



Universiteit  
Leiden  
The Netherlands

## Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins

Kozak, R.P.

### Citation

Kozak, R. P. (2017, January 24). *Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins*. Retrieved from <https://hdl.handle.net/1887/45434>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/45434>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/45434> holds various files of this Leiden University dissertation.

**Author:** Kozak, R.P.

**Title:** Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins

**Issue Date:** 2017-01-24

# Chapter 6

## Variation of human salivary O-glycome

**Radoslaw P. Kozak<sup>1</sup>, Paulina A. Urbanowicz<sup>1</sup>, Chamindie Punyadeera<sup>2</sup>, Karli R. Reiding<sup>3</sup>, Bas C. Jansen<sup>3</sup> Louise Royle<sup>1</sup>, Daniel I. Spencer<sup>1</sup>, Daryl L. Fernandes<sup>1</sup> and Manfred Wuhrer<sup>3</sup>**

<sup>1</sup>Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, United Kingdom

<sup>2</sup>School of Biomedical Sciences, Institute of Health and Biomedical Innovations, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, Australia

<sup>3</sup>Centre for Proteomics and Metabolomics Leiden University Medical Centre, The Netherlands

PLOS ONE (2016) 1-15.

## **Abstract**

The study of saliva O-glycosylation is receiving increasing attention due to the potential of glycans for disease biomarkers, but also due to easy access and non-invasive collection of saliva as biological fluid. Saliva is rich in glycoproteins which are secreted from the bloodstream or produced by salivary glands. Mucins which are highly O-glycosylated proteins are particularly abundant in human saliva. Their glycosylation is associated with blood group and secretor status, and represents a reservoir of potential disease biomarkers. This study aims to analyse and compare O-glycans released from whole human mouth saliva collected 3 times a day from a healthy individual over a 5 days period. O-linked glycans were released by hydrazinolysis, labelled with procainamide and analysed by ultra-high performance liquid chromatography with fluorescence detection (UHPLC-FLR) coupled to electrospray ionization mass spectrometry (ESI-MS/MS). The sample preparation method showed excellent reproducibility and can therefore be used for biomarker discovery. Our data demonstrates that the O-glycosylation in human saliva changes significantly during the day. These changes may be related to changes in the salivary concentrations of specific proteins.

## Introduction

Saliva is a unique fluid that lubricates and buffers the oral cavity.<sup>1,2</sup> It protects the oral cavity from microbial attack involving both the innate and acquired immune surveillances. Saliva is a structurally very complex secretion containing components from a variety of glands such as the parotid, submaxillary, sublingual and submandibular glands and some minor glands in the lip, cheek, tongue and palate.<sup>3</sup> Saliva contains a wide spectrum of microbes, epithelial cells, proteins/peptides, electrolytes, hormones, nasal and bronchial secretions and serum product.<sup>4-7</sup>

Saliva contains many glycoproteins, some of which are highly glycosylated (such as MUC5B and MUC7) as well as salivary agglutinin, secretory immunoglobulin A (sIgA), lactoferrin, amylase and proline-rich glycoproteins.<sup>8</sup> Mucins are dominant compounds in human saliva and play an important role in the maintenance of oral health and protection of teeth as part of the first line oral defence.<sup>9,10</sup> Many glycoproteins bind to a variety of bacteria and either aid or prevent adherence of bacteria to mucosal and tooth surfaces or lubricate tooth surfaces and possibly protect proteins from proteolytic attack.<sup>11</sup>

Bacteria usually have a net negative surface charge and therefore tend not to bind directly to negatively charged mucins. Instead, they may interact through bridging reagents involving divalent cations such as  $\text{Ca}^{2+}$  ions. However, many bacteria express cell surface proteins (lectins) that specifically recognize oligosaccharides present on the salivary mucins. In this manner salivary mucins are bound by multiple bacteria resulting in agglutination. Removal of the terminal sialic acids from the oligosaccharide side chains of salivary mucins was found to abolish agglutination of oral streptococci, indicating that the bacterial lectin has specificity for sialic acids.<sup>12</sup> Salivary agglutinin, a highly glycosylated protein frequently associated

with other salivary proteins and with secretory immunoglobulin A (IgA), is one of the main salivary components responsible for bacteria agglutination.<sup>13</sup>

Secretory IgA is the most abundant humoral immunologic component of saliva. It can neutralize viruses, bacteria, and enzyme toxins. It binds bacterial antigens and is able to aggregate bacteria and inhibit their adherence to oral tissues.<sup>14,15</sup> Other immunologic components, such as IgG and IgM, occur in lower quantity and probably originate from gingival fluid.<sup>16</sup>

Salivary composition shows associations with various oral and systemic diseases<sup>17-22</sup>, and significant age or sex related differences have been reported with respect to genomic<sup>23,24</sup> and proteomic expression levels.<sup>25-27</sup>

Recently interest in glycomics (the study of the spectrum of glycans expressed in a biological system) has increased due to the potential clinical utility of glycans as disease biomarkers.<sup>28</sup> Glycosylation of proteins is a very complex and common post-translational modification influencing many biological functions. There are two main types of glycosylation found in eukaryotes: N-glycosylation and O-glycosylation. On salivary glycoproteins, N-glycans are structurally diverse encompassing the high mannose, hybrid and complex N-glycan subclasses. The O-glycans comprise various O-glycan core types, with optionally fucoses in different linkages as well as sialylation and sulphation.<sup>29,30</sup> It has been reported that salivary MUC5B contains O-glycans which vary between individuals and that some of this variation is related to blood group and secretor status.<sup>31</sup> Also O-glycans attached to salivary MUC7 showed changes in the patients with recurrent aphthous stomatitis (RAS). Patients with RAS have less complex O-glycans suggesting that there are bacterial oligosaccharide degrading enzymes reducing glycan complexity.<sup>32</sup>

Human saliva proteomics has proven to be a novel approach in the search for

protein biomarkers for detection of diseases <sup>31</sup>, including cancer, autoimmune diseases, viral diseases, bacterial diseases, cardiovascular diseases, and human immunodeficiency virus (HIV).<sup>2,33</sup> It has become clear that proteomic composition of saliva can reflect the pathological as well as physiological state.<sup>34,35</sup> Recently, a growing number of proof-of-principle assays have been established using saliva to monitor diseases or bodily conditions such as infection, immune responses to viral infections, systemic levels of drugs, and the detection of illicit drug use.<sup>36</sup> Saliva profiling over the course of disease progression could reveal potential biomarkers indicative of different stages of diseases, which may be useful in medical diagnostics. This follows from several studies showing that glycosylation patterns in other body fluids such as plasma correlate with progression of certain diseases including inflammatory conditions and subclasses of diabetes.<sup>37-39</sup>

Here we describe the evaluation of longitudinal changes in salivary O-glycosylation profiles using samples collected 3 times a day over a 5 days period. The O-glycans were released by hydrazinolysis using a recently developed improved protocol <sup>40</sup> and fluorescently labelled with procainamide, followed by HILIC-UHPLC-FLR-ESI-MS analysis. The chosen techniques proved to be suitable to demonstrate major changes in the salivary O-glycome during the day.

## **Materials and methods**

### **Materials**

All reagents and kits for O-glycan release, labelling and clean-up were from Ludger Ltd. (Oxford, UK). Centrifugal filter devices with a molecular weight cut-off (MWCO) membrane of 10 kDa were obtained from Fisher Scientific UK (Leicestershire, UK). Acetonitrile (Romil; 190 SpS for UV/ gradient quality) was obtained from Charlton

Scientific (Charlton, Oxon, UK). Fetuin glycoprotein was from Ludger Ltd (Oxford, UK) and all other reagents were obtained from Sigma-Aldrich (Dorset, UK). Procainamide labelling was performed on a Hamilton Microlab STARLet liquid-handling robotic platform.

