

Optimized NMR-based metabolomic protocol for salivary analysis in clinical practice

Manon Campas,^{1,2} Constance Germay,^{1,2,4} Arianna Cirillo,^{1,2,5} Matthieu Schoumacher,^{1,2,6}*

Arnaud Pinsart,^{1,2} Inès Defauwes,^{1,2,3} Etienne Cavalier,^{1,2} and Pascal de Tullio^{1,2}

¹Clinical Metabolomics Group (CliMe), Centre for Interdisciplinary Research on Medicines (CIRM), University of Liege, 4000 Liege, Belgium

²Department of Clinical Chemistry, University of Liege, CIRM, CHU of Liege, 4000 Liege, Belgium

³Laboratory of Translational Research in Nephrology (LTRN), Metabolism and Cardiovascular Sciences, Interdisciplinary Group of Applied Genoproteomics (GIGA), University of Liege, 4000 Liege, Belgium

^{4,5,6}For present addresses, see author information caption

Corresponding author: Manon Campas, Avenue Hippocrate 15, Quartier Hôpital, B36 Tour 4 +5, 4000 Liege Belgium, +32 4 366 43 81, mcampas@uliege.be

KEYWORDS: Saliva, sample preparation, clinical metabolomics, nuclear magnetic resonance (NMR), intra and inter-day precision

ABSTRACT Saliva remains underexplored in metabolomics compared to widely used biofluids such as blood and urine. However, its non-invasive, rapid, and cost-effective collection, together with its suitability for self-sampling, makes it attractive for clinical and personalized medicine. Moreover, saliva is expected to provide complementary metabolic information to other biofluids.

In this context, this study aims to identify an optimal sample collection and preparation protocol for maximizing informative metabolomic data from salivary NMR profiles. Four preparation protocols, selected and adapted from the literature, were systematically compared using the number of identified metabolites, their concentrations, and intra-day repeatability as evaluation criteria. Among them, an in-house adapted method combining centrifugation, freeze-drying and ultrafiltration proved most effective. This approach provided the broadest metabolome coverage, with 42 metabolites quantified. This method was further assessed using ANOVA to determine intra- and inter-day precision. Most metabolites demonstrated excellent repeatability (coefficients of variation below 10%), confirming protocol reliability for quantitative metabolomics. Overall, the optimized approach yields high-quality spectra, wide metabolic coverage and strong analytical precision, supporting reproducible salivary NMR metabolomics. Beyond its methodological contribution, this work highlights saliva as a promising biofluid for diagnostics, disease monitoring and personalized medicine, provided that collection and preparation procedures are appropriately standardized.

1 INTRODUCTION

2
3 Saliva, a complex organic liquid secreted by the parotid, submandibular and sublingual
4 glands, is mostly constituted of water (99 %) including electrolytes, proteins and enzymes, mucus,
5 hormones and other small molecules such as amino acids and nucleic acids ¹⁻³. This biofluid plays
6 several roles in the organism such as contributing to mastication, deglutition and food tasting.
7 Moreover, saliva is included in the digestion process thanks to its lubricant role and also provides
8 an antimicrobial function to protect oral cavity from microorganisms ^{1,2}. Compared to blood and
9 urine which are well-established matrices in metabolomics, saliva has been comparatively
10 underexplored. However, thanks to its numerous advantages such as its non-invasive, simple, fast
11 and inexpensive collection, this biofluid has showed a growing interest in the last years in
12 metabolomics ⁴⁻⁷. Since many of the components pass in salivary glands via blood by passive
13 diffusion or active transport, saliva could in many cases reflect the metabolomic changes observed
14 in blood ¹. Salivary metabolomics has been explored across multiple fields including oral and sports
15 medicine, toxicology, pharmacology, microbiology and nutrition ^{8,9}. As an example, the salivary
16 metabolome is used as a fluid to discover biomarkers in different oral and systemic pathologies
17 such as oral cancer ¹⁰⁻¹³, Alzheimer's ^{14,15}, Sjögren's syndrome ¹⁶⁻¹⁸, periodontal disease ¹⁹ and
18 other oral disorders ²⁰. It is also worth noting that some salivary metabolites reflect the functional
19 status of the oral microbiome ²¹. Understanding the functional roles of salivary metabolites, as well
20 as the interplay between the human metabolome and the salivary microbiota, presents an exciting
21 avenue for future research ⁵.

22 The potential of using in clinic salivary biomarkers as prognostic or diagnostic tools is
23 promising but many challenges are still to overcome. Indeed, the literature based on salivary
24 metabolomics shows a high heterogeneity between studies. Regarding pre-analytical steps, sample

collection and preparation are considerably variable from one study to another⁵. It has been shown previously that saliva collection significantly impacts the resulting metabolic profiles. For instance, in a study of Figueira et al.²², differences in composition and concentration of metabolites were found between stimulated, unstimulated and pure parotid saliva. Consequently, the protocol for salivary sample collection must be rigorously standardized across metabolomics studies to minimize variability arising from differences in collection methods. Moreover, salivary sample preparation protocols in metabolomics remain poorly standardized. This lack of consistency can introduce substantial variability, potentially leading to conflicting or misleading conclusions across metabolomics studies. To overcome this lack of accuracy and to use saliva as an interesting biofluid in clinic, there is a need to identify a standardized and validated protocol for the saliva preparation in metabolomics.

It is well established that the principal analytical platforms employed in metabolomics are mass spectrometry (MS), typically coupled with separative techniques such as liquid chromatography (LC) or gas chromatography (GC) and nuclear magnetic resonance (NMR) spectroscopy²³. With the growing emphasis on personalized medicine, an approach that adapts healthcare based on individual biological profiles, the integration of metabolomics into precision medicine requires the collection of longitudinal data and the implementation of quantitative or semi-quantitative metabolite analysis²⁴. Saliva appears to be a matrix particularly well-suited for longitudinal sample collection thanks to its previously described advantages, as well as the fact that it can be self-collected by the patient. Furthermore, although NMR presents certain limitations, notably lower sensitivity and spectral resolution compared to MS-based methods, it offers significant advantages. These include high reproducibility, inherent quantitative capability and

non-destructive analysis, making it a particularly suitable technique for clinical metabolomics applications and longitudinal studies ²⁵.

In most NMR-based metabolomics studies, saliva samples are only centrifuged before the addition of phosphate buffer containing internal standards for quantification as well as 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) or sodium trimethylsilyl-[2,2,3,3-*d*₄]-propionate (TSP) which are commonly used for chemical shift calibration at 0 ppm ⁴. Other studies also implemented an ultrafiltration step to remove macromolecules from samples and avoid signals due to these compounds on NMR spectra ^{1,10,26}. Indeed, large molecules such as proteins affect the spectral quality, specifically the baseline, and may consequently mask the signals of certain metabolites ²⁷. More recently, a protocol integrating centrifugation, ultrafiltration and freeze-drying was developed with the goal of increasing metabolites concentration and enhancing spectral sensitivity ²⁸. Their protocol enabled the quantification of 47 metabolites versus 22 in samples prepared without lyophilization, with a five-fold increase in metabolite concentrations and the detection of previously unobserved metabolites such as phenylalanine and tyrosine. However, they also noted the loss of volatile metabolites during the lyophilization process.

