn: Western University |
| Publisher Tax ID | GB 494 6272 12 |
| Total | 0.00 CAD |

# 7 Curriculum Vitae

**Name:** Karla Tonelli Bicalho Crosara

**Post-secondary Education and Degrees:**

The University of Western Ontario
London, Ontario, Canada
2015-2019 and 2020-2022, Ph.D.

The University of Sao Paulo (Universidade de Sao Paulo-PROFIS)
Bauru, Sao Paulo, Brazil
2001-2003, Specialist Diploma, Orthodontics and Dentofacial Orthopedics

University of Brasilia (Universidade de Brasilia)
Brasilia, Distrito Federal, Brazil
1995-2000, D.D.S.

**Scholarships and Awards:**

**NSERC Postgraduate Scholarship Doctoral (PGSD3),**
May 2018-Aug. 2019 and Sept. 2020-Apr. 2021
The University of Western Ontario, CAD$63,000

**Ontario Graduate Scholarship (OGS),**
May 2018-Apr. 2019
The University of Western Ontario, CAD$15,000 (Granted and Declined)

**Queen Elizabeth II Graduate Scholarship in Science and Technology,**
May 2016-Apr. 2017
The University of Western Ontario, CAD$15,000

**CIHR Master's Research Award,**
May 2016-Apr. 2017
The University of Western Ontario, CAD$15,000 (Waitlisted)

**Western Graduate Research Scholarship (WGRS),**
Sept. 2015-Aug. 2021
The University of Western Ontario, CAD$22,500

**3rd Prize** Senior Poster Presentation Award, CAD$50 (Co-author)
Oct. 2018
The University of Western Ontario, Schulich Dentistry Research Day

**1st Prize** Senior Poster Presentation Award, CAD$200 (Presenter)
Oct. 2017
The University of Western Ontario, Schulich Dentistry Research Day

**1st Prize** Senior Poster Presentation Aard, CAD$200 (Presenter)

Oct. 2016
The University of Western Ontario, Schulich Dentistry Research Day

**Top 100 Poster Award** – 2016th London Health Research Day
Mar. 2016
The University of Western Ontario

PIBIC/CNPq Scholarship for Scientific Initiation,
Mar. 1999 -Jul. 1999
University of Brasilia, Brazil, $10,000 (Brazilian Real)

**Related Work Experience**

**Adjunct Clinical Professor, Schulich Dentistry**
The University of Western Ontario
2021-present

Graduate Teaching Assistant (Volunteer)
The University of Western Ontario
Sept. 2016-Aug. 2018

1st Lieutenant Dentist, Orthodontist
Brazilian Army (Exercito Brasileiro), Brasilia, Brazil
Feb. 2004-Feb. 2011

Associate Dentist/Orthodontist
Bicalho Ortodontia, Brasilia, Brazil
Apr. 2000-Jul. 2006

Instructor and Course Coordinator
Associacao Brasileira de Odontologia, Brasilia, Brazil
Oct. 2003-Jul. 2007

Undergraduate Teaching Assistant
Universidade de Brasilia, Brasilia, Brazil
Aug. 1995-Aug. 1998

**Publications in Peer Reviewed Journals (12):**

Siqueira, W. L., Canales, M. P., **Crosara, K. T. B.**, Marin, L. M., & Xiao, Y. (2021). Proteome difference among the salivary proteins adsorbed onto metallic orthodontic brackets and hydroxyapatite discs. *Plos one*, *16*(7), e0254909.

Caixeta DC, Aguiar EMG, Cardoso-Sousa L, Coelho LMD, Oliveira SW, Espindola FS, Raniero L, **Crosara KTB**, Baker MJ, Siqueira WL, and Sabino-Silva R. (2020). Salivary molecular spectroscopy: A sustainable, rapid and non-invasive monitoring tool for diabetes mellitus during insulin treatment. *PLoS One. 17*;15(3): e0223461.

Bicalho, J.S., **Crosara K.T.B.**, Bicalho R.F., and Oliveira S.B. (2020). Carta ao Editor. *Revista Clínica de Ortodontia Dental Press* 19(3):8-11

Ferreira EO, Mendes INVF, Monteiro SG, **Crosara KTB**, Siqueira WL, de Maria Pedroso Silva de Azevedo C, Moffa EB, and de Andrade Monteiro C. (2019). Virulence properties and sensitivity profile of Candida parapsilosis complex species and Kodamaea ohmeri isolates from onychomycosis of HIV/AIDS patients. *Microb Pathog*.;132:282-292.