## **Collection of saliva**

Whole unstimulated saliva (5 mL) was collected 3 times a day over a 5 days period from two human donors. Volunteers were asked to rinse their mouths with water for 30 seconds to remove any food particles. Donors were also asked to refrain from eating 2 h prior to collection<sup>3</sup>. PBS supplemented with protease inhibitor cocktail (2 mL, AEBSF at 2 mM, Aprotinin at 0.3  $\mu$ M, Bestatin at 116  $\mu$ M, E-64 at 14  $\mu$ M, Leupeptin at 1  $\mu$ M and EDTA at 1 mM) and antibiotic-antimycotic mixture containing 10,000 units penicillin, 10 mg streptomycin and 25  $\mu$ g amphotericin B per mL were added to each saliva sample. The saliva samples were clarified by centrifugation at 3000 x g at 4°C for 20 min to remove cellular debris. The supernatant was concentrated<sup>29</sup> using centrifugal filter device (10 kDa MWCO membrane). Samples were then stored at -20°C until required for further experimentation.

## **Release of O-glycans by hydrazinolysis**

The saliva samples and fetuin glycoprotein (used as a positive control) were buffer exchanged into 0.1% TFA prior to hydrazinolysis as described previously<sup>40</sup>. Samples were transferred to glass vials and dried down for 16 h by vacuum centrifugation prior to the addition of hydrazine.

Fetuin and saliva samples were incubated with 450  $\mu$ L anhydrous hydrazine (Ludger) at 60°C for 6 h.<sup>41,42</sup> Excess hydrazine was removed by centrifugal

evaporation. The samples were placed on ice for 20 min (0°C) and were re-*N*-acetylated by addition of 0.1 M sodium bicarbonate solution (200 µL) and acetic anhydride (21 µL). Samples were cleaned up by passing them through the Ludger Clean CEX cartridges (Ludger). The O-glycans were eluted off the cartridges using water (3 x 0.5 mL). Eluates were dried by vacuum centrifugation prior to fluorescent labelling.

### **Procainamide labelling**

Off the released O-glycans, 40% was fluorescently labelled with procainamide as described previously <sup>43</sup> using Ludger procainamide glycan labelling kit containing 2-picoline borane. Briefly, samples in 10 µL of water were mixed with 10 µL of procainamide labelling solution and incubated at 65°C for 1 h. The samples were evaporated to dryness by vacuum centrifugation and re-suspended in water (100 µL) for further analysis.

### **LC-ESI-MS and MS/MS analysis**

Procainamide labelled samples were analysed by HILIC-UHPLC-ESI-MS with fluorescence detection. Samples were injected in 10% water/90% acetonitrile (injection volume 25 µL) onto an ACQUITY UPLC® BEH-Glycan 1.7 µm, 2.1 x 150 mm column at 60°C on a Ultimate 3000 UHPLC instrument with a fluorescence detector ( $\lambda_{ex}$  = 310nm,  $\lambda_{em}$  = 370nm; Thermo Scientific). The running conditions used were: Solvent A was 50 mM ammonium formate pH 4.4 made from Ludger Stock Buffer (Ludger), and solvent B was acetonitrile. Gradient conditions were: 0 to 10 min, 90% B; 10 to 95 min, 90 to 57 % B at flow rate 0.4 mL/min; 95 to 98 min, 90 to 0 % B at a flow rate of 0.4 mL/min; 98 to 99 min, 0 % B at a flow rate of 0.25

mL/min; 99 to 100 min, 0% B at a flow rate of 0.25 mL/min. 100 to 103 min, 90% B at a flow rate 0.25 mL/min; 103 to 115 min, 90% B at a flow rate 0.4 mL/min. The UHPLC system was coupled on-line to an Amazon Speed ETD (Bruker Daltonics, Bremen, Germany) with the following settings: source temperature 250°C, gas flow 10 L/min; Capillary voltage 4500 V; ICC target 200,000; maximum accumulation time 50 ms; rolling average 2; number of precursors ions selected 3, release after 0.2 min; Positive ion mode; Scan mode: enhanced resolution; mass range scanned, 200-1600; Target mass, 900. A glucose homopolymer ladder (Ludger), labelled with procainamide was used as a system suitability standard as well as an external calibration standard for GU allocation.<sup>44</sup>

## Sample variation

Using in-house developed scripts, raw HILIC chromatograms for 2 individuals were exported from Chromeleon 7.1 (Dionex, USA) and aligned by applying the least variance linear fit through the highest data points observed within  $\pm 0.25$  min window around retention times 22.33, 27.13, 28.70, 29.28, 32.43 and 33.64 min. Then, for each chromatogram independently, the areas of the 7 most abundant peaks were integrated centering at 20.55 min (within a  $\pm 0.5$  min window), 22.26 ( $\pm 0.35$ ), 23.99 ( $\pm 0.35$ ), 24.74 ( $\pm 0.5$ ), 27.17 ( $\pm 0.35$ ), 32.39 ( $\pm 0.25$ ) and 33.57 ( $\pm 0.25$ ). The obtained peak areas were normalized to the total area sum within each chromatogram, and expressed in stripplot format using R version 3.1.2 (R Foundation for statistical computing, Vienna Austria).

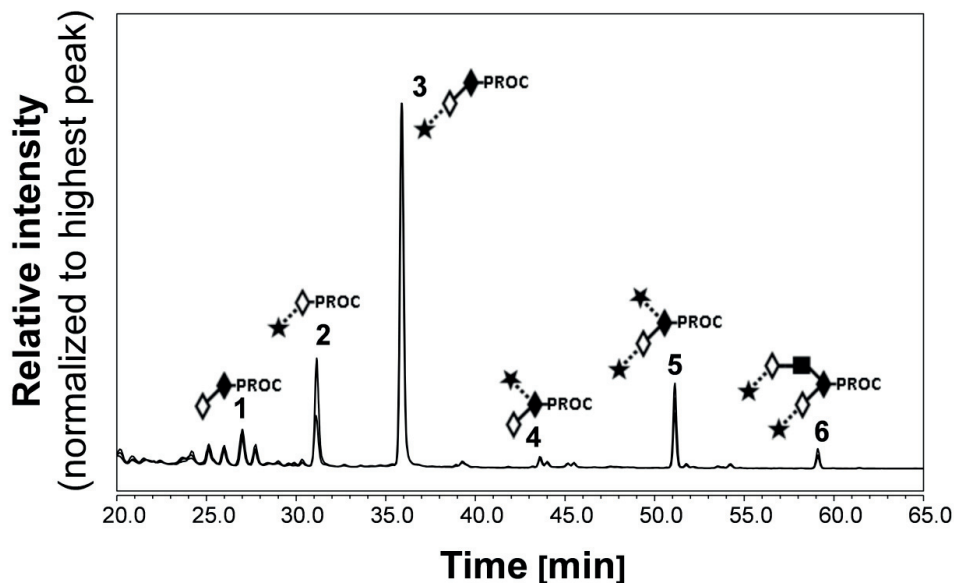
## **Ethics statement**

The Research Ethic Committee (NHS Health Research Authority, Surrey Borders REC London, UK) granted approval for this study with all participants giving written, informed consent (REC 16/LO/0883; IRAS project ID: 203515).