On the other hand, in MS-based salivary metabolomics, sample preparation often relies on centrifugation followed by solvent-based extraction protocols, typically using methanol and acetonitrile to precipitate proteins and improve metabolite recovery ^{1,3,29,30}. Although ultrafiltration could also be used ²², the addition of organic solvents remains central in many MS workflows. While these MS-oriented protocols differ from those used for NMR, outlining them helps to illustrate the diversity of salivary preparation strategies between analytical platforms.

Considering the advantages of NMR highlighted above, the present study aims to optimize a preparation protocol that maximizes relevant, informative and reproducible metabolic data from

70 salivary NMR profiles. In this work, we evaluated and compared three well-described and used
71 protocols from the literature, as well as a modified version of the protocol originally developed by
72 Quartieri et al.²⁸, in which we reversed the order of freeze-drying and ultrafiltration, by performing
73 freeze-drying prior to ultrafiltration (Figure 1). This modification was based on the hypothesis that
74 filtering the samples prior to lyophilization could increase metabolite loss during freeze-drying,
75 compared with lyophilizing the samples first. When ultrafiltration is initially performed,
76 macromolecules are removed, and volatile metabolites that were partially associated with them
77 may be more easily lost during lyophilization. All protocols have been compared and discussed
78 based on several criteria such as the number and the concentration of identified metabolites, the
79 repeatability of the method and the aspect and resolution of spectra. The intra and inter-day
80 precision of the best method will then be evaluated in order to identify the most informative and
81 repeatable protocol. Developing a method that ensures acceptable intra and inter-day repeatability
82 is crucial to confirm that the observed variations in samples reflect biological differences rather
83 than pre-analytical and analytical variability. Considering repeatability of 1D NMR is generally
84 around 1 % when the signal-to-noise ratio exceeds 200^{31,32}, observed variations above this value
85 reflect variability introduced during sample preparation, provided that the signal is however intense
86 enough. Besides, the repeatability of a method is acceptable if variations between replicate remains
87 below 10 %³³.

88 As highlighted above, stimulated and unstimulated saliva differ in composition and
89 metabolite concentrations. For these reasons, unstimulated whole saliva was selected and collected
90 using a passive drooling method into collection tubes. This type of saliva is the most commonly
91 used in salivary metabolomics studies⁸. In addition, its collection does not require the use of
92 chewing stimulators such as Salivette® or paraffin wax, or taste stimulators such as citric acid^{3,5,22}.

The use of such stimulants in NMR-based metabolomics may introduce unwanted spectral signals coming from compounds present in chewing materials or taste-stimulation agents. Finally, the use of unstimulated saliva avoids dilution effects typically observed in stimulated saliva. Indeed, the type and intensity of stimulation can induce unwanted dilution and alter metabolite levels, whereas passive secretion minimizes these effects ^{2,20}. Moreover, the collection protocol was designed to minimize external sources of contamination.

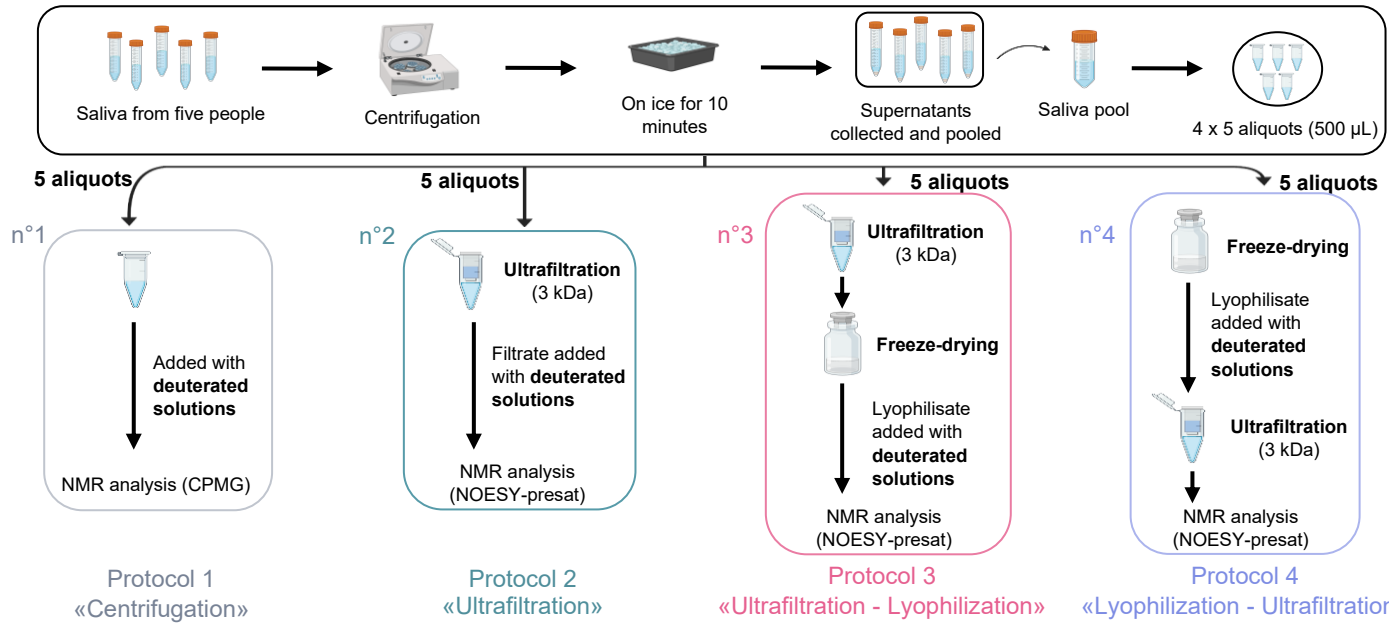


Figure 1. Description of the four salivary protocols. Protocols 1-3 are established methods described in the literature, whereas protocol 4 corresponds to our in-house procedure. To compare these methods, a single saliva pool was divided into 20 aliquots, and 5 aliquots were prepared by method.

EXPERIMENTAL PROCEDURES

Saliva samples were collected under standardized procedures from five healthy adult volunteers. The study was approved by the Hospitalo-Facultaire Ethics Committee of the University of Liège (reference: 707, September 23, 2014), under a general authorization for the collection of samples from healthy volunteers for technical research. All participants provided informed consent prior to inclusion in the study. All procedures were carried out in accordance with the principles of the Declaration of Helsinki.