**Crosara, K.T.B**., Zuanazzi, D., Moffa, E.B., Xiao, Y., Machado, M.A.D.A.M. and Siqueira, W.L. (2018). Revealing the Amylase Interactome in Whole Saliva Using Proteomic Approaches. *BioMed Research International, 2018*.

**Crosara, K.T.B.,** Moffa, E.B., Xiao, Y. and Siqueira, W.L. (2018). Merging in-silico and in vitro salivary protein complex partners using the STRING database: A tutorial. *Journal of proteomics, 171*, pp.87-94.

Valente, M.T., Moffa, E.B., **Crosara, K.T.B.,** Xiao, Y., Oliveira, T.M., Machado, M.A.D.A.M. and Siqueira, W.L. (2018). Acquired Enamel Pellicle Engineered Peptides: Effects on Hydroxyapatite Crystal Growth. *Scientific reports*, *8*(1), p.3766.

Tikhonova, S., Booij, L., D'Souza, V., **Crosara, K.T.B.,** Siqueira, W.L. and Emami, E. (2018). Investigating the association between stress, saliva and dental caries: a scoping review. *BMC oral health*, *18*(1), p.41.

Aboodi, G.M., Sima, C., Moffa, E.B., **Crosara, K.T.B.**, Xiao, Y., Siqueira, W.L., Glogauer, M. (2015). Salivary cytoprotective proteins in inflammation and resolution during experimental gingivitis--a pilot study. *Front Cell Infect Microbiol*. 5:92.

Silva Filho, O. D., Capelozza Filho, L., **Crosara, K. T. B.**, & Ozawa, T. O. (2007). Avaliação cefalométrica dos efeitos do aparelho Herbst no tratamento da deficiência mandibular na dentadura permanente. *Revista Dental Press de Ortodontia e Ortopedia Facial*, *12*, 101-118.

Bicalho, J.S., **Bicalho, K.T.** (2002). Descrição do método de contenção fixa com livre acesso do fio dental – Dica Clinica (Description of the method of fixed retention with free access of the dental floss-Clinical Tip). *R Clin Ortodon Dental Press*. 1(1): 9-13.

Bicalho, J. S., & **Bicalho, K. T.** (2001). Descrição do método de contençäo fixa, com livre acesso do fio dental. *Rev. dent. press ortodon. ortop. maxilar*, 97-104.

**Participation in Conferences (23):**

**Crosara, KTB,** et al (2018) "Orthodontic bracket pellicle proteome modulation after exposure to fluoride" 48th American Association of Dental Research/42st Canadian Association of Dental Research (CADR) Meeting. Fort Lauderdale, FL, USA *International-Poster*

Neilson, NW, **Crosara, KTB**, Mamandras, A, Siqueira, WL (2018) "Acquired bracket pellicle modulation via exposure to histatin 3." 2018 Annual CAO (Canadian Association of Orthodontists) Conference. London, ON, Canada, *National-Poster*

Neilson, NW, **Crosara, KTB**, Mamandras, A, Siqueira, WL (2018) "Acquired bracket pellicle modulation via exposure to histatin 3." Schulich Dentistry Research Day. *Institutional-Poster*

**Crosara, KTB,** et al (2017) "Foreseeing Protein Evolution: Benefiting from Engineered Salivary Peptides" 11th European Symposium on Saliva. Egmond Ann Zee, The Netherlands. *International-Oral*

**Crosara, KTB,** et al (2017) "Amylase Interactome in Whole Saliva using Proteomic Approaches" 95th International Association for Dental Research (IADR) Meeting. San Francisco, CA, USA. *International-Oral*

**Crosara, KTB**, et al (2017) "Amylase Interactome in Whole Saliva using Proteomic Approaches" Schulich Dentistry Research Day. *Institutional-Poster*

**Crosara, KTB**, et al (2017) "Amylase Interactome in Whole Saliva using Proteomic Approaches." London Health Research Day. London, ON. *Local-Poster*

Canales, M.P., **Crosara, K.T.B.**, et al. (2017) "Can exposure to fluoride solutions change the surface of orthodontic brackets?" Schulich Dentistry Research Day. *Institutional-Poster*