## **Results and discussion**

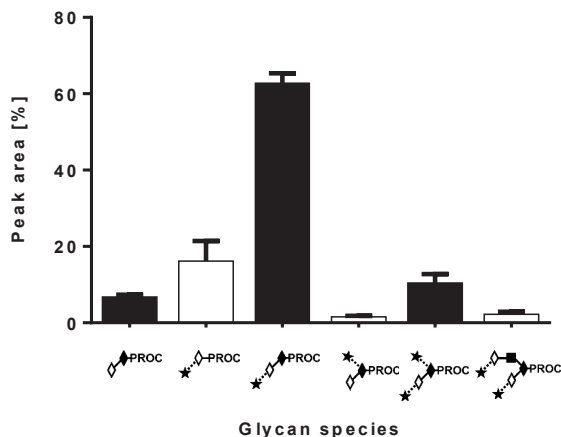
### **Evaluation of method repeatability**

The release and recovery of O-glycans still remains very challenging because there is no universal enzyme available for O-glycan release from glycoproteins. Therefore, the current best methods for universal O-glycan release rely on rather harsh chemistry. Over the last two decades a number of chemical methods for O-glycan removal have been developed, including reductive and non-reductive  $\beta$ -elimination and hydrazinolysis.<sup>45</sup> It has been reported that the most suitable method for universal O-glycan release in their non-reduced form in high yields is hydrazinolysis<sup>41</sup>. We decided to use this technique to release O-glycans from human saliva because the released O-glycans contain free reducing termini allowing further derivatization with fluorescent tags and analysis by HPLC-based methods.<sup>40,46</sup>



**Figure 1.** HILIC-UHPLC profiles for procainamide labelled O-glycans released in triplicate from fetuin..The following symbols are used to depict glycan structures:<sup>47</sup> open diamond = galactose; closed diamond = N-acetylgalactosamine; closed square = N-acetylglucosamine; closed star = N-acetyleneuraminic acid; dashed line,  $\alpha$ -linkage; solid line,  $\beta$ -linkage.

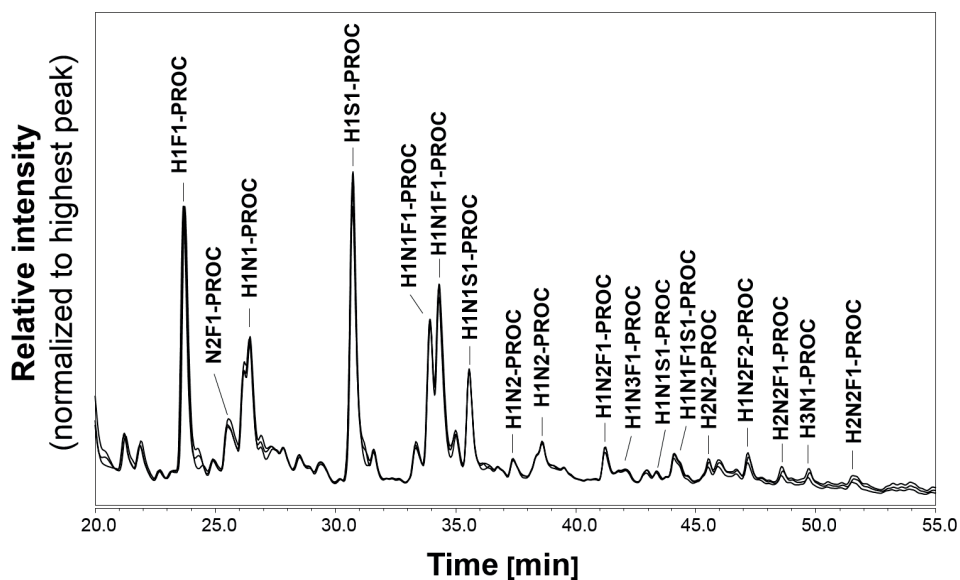
First, we investigated whether the sample preparation and glycan release method is suitable for saliva samples. Using fetuin glycoprotein and a saliva reference sample collected from a single individual as positive controls, we evaluated the O-glycan preparation and O-mode hydrazinolysis. Triplicates of bovine fetuin and saliva samples were buffer exchanged into 0.1% TFA prior to release.<sup>40</sup> The O-glycans were released using anhydrous hydrazine, labelled with procainamide and analysed by HILIC-UHPLC with fluorescence detection providing good glycan separation, coupled to ESI-MS and MS/MS providing glycan composition.



**Figure 2.** O-glycans patterns obtained from bovine fetuin. Samples were washed with 0.1% TFA followed by hydrazinolysis, procainamide labelling and HILIC-UHPLC analysis with fluorescence detection. The average abundance of the O-glycans was determined on multiple days. Symbols are as in Figure 1.

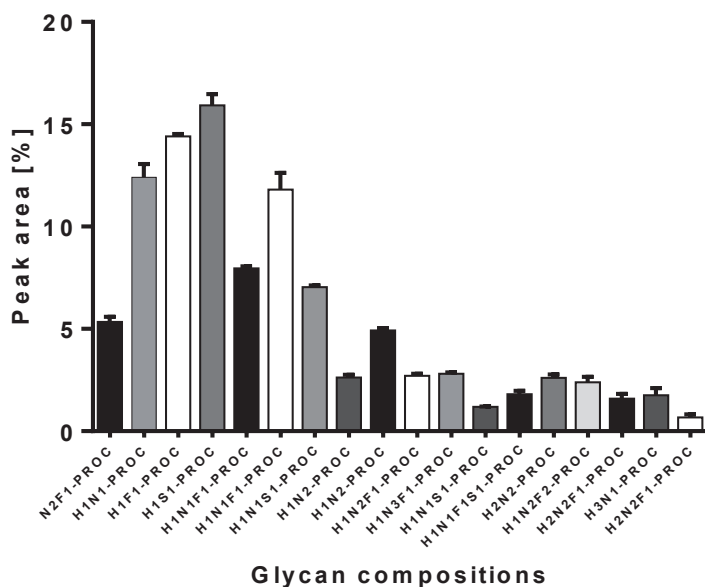
The HILIC-UHPLC profiles for triplicate releases of O-glycans from fetuin sample (Figure 1) were consistent with previously published data <sup>46,48</sup>. The fetuin profiles contain 6 main peaks corresponding to: core 1, Gal $\beta$ 1–3GalNAc (peak 1); monosialylated core 1 O-glycans, Neu5Ac $\alpha$ 2–3Gal $\beta$ 1–3GalNAc (peak 3); and Neu5Ac $\alpha$ 2–6(Gal $\beta$ 1–3)GalNAc (peak 4); disialylated core 1 O-glycan, Neu5Ac $\alpha$ 2–3Gal $\beta$ 1–3(Neu5Ac $\alpha$ 2–6)GalNAc (peak 5); and disialylated core 2 O-glycan, Neu5Ac $\alpha$ 2–3Gal $\beta$ 1–3(Neu5Ac $\alpha$ 2–3Gal $\beta$ 1–4GlcNAc $\beta$ 1–6)GalNAc (peak 6). The peeled product Neu5Ac $\alpha$ 2–3Gal (peak 2, Figure 1) was also detected. To confirm the identity of observed peaks (peaks 1-6, Figure 1) the procainamide labelled O-glycans were characterized by ESI-MS/MS. The results from HILIC-UHPLC coupled to ESI-MS/MS are summarised in Supplementary Table 1 and an example of fragmentation analysis of procainamide labelled O-glycans is shown in Supplementary Figure 1. The glycan peak areas for each release were very similar with the standard deviation between the samples below 5.5 % for all peaks (Figure

2).