Comparison of the four protocols

Sample collection

Saliva was collected under non-fasting conditions with participants instructed to wait at least 2 hours after food intake and toothbrushing with toothpaste. Volunteers rinsed their mouth twice with water without swallowing, after which unstimulated saliva was provided by a drooling method. Approximately 2-3 mL of saliva per participant was collected from five individuals into sterile tubes (Labsolute® Centrifuge Tube 15 mL, TH. Geyer Laboratory, Renningen, Germany). The tubes were then centrifuged for 15 min, 3730 g, 4°C, rested on ice for 10 min and supernatants were then collected and pooled. This pool was finally separated in twenty aliquots of 500 µL to have five aliquots for each method. All these samples were then stored at -80°C before the sample preparation.

Sample preparation: four protocols

All the filters (AmiconUltra® 0.5 mL 3kDa cut-off, Merck KGaA, Darmstadt, Germany) were first washed five times with 500 µL milli-Q water for 5 min, 13790 g at cooling temperature.

The different methods were all performed on 500 μ L of saliva per aliquot and finally diluted with phosphate buffer containing TSP and a solution of calcium formate as internal standard to have the same final concentration of calcium formate in all NMR tubes. Before selecting calcium formate as the internal standard, the presence of endogenous formate was investigated in saliva samples using a pooled sample prepared with protocol 1. Additional details are provided in Supporting Information.

Protocol 1 “Centrifugation”: Five aliquots from centrifuged pooled saliva were thawed and next added with 125 μ L of a buffer phosphate solution containing 0.5 mM of TSP as reference standard for the zero calibration and 50 μ L of a 3.5 mM calcium formate deuterated solution as internal standard for quantification. Samples were finally analysed by ^1H -NMR in 5 mm tubes.

Protocol 2 “Ultrafiltration”: Five aliquots from centrifuged pooled saliva were thawed and filtered through 3 kDa filters for 90 min, 13790 g, 4°C. Whole filtrates (385 μ L) were then collected and added with 240 μ L of a buffer phosphate solution containing 0.2 mM of TSP and 50 μ L of a 3.5 mM calcium formate deuterated solution. Samples were finally analysed by ^1H -NMR in 5 mm tubes.

Protocol 3 “Ultrafiltration – Freeze-drying”: Five aliquots from centrifuged pooled saliva were thawed and filtered through 3 kDa filters for 90 min, 13790 g, 4°C. Whole filtrates were then collected and freeze-dried for 22 hours. Lyophilizates were then added with 278 μ L of a buffer phosphate solution containing 0.075 mM of TSP and 22 μ L of a 3.5 mM calcium formate deuterated solution. Samples were finally analysed by ^1H -NMR in 3 mm tubes.

Protocol 4 “Freeze-drying - Ultrafiltration”: Five aliquots from centrifuged pooled saliva were thawed and freeze-dried for 22 hours. Lyophilizates were then added with 278 μ L of a buffer

phosphate solution containing 0.075 mM of TSP and 22 μ L of a 3.5 mM calcium formate deuterated solution. Samples were then filtered through 3 kDa filters for 90 min, 13790 g, 4°C. Filtrates were finally analysed by ^1H -NMR in 3 mm tubes.

NMR spectroscopy

All samples were recorded at 298 K on a Bruker® Neo spectrometer operating at 500.13 MHz for the proton acquisition. The instrument is equipped with a 5 mm TCI cryoprobe with a Z-gradient. ^1H -NMR spectra were acquired using a Carr-Purcell-Meiboom-Gill (CPMG) sequence (protocol 1) and a 1D nuclear Overhauser effect spectroscopy (NOESY) presat sequence with water presaturation (protocols 2,3 and 4). CPMG experiment used a RD-90-(t-180-t)_n acquire sequence with a relaxation delay (RD) of 4 s and a time between 180° pulses (t) of 0.0008 s. The number of transients was 64 and a number of 4 dummy scans was chosen. The acquisition time was set to 3.2767999 s. The NOESY presat experiment used a RD-90°-T1-90°-T_m-90°-acquire sequence with a relaxation delay of 4 s, a mixing time (T_m) of 0.01 s, and a fixed T1 delay of 0.000004 s. Water suppression pulse was placed during the RD. The number of transients was 64 and a number of 4 dummy scans was chosen. The acquisition time was set to 3.1457281 s.

Global spectral integration

Spectra were processed using MestReNova® software (version 14.3.1, Mestrelab Research, Santiago de Compostela, Spain). Baseline and phase correction was applied after which overall aliphatic (0.6 - 4.2 ppm) and overall aromatic (6.5 – 8.0 ppm) zones were integrated for each spectrum, assigning the integral value of 1 to the reference calcium formate. By method, the

mean of integral and its standard deviation (SD) have been calculated using Excel® (version 16.0, Microsoft, Redmond, WA, USA).

Quantification

Spectra were introduced in Chenomx® NMR Suite software (version 9.02, Chenomx Inc, Edmonton, Canada) and baseline and phase were corrected. Metabolite identification was first performed based on published literature data ¹, two-dimensional NMR experiments (¹H-¹H correlation spectroscopy (COSY) and ¹H-¹³C heteronuclear single quantum coherence (HSQC)), and the Human Metabolome Database (HMDB) ³⁴, and subsequently confirmed using the Chenomx® library. Based on the formate peak (8.46 ppm), a total of 44 distinct metabolites present in NMR spectra were then quantified across all four methods. Concentrations were finally inserted in Excel® (version 16.0, Microsoft, Redmond, WA, USA) to calculate means and standard deviations of each protocol as well as fold changes between protocols.

PCA analysis on buckets

Spectra were processed using MestReNova® software (version 14.3.1, Mestrelab Research, Santiago de Compostela, Spain). Baseline and phase corrections were applied, and the spectral intensities were normalized on the formate peak and reduced to integrated regions of equal width (0.04 ppm) corresponding to the 0.5-9.0 ppm region. The residual signals of water (4.5-5.1 ppm), TSP (-0.5-0.5 ppm) and the glycerol signals coming from the filter (3.53-3.60 ppm and 3.63-3.68 ppm) were removed before analysis. The reduced and normalized NMR spectral data were imported into SIMCA® (version 16.0.1, Sartorius, Goettingen, Germany). Buckets corresponding to calcium formate were excluded and unit variance (UV) scaling was applied to bucket tables. Unsupervised analysis (principal component analysis, PCA) was performed and a list with all PC

scores of samples from the first three components was extracted. Per method, the standard deviations of PC scores from five replicates were calculated with Excel® (version 16.0, Microsoft, Redmond, WA, USA).