Canales, M.P., **Crosara, K.T.B.**, et al. (2017) "Can exposure to fluoride solutions change the surface of orthodontic brackets?" 2017 Annual CAO (Canadian Association of Orthodontists) Conference. Toronto, ON. *International-Poster*

Heller, D., **Crosara, K.T.B.**, et al. (2017). "In Vitro Zirconia Salivary Pellicle Proteome." 34ª Reunião Anual 2017-SBPqO (Sociedade Brasileira de Pesquisa Odontologica Divisao Brasileira da IADR) Meeting, Campinas, Brazil. *International-Oral*

Nunes, P.L.S., Moffa, E.B., Siqueira, W.L., **Crosara, K.T.B.**, et al. (2017). "Caracterização de uma amostra de pacientes com Mucopolissacaridose do Estado do Maranhão." 34ª Reunião Anual 2017-SBPqO (Sociedade Brasileira de Pesquisa Odontologica Divisao Brasileira da IADR). Campinas, Brazil. *International-Oral*

Jorge, P.K., Cardoso, M., Santos, M.B., **Crosara, K.T.B.**, et al. (2017). "Análise proteômica de dentes decíduos traumatizados." 34ª Reunião Anual 2017-SBPqO (Sociedade Brasileira de Pesquisa Odontologica Divisao Brasileira da IADR). Campinas, Brazil. *International-Oral*

**Crosara, KTB**, et al (2016) "A creative, inexpensive and effective approach to remove amylase from oral fluids prior to mass spectrometry" Schulich Dentistry Research Day. *Institutional-Poster*

**Crosara, KTB**, et al (2016) "A creative, inexpensive and effective approach to remove amylase from oral fluids prior to mass spectrometry." London Health Research Day. *Local-Poster*

Kohara, E.K., Zuanazzi, D., **Crosara, K.T.B.**, et al. (2016). "Protein salivary contents of teenagers with early caries lesions." (2016) AADR/CADR Annual Meeting, Los Angeles, USA. *International-Poster*

**Crosara, K.T.B.** (2003). "Treatment of the mandibular retrognathism with the Herbst appliance." V GREO (Meeting of Residents and Specialist from Profis-Funcraf), Bauru, Brazil. *Institutional-Oral*

**Crosara, K.T.B.** (2002). "The use of NIKON COOLPIX camera in dentistry." 13th Brazilian Orthodontic Conference-Orto 2002-SPO, Sao Paulo, Brazil. *National-Oral*

Bicalho, J.S. and **Crosara, K.T.B.** (2002). "Oral Digital photography." Continuing Education Program – Brazilian Dentistry Association, Brasilia, Brazil. *Local-Oral*

Gama, F.G.V., **Bicalho, K.T.** (1999). "Evaluation of the effects of an intra-oral mandibular repositioner for the treatment of snoring and sleep apnea: analysis of the profile of the patients." V Dental Journey of the University of Brasilia, Brazil. *Institutional-Poster*

**Bicalho, K.T.**, Gama, F.G.V. (1999). "Evaluation of the effects of an intra-oral mandibular repositioner for the treatment of snoring and sleep apnea: analysis of the efficiency of the device." V Dental Journey of the University of Brasilia, Brazil. *Institutional-Poster*

**Bicalho, K.T.** (1999). "Evaluation of the effects of an intra-oral mandibular repositioner for the treatment of snoring and sleep apnea: analysis of the efficiency of the device." V Conference of Scientific Initiation of the University of Brasilia, Brasilia, Brazil *Institutional-Poster*

**Bicalho, K.T.**, Gama, F.G.V. (1999). "Making of a hybrid functional device for treatment of snoring and obstructive sleep apnea." VII International Conference of Dentistry of Minas Gerais, Belo Horizonte, Brazil. *Interational, Oral*

**Bicalho, K.T.**, Gama, F.G.V. (1998). "Making of a hybrid functional device for treatment of snoring and obstructive sleep apnea." IV Dental Journey of the University of Brasilia, Brasilia, Brazil. *Institutional-Poster*

**Scientific Editorial:**

**Reviewer,** BioMed Research International, Hindawi (2021-present)

**Reviewer**, Clinical Oral Investigations, Springer Nature (2018-2019)

**Reviewer**, BMC Oral Health, BioMed Central, Springer Nature (2017-2020)

**Languages:**

Fluent in English and Portuguese (spoken and written)

Limited proficiency in French and Spanish (spoken and written)