**Figure 3.** HILIC-UHPLC O-glycan profiles of saliva reference sample collected from a single individual in the morning. Sample was buffer exchange into 0.1% TFA, and the O-glycans were released in triplicate by hydrazinolysis followed by procainamide labelling and HILIC-UHPLC analysis with fluorescence detection.

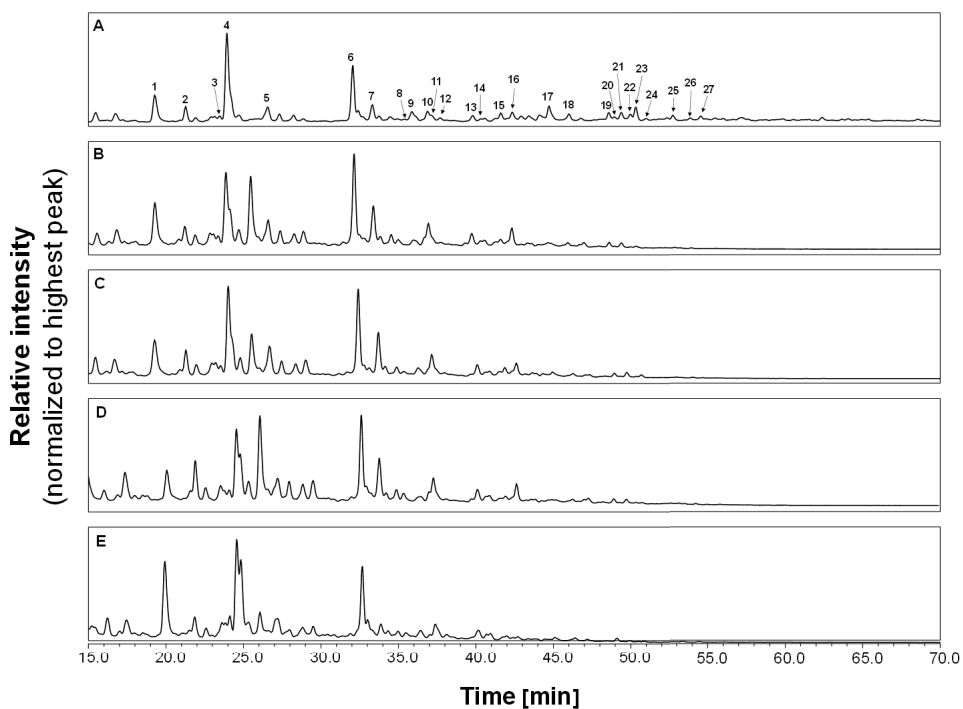
The robustness of the O-glycosylation analysis method was also tested on the saliva standard. The O-glycans were released by anhydrous hydrazine, labelled with procainamide and analysed by HILIC-UHPLC-ESI-MS with fluorescence detection. The HILIC-UHPLC profiles for triplicate releases of O-glycans from whole saliva sample are shown in Figure 3. As the profiles were very complex we selected a subset of 18 peaks for comparison. These peaks were analysed by ESI-MS/MS to determine O-glycan composition. The glycan peak areas for selected peaks were very similar with the standard deviation between the samples below 1 % (Figure 4). Thus, O-mode hydrazinolysis, procainamide labelling systems and UHPLC-ESI-MS/MS analysis are highly reproducible and suitable for the analysis of human saliva samples.



**Figure 4.** O-glycan patterns obtained from triplicate releases from saliva sample collected in the morning. Samples were washed with 0.1% TFA followed by hydrazinolysis, procainamide labelling and HILIC-UHPLC analysis with fluorescence detection. 18 O-glycan structures were selected for comparison.

## Longitudinal analysis of O-glycans from human saliva

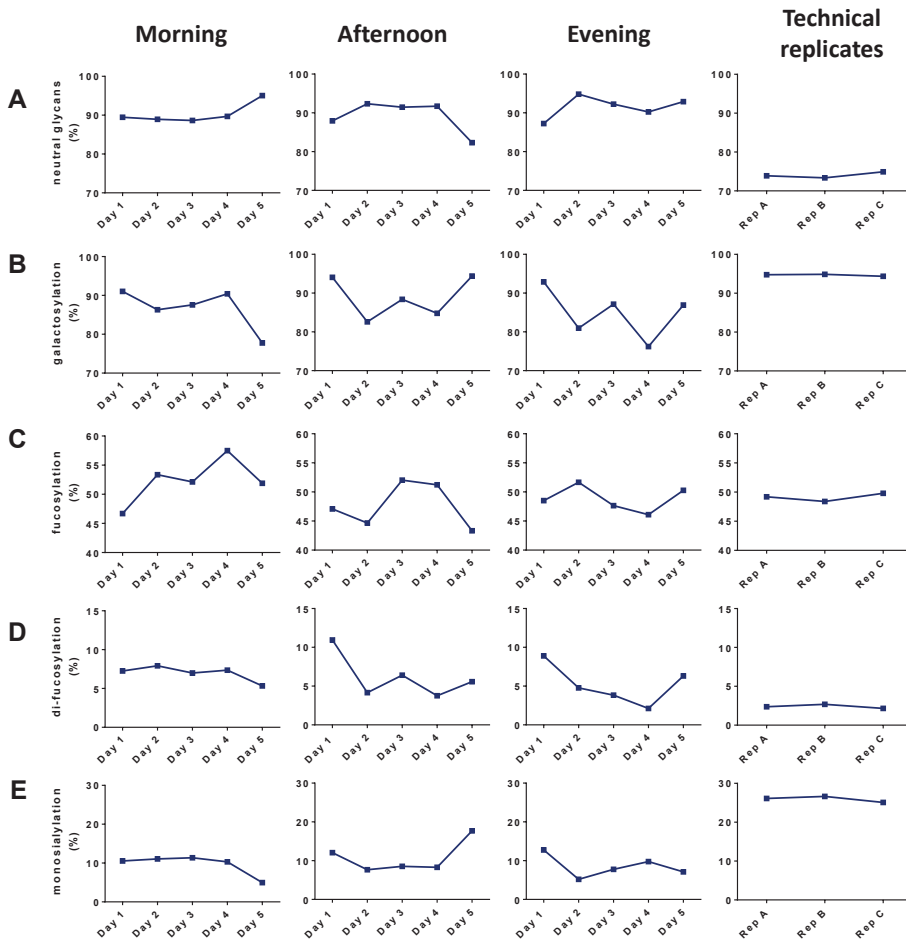
O-glycans were released from saliva collected 3 times a day over a 5 days period from 2 individuals by hydrazinolysis after acidic wash, labelled with procainamide and analysed by UHPLC-ESI-MS/MS. Only one sample was taken through ESI-MS/MS analysis. As the obtained HILIC profiles were very complex we selected a subset of 27 peaks for comparison and detailed analysis by ESI-MS/MS (Supplementary Figure 2). A broad range of O-glycan compositions were detected in the samples and although there was some overlap in LC traces, the sample profiles showed prominent differences. The observed molecular masses and compositions of 27 selected structures (Figure 5) are summarised in Table 1.



**Figure 5.** HILIC-UHPLC profile of saliva O-glycan patterns. Samples were collected in days 1-5 in the morning. Samples were buffer exchanged into 0.1% TFA, and the O-glycans were released by hydrazinolysis followed by procainamide labelling and HILIC-UHPLC analysis with fluorescence detection.