Intra and inter-day precision of the chosen method

Sample collection and preparation

Saliva was collected under non-fasting conditions with participants instructed to wait at least 2 hours after food intake and toothbrushing with toothpaste. Volunteers rinsed their mouth twice with water without swallowing, after which unstimulated saliva was provided by a drooling method. Approximately 2-3 mL of saliva per participant was collected from four individuals into sterile tubes (Labsolute® Centrifuge Tube 15 mL, TH. Geyer Laboratory, Renningen, Germany). The tubes were then centrifuged for 15 min, 3730 g, 4°C, rested on ice for 10 min and supernatants were collected and pooled. This pool was finally separated in fifteen aliquots of 500 µL to have five aliquots by day and to repeat the analysis on three days. All these samples were then stored at -80°C before the sample preparation according to the protocol “Freeze-drying – Ultrafiltration”.

Statistical analyses

Spectra were first acquired using a NOESY presat experiment as described above. They were then introduced in Chenomx® NMR Suite software (version 9.02, Chenomx Inc, Edmonton, Canada) and baseline and phase were corrected. Using the formate peak (8.46 ppm) as a reference, 41 metabolites were identified and quantified. For each metabolite, an analysis of variance (ANOVA) was then performed to give both intra and inter-day precision of the protocol by extracting standard deviations of intra and inter-day variability using R version 4.3.1. Coefficients

210 of variation (CV) were then calculated with Excel® (version 16.0, Microsoft, Redmond, WA,
211 USA).

212

RESULTS

Comparison of the four protocols

First, the ^1H NMR spectra of saliva from methods 1, 2, 3 and 4 (centrifuged, ultrafiltered, ultrafiltered/freeze-dried and freeze-dried/ultrafiltered saliva, respectively) were qualitatively compared by assigning the same integral value to the formate peak (Figure 2A). As seen in Figures 2B and C, which display the different spectra in the aliphatic region between 0.6 and 4.4 ppm, the spectrum from protocol 1 still exhibits huge broad signals attributable to macromolecules (0.8-1.0 ppm; 1.6-1.8 ppm; 1.9-2.5 ppm) affecting the baseline compared to the other three methods. Moreover, the ultrafiltration step (methods 2, 3 and 4) results in a flatter baseline and better resolved peaks than the centrifugation-only protocol (method 1), with no residual macromolecular signals. In addition, methods 3 and 4, which include both ultrafiltration and freeze-drying steps (but in different order) show an overall more intense signal in the aliphatic zone compared to the methods without this step. This is further illustrated in Table 1, which presents the global integral values of each spectrum for both the aliphatic and aromatic zones. Protocol 2, which involves only ultrafiltration, shows an overall aliphatic integral almost four times lower than protocols 3 and 4 and nearly three times lower than protocol 1. These observations indicate that freeze-drying step has improved the signal-to-noise ratio in the aliphatic zone by concentrating samples, whereas ultrafiltration alone, despite yielding cleaner spectra, results in lower signal intensity than method 1 employing only centrifugation.

Besides, in the aromatic zone between 6.5 and 8.0 ppm (Figure 2D), the spectrum of saliva prepared with protocol 1 shows the most intense overall signal compared to the other methods, followed by protocol 4, 3 and finally 2 (see Table 1). Thus, while ultrafiltration contributes to the

loss of aromatic signals, performing freeze-drying prior to ultrafiltration (protocol 4) allows to minimize this loss of signal. Additionally, the standard deviations of the integrals from five replicates in protocol 1 are higher than those observed in the other methods for both aliphatic and aromatic regions, concluding in a higher analytical variability.

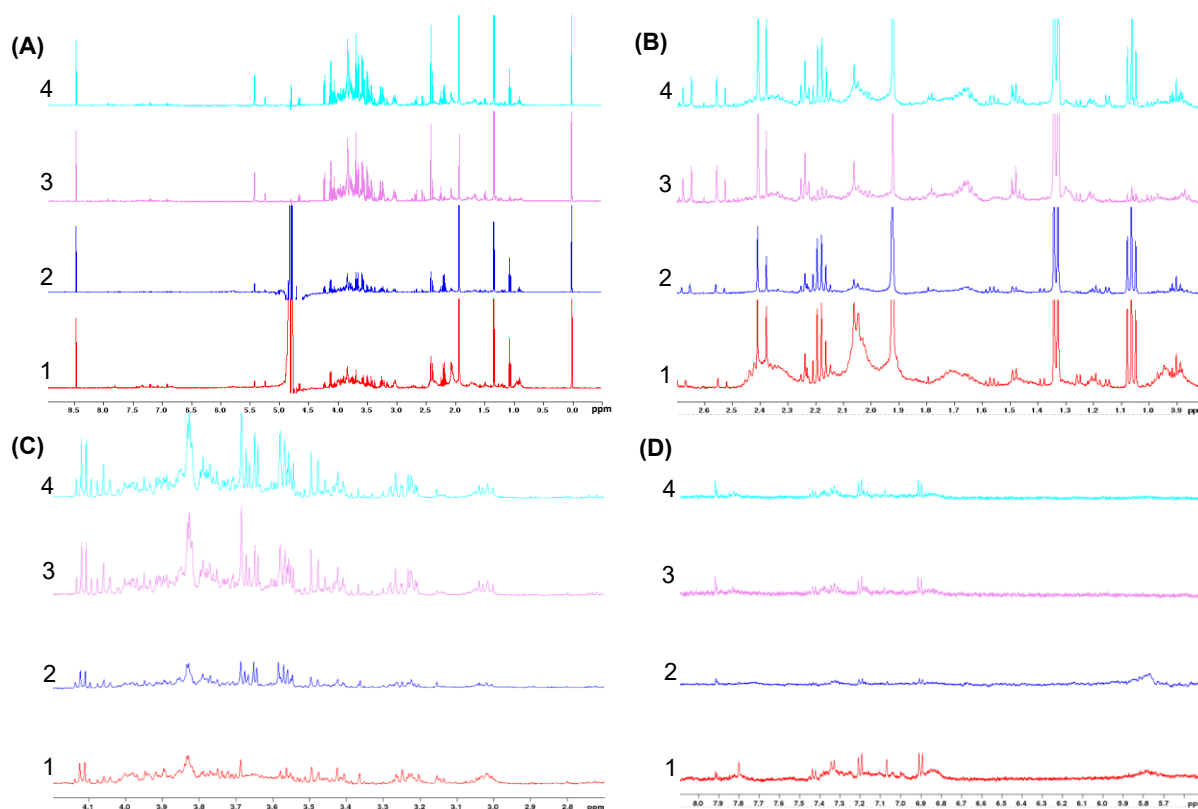


Figure 2. (A) Whole ¹H-NMR spectra of salivary samples (1) only centrifuged, (2) centrifuged and ultrafiltered, (3) centrifuged, ultrafiltered and then freeze-dried, (4) centrifuged, freeze-dried and then ultrafiltered. (B), (C), (D) Expanded regions of the ¹H-NMR spectra : (B) 0.8 to 2.7 ppm; (C) 2.7 to 4.2 ppm; (D) 5.5 to 8.1 ppm.