The O-glycans displayed a heterogeneous mixture of neutral and acidic oligosaccharides and most of the structures were fucosylated. The MS/MS fragmentation data of procainamide labelled O-glycans contained diagnostic fragment ions characteristic of core type or blood group epitope consistent with previously reported data.<sup>29,31,49,50</sup> The O-glycans carried the blood group H-epitopes Fuc $\alpha$ 1-2Gal $\beta$ 1 (diagnostic structure: Fuc-Hex) which is usually modified by a single monosaccharide residue (Gal $\alpha$ 1-3- or GalNAc $\alpha$ 1-3-) and determines ABO blood groups.<sup>31,51,52</sup> The O-glycans found in this study were released from whole saliva from an individual of blood group A. The terminal structures of glycans 13 and 18 (Figure 5, Table 1) have a HexNAc attached to (Fuc)-Hex-HexNAc-Proc (blood

group H-epitope), suggesting expression of blood group A antigen GalNAc $\alpha$ 1-3(Fuc $\alpha$ 1-2)Gal $\beta$ 1- on saliva glycoproteins. The presence of ABO blood group specific epitopes matching the blood group status suggest that the saliva donor is secretor.



**Figure 6.** Saliva O-glycosylation variation during collection. Several differences in saliva O-glycosylation between different collection times and days can be observed with respect to (A) neutral glycans, (B) galactosylation, (C) fucosylation, (D) difucosylation, (E) monosialylation. Data for the technical replicates corresponds to the saliva reference sample collected from a single individual and analysed during evaluation of method repeatability and robustness.

| UHPLC Peak Id | Average GU (PROC) | HILIC-UHPLC-ESI-MS |            |            |            |                               |                               |                     |
|---------------|-------------------|--------------------|------------|------------|------------|-------------------------------|-------------------------------|---------------------|
|               |                   | Composition        |            |            |            | Calculated [m/z] <sup>+</sup> | Registered [m/z] <sup>+</sup> | Probable structures |
|               |                   | Hex (H)            | HexNAc (N) | Fucose (F) | Neu5Ac (S) |                               |                               |                     |
| 1             | 1.40              | 0                  | 1          | 1          | 0          | 587.38                        | 587.32                        | -                   |
| 2             | 1.51              | 1                  | 0          | 1          | 0          | 546.38                        | 546.30                        | F-H-PROC            |
| 3             | 1.63              | 0                  | 2          | 1          | 0          | 791.45                        | 790.41                        | F-N-N-PROC          |
| 4             | 1.66              | 1                  | 1          | 0          | 0          | 603.39                        | 603.32                        | H-N-PROC            |
| 5             | 1.81              | 1                  | 1          | 1          | 0          | 749.45                        | 749.38                        | F-H-N-PROC          |
| 6             | 2.21              | 1                  | 1          | 1          | 0          | 749.45                        | 749.38                        | F-H-N-PROC          |
| 7             | 2.30              | 1                  | 1          | 0          | 1          | 894.48                        | 894.42                        | S-H-N-PROC          |
| 8             | 2.45              | 1                  | 2          | 0          | 0          | 806.46                        | 806.40                        | N-(H)-N-PROC        |
| 9             | 2.48              | 1                  | 2          | 0          | 0          | 806.46                        | 806.40                        | N-(H)-N-PROC        |
| 10            | 2.57              | 1                  | 1          | 2          | 0          | 895.5                         | 895.44                        | S-(S-H)-N-ROC       |
| 11            | 2.61              | 1                  | 2          | 1          | 0          | 952.51                        | 952.46                        | N-(S-H)-N-ROC       |
| 12            | 2.64              | 1                  | 3          | 0          | 0          | 1010.53                       | 1009.48                       | H-N-(N)-N-ROC       |
| 13            | 2.82              | 1                  | 2          | 1          | 0          | 952.51                        | 952.46                        | N-(F)-H-N-ROC       |
| 14            | 2.87              | 2                  | 2          | 0          | 0          | 969.5                         | 968.46                        | -                   |
| 15            | 2.98              | 1                  | 1          | 0          | 1          | 894.48                        | 894.42                        | S-(H)-N-PROC        |
| 16            | 3.05              | 1                  | 3          | 1          | 0          | 1156.59                       | 1155.54                       | S-H-N-(N)-N-PROC    |
| 17            | 3.28              | 2                  | 2          | 0          | 0          | 968.52                        | 968.46                        | H-(H-N)-N-PROC      |

|    |      |   |   |   |   |                        |                        |                    |
|----|------|---|---|---|---|------------------------|------------------------|--------------------|
| 18 | 3.41 | 1 | 2 | 2 | 0 | 1098.59                | 1098.52                | N-(F)-H-(F)-N-PROC |
| 19 | 3.69 | 3 | 1 | 0 | 0 | 927.1                  | 927.43                 | -                  |
| 20 | 3.73 | 2 | 2 | 1 | 0 | 1114.59                | 1114.51                | H-N-(F-H)-N-PROC   |
| 21 | 3.78 | 2 | 2 | 1 | 0 | 1114.59                | 1114.51                | H-N-(F-H)-N-PROC   |
| 22 | 3.85 | 2 | 2 | 0 | 1 | 1259.64                | 1259.55                | S-H-N-H-N-PROC     |
| 23 | 3.89 | 2 | 2 | 1 | 0 | 1114.59                | 1114.51                | H-N-(F-H)-N-PROC   |
| 24 | 3.97 | 2 | 2 | 1 | 0 | 1114.58                | 1114.51                | N-H-N-(F)-H-PROC   |
| 25 | 4.18 | 3 | 2 | 0 | 0 | 1130.58                | 1130.51                | -                  |
| 26 | 4.31 | 2 | 2 | 2 | 0 | 1260.65                | 1260.57                | F-H-N-(F-H)-N-PROC |
| 27 | 4.40 | 2 | 2 | 1 | 1 | [703.00] <sup>2+</sup> | [703.31] <sup>2+</sup> | F-H-N-(S-H)-N-PROC |

**Table 1.** Mass spectrometric data for selected 27 O-glycans obtained from human saliva.

Derived glycosylation features such as galactosylation and fucosylation were calculated for the selected O-glycan peaks to analyse the range of variability. Firstly we compared levels of galactosylation (Figure 6B). A pronounced decrease in galactosylation (from 91% on the first day to 78% on the fifth day) was observed for the samples collected in the morning. For the samples collected in the afternoon or evening variation was also high (between 83% and 94% for afternoon collection and 76% and 93% for evening collection).

Next we looked at changes in fucosylation for the samples collected three times a day during a 5 days period (Figure 6C). High biological variation was observed both intra-day and inter-day with fucosylation levels ranging from 43% to 57%.

We also looked at changes in sialylation. Significant variation (between 5% and 13%) between samples collected at different points of the day during the 5 days period was observed (Figure 6E).

The peak alignment across all HILIC chromatograms from two individuals shows that the individuals differ in relative peak abundance, while displaying similar variation within each peak (Supplementary Figure 3C.).

The main aim of this study was to analyse and compare entire O-glycan populations released from whole human saliva samples collected 3 times a day over a 5 days period. Our results highlighted that there is extensive natural variation in the salivary O-glycosylation. Both the intra-day and the inter-day variation appeared to be high.

We used whole human saliva because the wide spectrum of compounds (including glycoproteins) present in saliva may provide information for potentially important biomarkers and it has been reported that 30% of blood proteins are also present in saliva.<sup>2</sup> Also changes in O-glycans have been associated with different disease states.<sup>53,54</sup> It has been previously published that the variation in salivary MUC5B glycosylation is high and dependent on blood group or secretor status.<sup>31</sup> Likewise, one may expect that salivary O-glycosylation, along with dynamic changes in composition of other salivary analytes may depend on diet, lifestyle and inflammatory responses in the mucus layer, albeit additional studies will be necessary to study this. One of the biggest advantages of saliva is easy collection, storage and transportation. Donation is non-invasive, stress-free and it does not require highly-trained personnel.<sup>6</sup>

## Conclusions

Our study showed that the composition and relative abundance of the O-glycans in human saliva significantly change during the day. This is an interesting observation, that should be taken into consideration when evaluating salivary O-glycosylation e.g. for its diagnostic biomarker potential.

The changes in the specific O-glycan types within one day as well as between days may be caused by changes in the salivary concentrations of specific proteins, but may likewise be attributed to changes in protein-specific glycosylation profiles. It may, therefore make sense to follow the current study up and study the glycosylation of one of the mucins abundantly present in saliva with respect to its longitudinal robustness and biomarker potential. Mucins found in saliva have high turnover rates and respond quickly to altered conditions within the body.<sup>32,53</sup> Mucins and their O-glycans affect a range of aspects of oral health and their changes in the response to the infection and inflammation.<sup>32</sup>

## Acknowledgements

This work was supported by the European Union's Seventh Framework Program (FP7-Health-F5-2011) under grant agreement n°278535 (HighGlycan).

Ludger Ltd is a commercial bioscience company specialising in development and validation of glycoprofiling technology. The following authors declare competing financial interest(s): RPK, PAU, LR, DIS, DLF are employed by Ludger Ltd. In addition, products from Ludger were used in this research. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

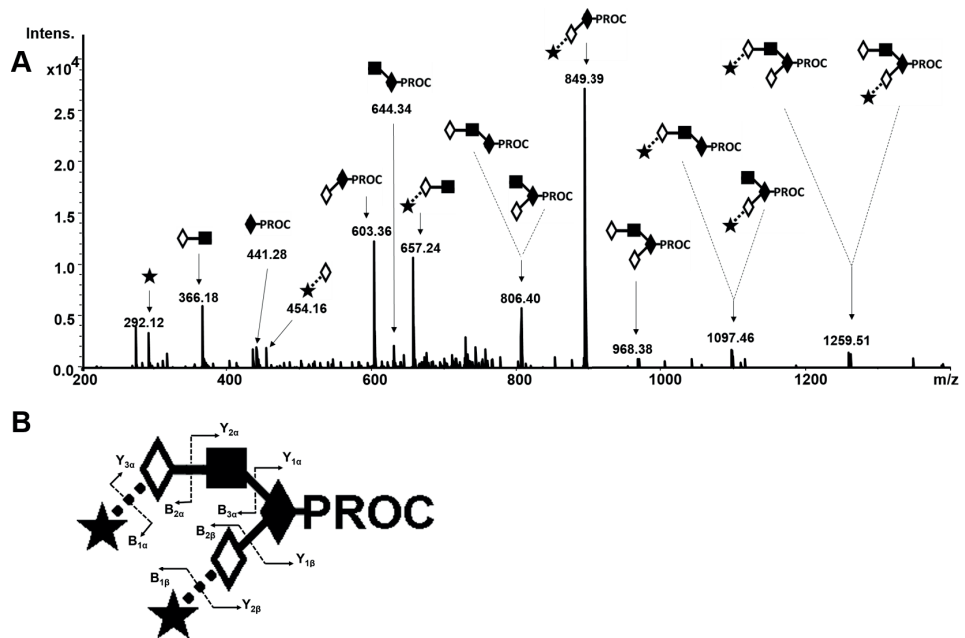
## Reference List:

- (1) Punyadeera, C.; Dimeski, G.; Kostner, K.; Beyerlein, P.; Cooper-White, J. J. *Immunol. Methods* **2011**, *373*, 19-25.
- (2) Schulz, B. L.; Cooper-White, J.; Punyadeera, C. K. *Crit Rev. Biotechnol* **2013**, *33*, 246-259.
- (3) Mohamed, R.; Campbell, J. L.; Cooper-White, J.; Dimeski, G.; Punyadeera, C. *Clin. Transl. Med* **2012**, *1*, 19.
- (4) Hofman, L. F. *J. Nutr* **2001**, *131*, 1621S-1625S.
- (5) Kaufman, E.; Lamster, I. B. *Crit Rev. Oral Biol. Med* **2002**, *13*, 197-212.
- (6) Pfaffe, T.; Cooper-White, J.; Beyerlein, P.; Kostner, K.; Punyadeera, C. *Clin. Chem* **2011**, *57*, 675-687.
- (7) Sondej, M.; Denny, P. A.; Xie, Y.; Ramachandran, P.; Si, Y.; Takashima, J.; Shi, W.; Wong, D. T.; Loo, J. A.; Denny, P. C. *Clin. Proteomics* **2009**, *5*, 52-68.
- (8) FML, A.; RMP, V.; MJC, I.; PMDN, D., 2007, pp 347-376.
- (9) Frohm, N. M.; Sandstedt, B.; Sorensen, O.; Weber, G.; Borregaard, N.; Stahle-Backdahl, M. *Infect. Immun* **1999**, *67*, 2561-2566.
- (10) Helmerhorst, E. J.; Oppenheim, F. G. *J. Dent. Res* **2007**, *86*, 680-693.
- (11) Amerongen, A. V.; Bolscher, J. G.; Veerman, E. C. *Glycobiology* **1995**, *5*, 733-740.
- (12) R.J., L.; H.J., J. *Oral Microbiology at a Glance*; Wiley-Blackwell, 2010.
- (13) Amerongen, A. V.; Veerman, E. C. *Oral Dis* **2002**, *8*, 12-22.
- (14) Schenkels, L. C.; Veerman, E. C.; Nieuw Amerongen, A. V. *Crit Rev. Oral Biol. Med* **1995**, *6*, 161-175.
- (15) AR, T. C. *Oral Histology: Development, Structure And Function*, 1998.
- (16) Edgar, W. M. *Br. Dent. J* **1992**, *172*, 305-312.
- (17) Borzym-Kluczyk, M.; Radziejewska, I.; Cechowska-Pasko, M. *Clin. Chem. Lab Med* **2011**, *49*, 885-889.
- (18) Dowling, P.; Wormald, R.; Meleady, P.; Henry, M.; Curran, A.; Clynes, M. *J. Proteomics* **2008**, *71*, 168-175.
- (19) Foo, J. Y.; Wan, Y.; Kostner, K.; Arivalagan, A.; Atherton, J.; Cooper-White, J.; Dimeski, G.; Punyadeera, C. *PLoS. One* **2012**, *7*, e48452.
- (20) Hu, S.; Arellano, M.; Boontheung, P.; Wang, J.; Zhou, H.; Jiang, J.; Elashoff, D.; Wei, R.; Loo, J. A.; Wong, D. T. *Clin. Cancer Res* **2008**, *14*, 6246-6252.