Table 1. Global integral values $\pm$ SD of aliphatic and aromatic zones of the different methods.

Integrals were calculated and normalized to the internal standard (calcium formate)

	Protocol 1	Protocol 2	Protocol 3	Protocol 4
Aliphatic zone (0.6 – 4.2 ppm)	62.6 $\pm$ 16.1	23.6 $\pm$ 2.0	74.7 $\pm$ 4.0	82.4 $\pm$ 6.5
Aromatic zone (6.5 – 8.0 ppm)	1.7 $\pm$ 1.2	0.5 $\pm$ 0.1	0.6 $\pm$ 0.1	1.1 $\pm$ 0.03

Furthermore, to deepen the analysis, metabolite identification and quantification were performed using Chenomx software, aiming to detect the maximum number of compounds possible. It is important to note, however, that the quantification of metabolites obtained with method 1 is not optimal, as macromolecular signals remain in the spectra despite the use of a CPMG sequence, resulting in a less flat baseline compared with protocols 2, 3 and 4 including a filtration step. As expected, the number and concentration of metabolites vary between methods.

A total of 44 distinct metabolites is detected across all four methods. Among these, 40, 39, 37 and 42 metabolites are identified and quantified in the 1D spectra obtained with methods 1 (centrifugation), 2 (ultrafiltration), 3 (ultrafiltration followed by freeze-drying) and 4 (freeze-drying followed by ultrafiltration), respectively. Importantly, most of these metabolites are shared across all protocols, indicating a substantial overlap in metabolite coverage despite the different preparation procedures. However, while the qualitative profiles are similar, the quantitative levels of metabolite vary between protocols. As presented in Table 2, which reports the mean concentrations of each metabolite by method, protocol 4 reveals the highest concentrations for 26 of the 44 detected metabolites, compared to 10, 2 and 8 for protocols 1, 2 and 3, respectively.

Besides, the protocols without lyophilization (1 and 2) allow an increased detection of acetate, acetoin, ethanol, trimethylamine and urea, in comparison to those including a freeze-drying step (3 and 4). However, nearly all other metabolites show decreased concentrations in the absence of lyophilization. Additionally, 3-phenylpropionate and glycine are not detected in protocols 1 and 2 but are identified in protocols 3 and 4. Moreover, method 1 does not allow the quantification of acetoacetate and creatine while aspartate, isoleucine and ornithine are not detected in method 2. Furthermore, compared to the “freeze-drying followed by ultrafiltration” protocol (method 4), the protocol involving ultrafiltration followed by freeze-drying (method 3) does not allow the detection of acetoin, butyrate, isocaproate, isovalerate and urea (Table 2).

Table 2. Final metabolite concentrations determined by 1D ¹H NMR after only centrifugation or ultrafiltration or ultrafiltration followed by freeze-drying or freeze-drying followed by ultrafiltration

Metabolite (μM)	Protocol 1	Protocol 2	Protocol 3	Protocol 4
<i>3-Phenylpropionate</i>	nd	nd	3.6 ± 0.5	4.04 ± 0.2
<i>5-Aminopentanoate</i>	17.6 ± 3.6	15.5 ± 0.6	55.2 ± 2.7	54.8 ± 3.4
<i>Acetate</i>	449.3 ± 8.0	319.9 ± 3.5	83.3 ± 4.0	312.0 ± 25.0
<i>Acetoacetate</i>	nd	0.7 ± 0.1	1.9 ± 0.2	2.3 ± 0.3
<i>Acetoin</i>	3.6 ± 0.7	2.62 ± 0.2	nd	1.6 ± 0.4
<i>Alanine</i>	6.2 ± 0.4	4.3 ± 0.2	15.8 ± 0.8	15.7 ± 1.3
<i>Aspartate</i>	5.6 ± 1.8	nd	6.4 ± 0.9	6.1 ± 0.3
<i>Betaine</i>	0.5 ± 0.2	0.6 ± 0.1	1.7 ± 0.5	1.9 ± 0.2
<i>Butyrate</i>	13.6 ± 1.5	13.2 ± 0.5	nd	22.4 ± 2.4
<i>Choline</i>	2.9 ± 0.2	1.2 ± 0.04	4.0 ± 0.1	4.2 ± 0.2
<i>Citrate</i>	9.2 ± 0.8	11.6 ± 0.4	41.4 ± 1.8	40.5 ± 1.7
<i>Creatine</i>	nd	3.1 ± 0.3	9.8 ± 0.5	9.8 ± 0.9
<i>Creatinine</i>	3.1 ± 0.3	1.6 ± 0.4	3.6 ± 0.3	3.8 ± 0.3
<i>Dimethylamine</i>	0.8 ± 0.1	0.7 ± 0.1	1.6 ± 0.3	1.9 ± 0.2
<i>Ethanol</i>	6.9 ± 0.3	7.6 ± 0.3	nd	nd
<i>Ethanolamine</i>	21.1 ± 1.2	4.5 ± 0.9	15.9 ± 0.7	15.5 ± 1.1
<i>Fucose</i>	10.5 ± 1.1	1.9 ± 0.2	7.5 ± 0.3	7.1 ± 0.6