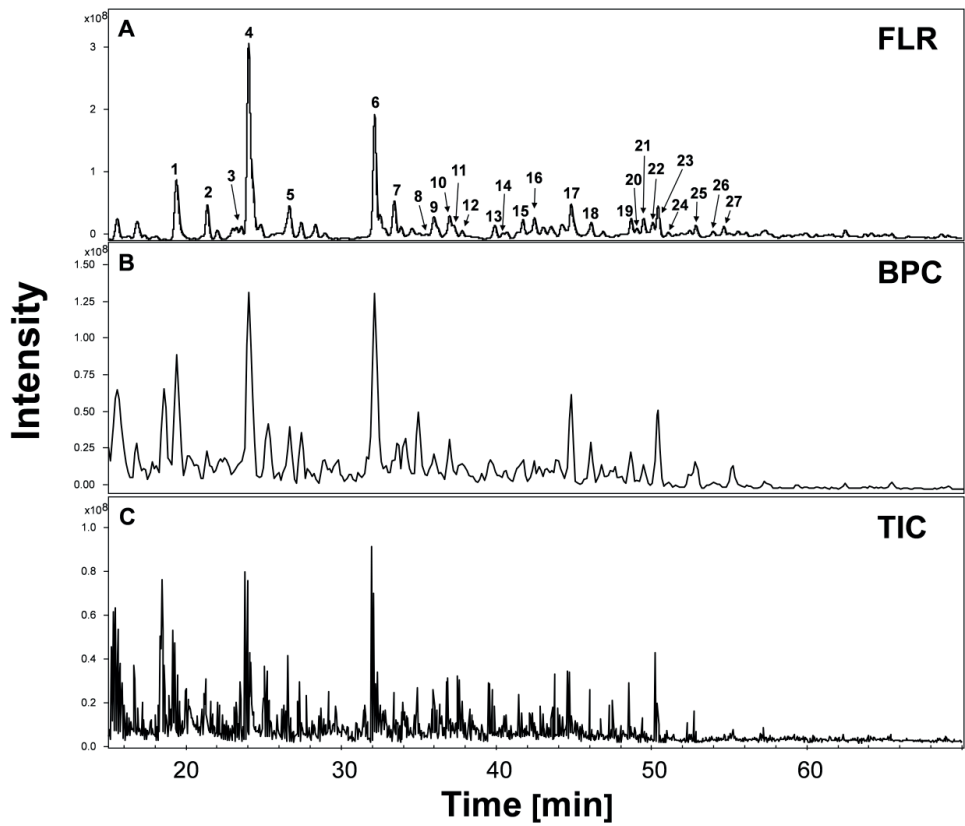
- (21) Ovchinnikov, D. A.; Wan, Y.; Coman, W. B.; Pandit, P.; Cooper-White, J. J.; Herman, J. G.; Punyadeera, C. *Biomark. Insights* **2014**, *9*, 53-60.
- (22) Zhang, L.; Xiao, H.; Karlan, S.; Zhou, H.; Gross, J.; Elashoff, D.; Akin, D.; Yan, X.; Chia, D.; Karlan, B.; Wong, D. T. *PLoS. One* **2010**, *5*, e15573.
- (23) Bocklandt, S.; Lin, W.; Sehl, M. E.; Sanchez, F. J.; Sinsheimer, J. S.; Horvath, S.; Vilain, E. *PLoS. One* **2011**, *6*, e14821.
- (24) Srivastava, A.; Wang, J.; Zhou, H.; Melvin, J. E.; Wong, D. T. *Arch. Oral Biol* **2008**, *53*, 1058-1070.
- (25) Ambatipudi, K. S.; Lu, B.; Hagen, F. K.; Melvin, J. E.; Yates, J. R. *J. Proteome. Res* **2009**, *8*, 5093-5102.
- (26) Cabras, T.; Pisano, E.; Boi, R.; Olianias, A.; Manconi, B.; Inzitari, R.; Fanali, C.; Giardina, B.; Castagnola, M.; Messina, I. *J. Proteome. Res* **2009**, *8*, 4126-4134.
- (27) Fleissig, Y.; Reichenberg, E.; Redlich, M.; Zaks, B.; Deutsch, O.; Aframian, D. J.; Palmon, A. *Oral Dis* **2010**, *16*, 831-838.
- (28) Gornik, O.; Pavic, T.; Lauc, G. *Biochim. Biophys. Acta* **2012**, *1820*, 1318-1326.
- (29) Everest-Dass, A. V.; Jin, D.; Thaysen-Andersen, M.; Nevalainen, H.; Kolarich, D.; Packer, N. H. *Glycobiology* **2012**, *22*, 1465-1479.
- (30) Everest-Dass, A. V.; Abrahams, J. L.; Kolarich, D.; Packer, N. H.; Campbell, M. P. *J. Am. Soc. Mass Spectrom* **2013**, *24*, 895-906.
- (31) Thomsson, K. A.; Schulz, B. L.; Packer, N. H.; Karlsson, N. G. *Glycobiology* **2005**, *15*, 791-804.
- (32) Zad, M.; Flowers, S. A.; Bankvall, M.; Jontell, M.; Karlsson, N. G. *Clin. Oral Investig* **2015**, *19*, 2147-2152.
- (33) Hu, S.; Jiang, J.; Wong, D. T. *Methods Mol. Biol* **2010**, *666*, 31-41.
- (34) Epstein, J. B.; Gorsky, M.; Guglietta, A.; Le, N.; Sonis, S. T. *Cancer* **2000**, *89*, 2258-2265.
- (35) Zauber, H.; Mosler, S.; von, H. A.; Schulze, W. X. *Proteomics* **2012**, *12*, 2221-2235.
- (36) Zhang, A.; Sun, H.; Wang, X. *Appl. Biochem. Biotechnol* **2012**, *168*, 1718-1727.
- (37) Clerc, F.; Reiding, K. R.; Jansen, B. C.; Kammeijer, G. S.; Bondt, A.; Wuhrer, M. *Glycoconj. J* **2015**.
- (38) Adamczyk, B.; Tharmalingam, T.; Rudd, P. M. *Biochim. Biophys. Acta* **2012**, *1820*, 1347-1353.

- (39) Jaeken, J.; Matthijs, G. *Annu. Rev. Genomics Hum. Genet* **2007**, *8*, 261-278.
- (40) Kozak, R. P.; Royle, L.; Gardner, R. A.; Fernandes, D. L.; Wuhrer, M. *Anal. Biochem* **2012**, *423*, 119-128.
- (41) Merry, A. H.; Neville, D. C.; Royle, L.; Matthews, B.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. *Anal. Biochem* **2002**, *304*, 91-99.
- (42) Patel, T.; Bruce, J.; Merry, A.; Bigge, C.; Wormald, M.; Jaques, A.; Parekh, R. *Biochemistry* **1993**, *32*, 679-693.
- (43) Kozak, R. P.; Tortosa, C. B.; Fernandes, D. L.; Spencer, D. I. *Anal. Biochem* **2015**, *486*, 38-40.
- (44) Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. *Methods Mol. Biol* **2006**, *347*, 125-143.
- (45) Zauner, G.; Kozak, R. P.; Gardner, R. A.; Fernandes, D. L.; Deelder, A. M.; Wuhrer, M. *Biol. Chem* **2012**, *393*, 687-708.
- (46) Kozak, R. P.; Royle, L.; Gardner, R. A.; Bondt, A.; Fernandes, D. L.; Wuhrer, M. *Anal. Biochem* **2014**, *453*, 29-37.
- (47) Harvey, D. J.; Merry, A. H.; Royle, L.; Campbell, M. P.; Dwek, R. A.; Rudd, P. M. *Proteomics* **2009**, *9*, 3796-3801.
- (48) Spiro, R. G.; Bhoyroo, V. D. *J. Biol. Chem* **1974**, *249*, 5704-5717.
- (49) Chaudhury, N. M.; Proctor, G. B.; Karlsson, N. G.; Carpenter, G. H.; Flowers, S. A. *Mol. Cell Proteomics* **2015**.
- (50) Karlsson, N. G.; Thomsson, K. A. *Glycobiology* **2009**, *19*, 288-300.
- (51) Karlsson, N. G.; Schulz, B. L.; Packer, N. H. *J. Am. Soc. Mass Spectrom* **2004**, *15*, 659-672.
- (52) Schulz, B. L.; Packer, N. H.; Karlsson, N. G. *Anal. Chem* **2002**, *74*, 6088-6097.
- (53) Kawashima, H. *Biol. Pharm. Bull* **2012**, *35*, 1637-1641.
- (54) Schulz, B. L.; Sloane, A. J.; Robinson, L. J.; Prasad, S. S.; Lindner, R. A.; Robinson, M.; Bye, P. T.; Nielson, D. W.; Harry, J. L.; Packer, N. H.; Karlsson, N. G. *Glycobiology* **2007**, *17*, 698-712.
- (55) Domon, B.; Costello, C. E. *Glycoconjugate journal* **1988**, *5*, 397-409.

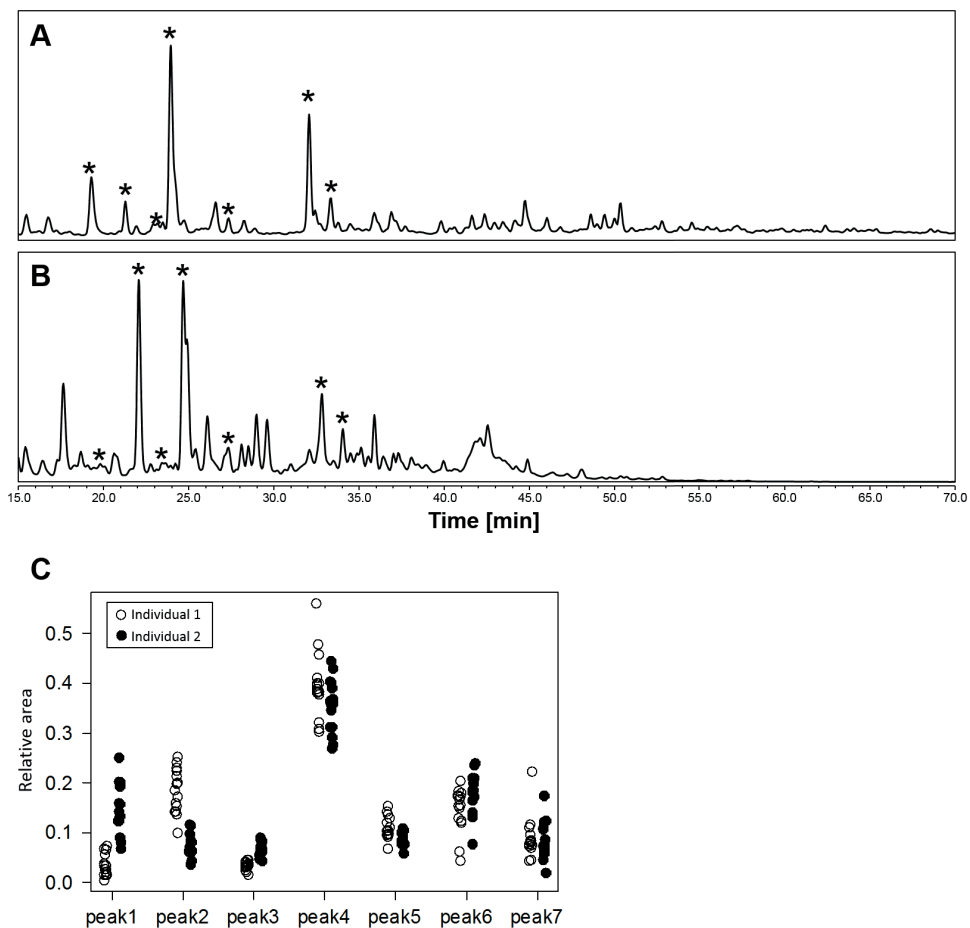
## Supplementary information



**Supplementary Figure 1.** MS/MS characterization of procainamide labelled fetuin O-glycan following online HILIC-UHPLC. (A) MS/MS characterization of the glycan eluting at retention time equivalent to that of GU 4.91. (B) Fragmentation scheme of the  $[M+2H]^{2+}$  ion at  $m/z$  775.84. The nomenclature used for fragmentation is based on literature.<sup>55</sup>



**Supplementary Figure 2.** Fluorescence, MS and MS/MS traces of saliva O-glycan patterns. Sample was collected in days 1 in the morning. (A) HILIC-FLR trace; (B) Base Peak Chromatogram (BPC); (C) Total Ion Chromatogram (TIC).



**Supplementary Figure 3.** Comparison of HILIC-UHPLC O-glycan profiles of saliva samples collected from two individuals. The 7 peaks labelled with asterisks correspond to signals suitable for alignment and integration. (A) Saliva sample from individual 1 collected at day 1 in the morning; (B) Saliva sample from individual 2 collected at day 1 in the morning; (C) Comparison between individual 1 and 2 of the variation across time within 7 O-glycans measured by HILIC-UHPLC.

After chromatogram alignment, the signals were integrated and the obtained areas normalized to the total area sum. The individuals show to differ in relative peak abundance, while displaying similar variation within each peak.

| UHPLC Peak Id | Average   |        | Structure | HILIC-UHPLC-ESI-MS and MS/MS |            |            |                               |                               |                               |                               |   |
|---------------|-----------|--------|-----------|------------------------------|------------|------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|---|
|               | GU (PROC) | % Area |           | Composition                  |            |            | [m/z] <sup>+</sup> calculated | [m/z] <sup>+</sup> calculated | [m/z] <sup>+</sup> registered | [m/z] <sup>+</sup> registered | [m/z] characteristic B and Y fragment ions (composition)  |
|               |           |        |           | Hex (H)                      | HexNAc (N) | Neu5Ac (S) |                               |                               |                               |                               |   |
| 1             | 1.93      | 6.80   |           | 1                            | 1          | 0          | 603.32                        | 302.17                        | 603.35                        | nd                            | 441.30 (N1-PROC)  |
| 2             | 2.18      | 16.15  |           | 1                            | 0          | 1          | 691.34                        | 346.17                        | 691.36                        | nd                            | 292.12 (S); 400.27 (H1-PROC)  |
| 3             | 2.50      | 62.82  |           | 1                            | 1          | 1          | 894.42                        | 447.71                        | 894.41                        | nd                            | 441.31 (N1-PROC); 603.34 (H1N1-PROC)  |
| 4             | 3.13      | 1.59   |           | 1                            | 1          | 1          | 894.42                        | 447.71                        | 894.41                        | nd                            | 441.32 (N1-PROC); 603.37 (H1N1-PROC)  |
| 5             | 3.90      | 10.42  |           | 1                            | 1          | 2          | 1185.51                       | 593.26                        | 1185.54                       | 593.30                        | 292.13 (S); 441.26 (N1-PROC); 454.16 (HS); 603.33 (H1N1-PROC); 732.36 (N1S1-PROC); 894.38 (H1N1S1-PROC)   |
| 6             | 4.91      | 2.22   |           | 2                            | 2          | 2          | 1550.65                       | 775.83                        | nd                            | 775.84                        | 292.12 (S); 366.18 (H1N1); 441.28 (N1-PROC); 454.16 (H1S1); 603.36 (H1N1-PROC); 644.34 (N2-PROC); 657.24 (H1N1S1); 806.40 (H1N2-PROC); 894.39 (H1N1S1-PROC); 968.38 (H2N2-PROC); 1097.46 (H1N2S1-PROC); 1259.51 (H2N2S1-PROC) |

**Supplementary Table 1.** Summary of O-glycan structures from fetuin with GU values, average relative abundance and mass values.