<i>Glucose</i>	76.7 ± 6.3	32.7 ± 4.1	116.2 ± 9.1	119.4 ± 13.6
<i>Glycerophosphocholine</i>	1.4 ± 0.2	1.5 ± 0.2	5.5 ± 0.3	5.6 ± 0.4
<i>Glycine</i>	nd	nd	22.4 ± 1.6	22.6 ± 0.7
<i>Glycolate</i>	10.5 ± 0.4	7.6 ± 0.5	27.0 ± 2.1	29.0 ± 2.6
<i>Histidine</i>	15.1 ± 0.4	2.3 ± 0.3	4.9 ± 0.3	5.7 ± 0.3
<i>Isocaproate</i>	1.1 ± 0.07	1.2 ± 0.1	nd	2.4 ± 0.3
<i>Isoleucine</i>	1.7 ± 0.3	nd	1.7 ± 0.3	1.2 ± 0.1
<i>Isovalerate</i>	1.4 ± 0.1	1.5 ± 0.1	nd	4.0 ± 0.4
<i>Lactate</i>	150.4 ± 17.7	170.1 ± 9.6	481.0 ± 14.9	554.6 ± 26.1
<i>Leucine</i>	3.2 ± 0.1	1.3 ± 0.1	4.7 ± 0.2	5.1 ± 0.5
<i>Methylamine</i>	0.9 ± 0.1	0.6 ± 0.1	2.2 ± 0.2	2.3 ± 0.2
<i>O-Phosphoethanolamine</i>	29.8 ± 2.6	22.9 ± 0.9	82.1 ± 5.4	81.6 ± 4.6
<i>Ornithine</i>	3.9 ± 0.8	nd	5.7 ± 0.8	5.8 ± 0.9
<i>Phenylacetate</i>	1.5 ± 0.2	1.4 ± 0.5	3.1 ± 0.2	3.5 ± 0.5
<i>Phenylalanine</i>	7.5 ± 0.8	2.0 ± 0.3	6.0 ± 0.5	6.7 ± 0.5
<i>Proline</i>	12.8 ± 2.3	5.0 ± 1.3	17.3 ± 4.4	18.2 ± 3.4
<i>Propionate</i>	90.8 ± 4.7	79.0 ± 5.3	13.1 ± 1.0	81.7 ± 7.2
<i>Putrescine</i>	3.3 ± 0.7	3.0 ± 0.3	12.0 ± 0.7	12.2 ± 0.4
<i>Pyruvate</i>	22.5 ± 1.5	18.1 ± 0.5	34.6 ± 1.7	50.7 ± 2.9
<i>Sarcosine</i>	0.6 ± 0.1	0.7 ± 0.1	2.6 ± 0.3	2.5 ± 0.2
<i>Succinate</i>	16.7 ± 0.5	17.5 ± 0.6	60.9 ± 1.6	66.7 ± 4.0
<i>Sucrose</i>	33.3 ± 1.7	50.5 ± 1.7	196.3 ± 9.5	199.4 ± 15.4
<i>Taurine</i>	17.4 ± 2.6	15.4 ± 1.2	52.9 ± 6.6	65.2 ± 4.9
<i>Trimethylamine</i>	0.6 ± 0.1	0.4 ± 0.1	nd	nd
<i>Tyrosine</i>	13.4 ± 0.4	3.2 ± 0.1	10.1 ± 0.8	9.8 ± 0.6
<i>Urea</i>	89.7 ± 12.4	204.3 ± 16.5	nd	29.5 ± 1.2
<i>Valine</i>	1.3 ± 0.2	1.0 ± 0.1	3.5 ± 0.2	3.8 ± 0.3

273 Note : nd = not detected

274

275 Finally, as shown in Table 3, presenting the fold changes of each metabolite between two
276 methods, the centrifugation-only protocol (method 1) yields higher signal intensities for most
277 metabolites compared to protocol 2 (centrifugation followed by ultrafiltration), with the exception
278 of acetoacetate, betaine, citrate, creatine, ethanol, glycerophosphocholine, isocaproate, isovalerate,
279 lactate, sarcosine, succinate, sucrose and urea. In contrast, when a freeze-drying step is added to
280 the ultrafiltration process (protocols 3 and 4), the trend is reversed. Indeed, most metabolites
281 exhibit higher concentrations in protocols 3 and 4 compared to protocol 1, with the exception of

fucose and urea as well as certain volatile aliphatic compounds (acetate, acetoin, ethanolamine and propionate) and aromatic amino acids such as histidine, phenylalanine and tyrosine.

Furthermore, to evaluate the impact of the order of manipulations, fold changes between methods 3 and 4 were also calculated. Results indicate that performing freeze-drying prior to ultrafiltration (protocol 4) gives higher concentrations for nearly all metabolites, except for 5-aminopentanoate, alanine, aspartate, citrate, ethanolamine, fucose, isoleucine, O-phosphoethanolamine, sarcosine and tyrosine. However, these ten last metabolites have a fold change very close to 1, indicating minimal differences between the two protocols. An interesting finding is that volatile compounds such as acetoin, acetate, butyrate, isocaproate, isovalerate, propionate and urea are decreased or even absent in protocol 3 in contrast to protocol 4. This observation suggests that the sequence of preparative steps (freeze-drying and ultrafiltration) significantly impacts the recovery and detection of volatile metabolite signals in the resulting spectra.

Table 3. Fold changes of mean metabolites concentration between different protocols: only centrifuged saliva (method 1); ultrafiltered saliva (method 2); ultrafiltered then freeze-dried saliva (method 3); freeze-dried then ultrafiltered saliva (method 4)

Metabolite	Fold change between method 1 and 2 ^a	Fold change between method 1 and 3 ^b	Fold change between method 1 and 4 ^c	Fold change between method 3 and 4 ^d
<i>3-Phenylpropionate</i>	n.a	n.a	n.a	0.89
<i>5-Aminopentanoate</i>	1.13	0.32	0.32	1.01
<i>Acetate</i>	1.40	5.39	1.44	0.27
<i>Acetoacetate</i>	n.a	n.a	n.a	0.83
<i>Acetoin</i>	1.36	n.a	2.23	n.a
<i>Alanine</i>	1.42	0.39	0.39	1.01
<i>Aspartate</i>	n.a	0.88	0.92	1.05
<i>Betaine</i>	0.83	0.29	0.26	0.89
<i>Butyrate</i>	1.03	n.a	0.61	n.a
<i>Choline</i>	2.34	0.72	0.67	0.94
<i>Citrate</i>	0.80	0.22	0.23	1.02
<i>Creatine</i>	n.a	n.a	n.a	1.00

<i>Creatinine</i>	1.94	0.86	0.82	0.95
<i>Dimethylamine</i>	1.14	0.50	0.42	0.84
<i>Ethanol</i>	0.90	n.a	n.a	n.a
<i>Ethanolamine</i>	4.71	1.33	1.36	1.03
<i>Fucose</i>	5.56	1.39	1.47	1.06
<i>Glucose</i>	2.35	0.66	0.64	0.97
<i>Glycerophosphocholine</i>	0.93	0.25	0.25	0.98
<i>Glycine</i>	n.a	n.a	n.a	0.99
<i>Glycolate</i>	1.38	0.39	0.36	0.93
<i>Histidine</i>	6.44	3.05	2.62	0.86
<i>Isocaproate</i>	0.51	n.a	0.25	n.a
<i>Isoleucine</i>	n.a	1.00	1.39	1.39
<i>Isovalerate</i>	0.91	n.a	0.34	n.a
<i>Lactate</i>	0.88	0.31	0.27	0.87
<i>Leucine</i>	2.48	0.67	0.62	0.92
<i>O-Phosphoethanolamine</i>	1.30	0.36	0.37	1.01
<i>Ornithine</i>	n.a	0.68	0.67	0.98
<i>Phenylacetate</i>	1.59	0.71	0.62	0.88
<i>Phenylalanine</i>	3.74	1.25	1.11	0.89
<i>Proline</i>	2.56	0.74	0.70	0.95
<i>Propionate</i>	1.15	6.93	1.11	0.16
<i>Putrescine</i>	1.09	0.27	0.27	0.98
<i>Pyruvate</i>	1.24	0.65	0.44	0.68
<i>Sarcosine</i>	0.86	0.23	0.24	1.04
<i>Succinate</i>	0.95	0.27	0.25	0.91
<i>Sucrose</i>	0.66	0.17	0.17	0.98
<i>Taurine</i>	1.13	0.33	0.27	0.81
<i>Trimethylamine</i>	1.50	n.a	n.a	n.a
<i>Tyrosine</i>	4.17	1.33	1.36	1.03
<i>Urea</i>	0.49	n.a	3.05	n.a
<i>Valine</i>	1.35	0.37	0.35	0.94

Note: n.a = not applicable (at least one metabolite mean concentration is equal to 0). ^{a,b,c}Values > 1 indicates higher and <1 lower metabolite concentration in centrifugation method (protocol 1) compared to method 2, 3 or 4. ^dValues > 1 indicates higher and <1 lower metabolite concentration in ultrafiltration/freezing-drying method (protocol 3) compared to freezing-drying/ultrafiltration method (protocol 4).

In addition to the previously discussed results, it is also important to assess the intra-day repeatability of each method. For this purpose, a PCA was performed on the binned spectral data, as described in the Methods section (Figure 3). PCA scores for PC1, PC2 and PC3 were extracted from each sample, and the standard deviations of these components were then calculated for each protocol (Table 4). The results indicate that the least repeatable method is the centrifugation-only protocol (method 1), followed by method 2 (centrifugation plus ultrafiltration). When comparing the two methods that include both ultrafiltration and freeze-drying steps (protocols 3 and 4), the standard deviations are similar, although protocol 3 (the ultrafiltration followed by freeze-drying) shows the lowest SD values. To further investigate the origin of the higher variability observed for protocol 1, the spectral regions contributing most to this variability were identified using the PCA loading plots (data not shown). The regions mainly responsible for the dispersion were found between 1.98 and 3.02 ppm, containing signals from macromolecules, and between 4.18 and 4.50 ppm, a region close to the residual water signal. Inspection of the overlapped ¹H-NMR spectra from the five replicates prepared using protocol 1 (see Figure S2 in Supporting Information) showed that these regions do not properly overlap, reflecting differences in baseline height across samples.

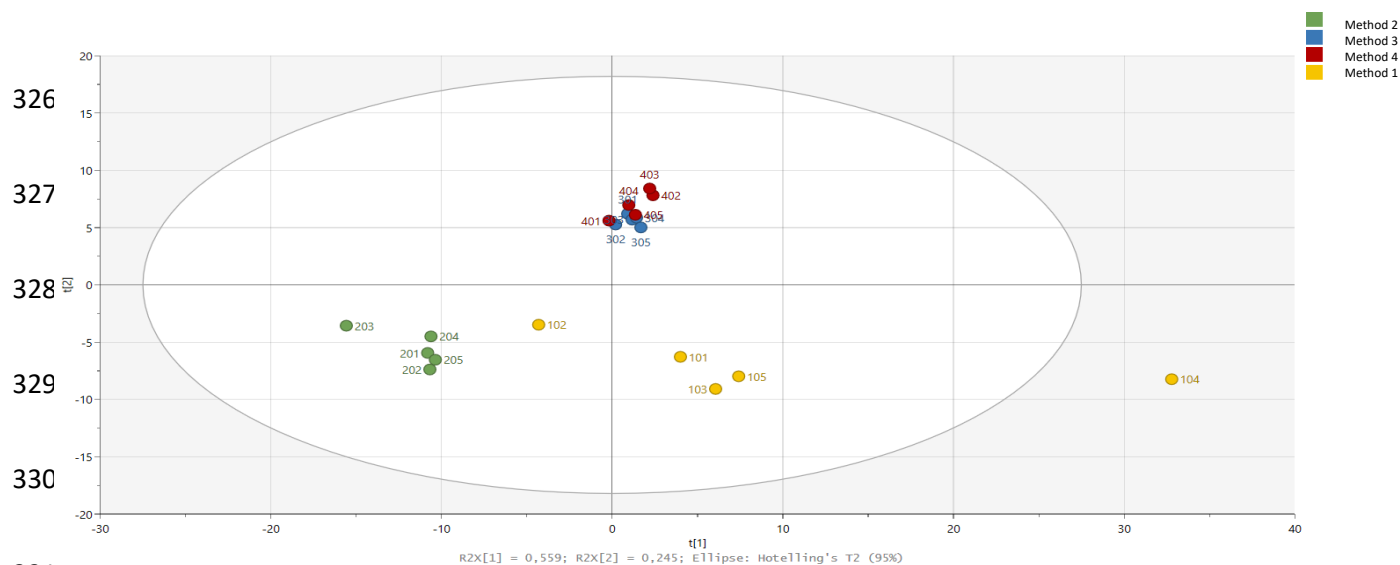


Figure 3. PCA analysis (based on buckets) of all samples including the five replicates per method.

Table 4. Standard deviations of the 3 PCA scores for each of the saliva preparation methods (coming from the PCA plots in Figure 3)

	PC1 (72.1 %)	PC2 (23.7 %)	PC3 (3.8 %)
Protocol 1	13.94	2.23	7.16
Protocol 2	2.22	1.55	2.01
Protocol 3	0.58	0.48	0.58
Protocol 4	1.03	1.14	0.92

Taken together, these observations indicate that although protocol 3 exhibited the highest intra-day repeatability, protocol 4 showed the second-highest repeatability, with values close to those observed for protocol 3. Accordingly, the protocol employing a centrifugation followed by a freeze-drying and finally an ultrafiltration (protocol 4) yields to a good spectral resolution, an overall better signal-to-noise ratio, the best covering of saliva metabolome and a good intra-day repeatability, representing the best compromise among the evaluated protocols. A further analysis of the chosen protocol needs however to be done to study both the intra and inter-day precision in order to fully validate this protocol for routine metabolomics applications.

To achieve this goal, saliva samples collected from four individuals were centrifuged, and the resulting supernatants were pooled. This pooled sample was then divided into fifteen aliquots, stored at -80°C. Each day, five replicates were prepared according to the selected “Freeze-drying/ultrafiltration” protocol. After quantifying 41 metabolites, an ANOVA was performed for each metabolite to separately assess intra and inter-day variability. CV were finally extracted from the analysis (see Figure 4).

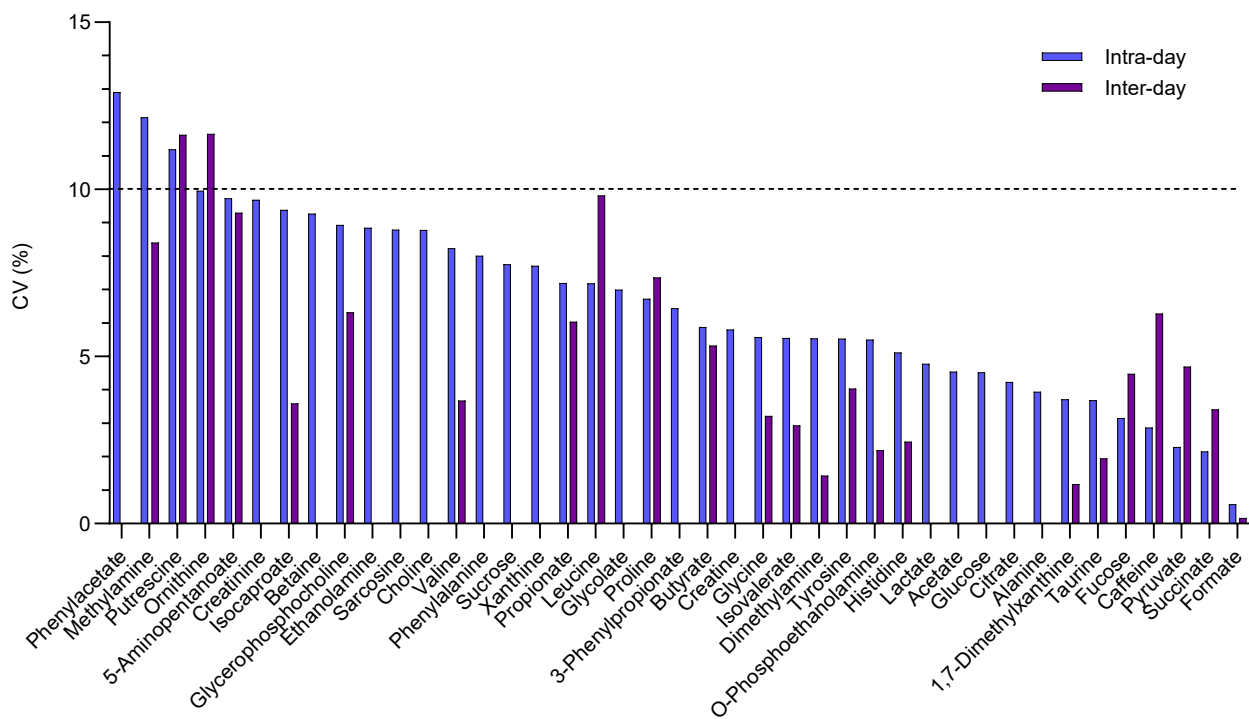


Figure 4. Analysis of the intra and inter-day precision of the chosen protocol « freeze-drying-ultrafiltration »: Intra-day and inter-day Coefficients of Variation (CV %) of each metabolite extracted from the ANOVA analyses.

As explained above, the repeatability of the method is acceptable if variations between replicate remains below 10 %. As results, three metabolites, phenylacetate, methylamine and putrescine have intra-day CVs slightly above 10 %, with values of 12.91, 12.16 and 11.21 %, respectively. Five other metabolites, ornithine, 5-aminopentanoate, creatinine, isocaproate and betaine have a CV value just below the threshold of 10 % (9.96, 9.73, 9.68, 9.39 and 9.28 %, respectively). Regarding inter-day variability, putrescine again shows a CV above 10 % (11.64 %) as well as ornithine (11.66 %), while 5-aminopentanoate and leucine have CVs slightly below but close to the 10% cut-off (9.30 and 9.82 %, respectively). To further investigate whether higher CV values were related to metabolite concentration, Table 5 indicates that metabolites exhibiting higher intra-day, inter-day, or both CVs (> 8 %) are generally those with the lowest mean concentrations, corresponding to lower signal-to-noise ratio.

Table 5. Mean metabolite concentrations obtained per day (5 replicates) for intra and inter-day precision evaluation of protocol 4, along with intra- and inter-day CVs. Metabolites are ranked from the lowest to the highest concentration.

Metabolite	Day 1 (μ M)	Day 2 (μ M)	Day 3 (μ M)	Intra-day CV (%)	Inter-day CV (%)
<i>Valine</i>	0.6	0.6	0.6	8.2	3.7
<i>Isocaproate</i>	0.7	0.7	0.8	9.4	3.6
<i>Methylamine</i>	0.7	0.8	0.6	12.2	8.4
<i>Dimethylamine</i>	0.8	0.8	0.8	5.5	1.4
<i>Choline</i>	0.8	0.9	0.8	8.8	0.0
<i>Sarcosine</i>	1.0	1.0	1.0	8.8	0.0
<i>Isovalerate</i>	1.0	1.0	1.1	5.6	2.9
<i>Betaine</i>	1.2	1.2	1.2	9.3	0.0
<i>Glycerophosphocholine</i>	1.6	1.4	1.4	8.9	6.3
<i>1,7-Dimethylxanthine</i>	1.7	1.7	1.7	3.7	1.2
<i>3-Phenylpropionate</i>	1.7	1.7	1.7	6.4	0.0
<i>Creatinine</i>	1.8	1.9	1.9	9.7	0.0
<i>Leucine</i>	1.8	1.6	1.5	7.2	9.8
<i>Phenylacetate</i>	2.6	2.5	2.5	12.9	0.0
<i>Phenylalanine</i>	4.1	3.8	3.9	8.0	0.0
<i>Fucose</i>	4.3	4.1	3.9	3.2	4.5
<i>Histidine</i>	4.3	4.2	4.1	5.1	2.4
<i>Creatine</i>	4.4	4.3	4.4	5.8	0.0
<i>Caffeine</i>	5.1	4.9	4.5	2.9	6.3
<i>Ethanolamine</i>	5.2	5.3	5.1	8.8	0.0
<i>Alanine</i>	6.2	6.0	6.0	3.9	0.0
<i>Ornithine</i>	6.2	5.1	4.9	10.0	11.7
<i>Proline</i>	8.7	7.4	8.1	6.7	7.4
<i>Tyrosine</i>	10.8	10.0	10.0	5.5	4.0
<i>Putrescine</i>	11.2	9.2	8.9	11.2	11.6
<i>Sucrose</i>	12.7	12.1	12.1	7.8	0.0
<i>Xanthine</i>	13.5	13.5	13.4	7.7	0.0
<i>Butyrate</i>	14.7	13.2	13.5	5.9	5.3
<i>Glycine</i>	17.6	16.3	17.3	5.6	3.2
<i>Taurine</i>	18.9	18.4	17.9	3.7	2.0
<i>Glucose</i>	19.7	19.3	19.3	4.5	0.0
<i>Glycolate</i>	20.8	20.6	19.7	7.0	0.0
<i>Pyruvate</i>	29.6	27.2	27.3	2.3	4.7
<i>Succinate</i>	31.9	30.0	29.9	2.2	3.4
<i>Citrate</i>	35.8	34.9	36.2	4.2	0.0
<i>O-Phosphoethanolamine</i>	47.0	44.8	44.1	5.5	2.2
<i>5-Aminopentanoate</i>	54.9	48.3	44.9	9.7	9.3
<i>Propionate</i>	72.4	63.9	65.1	7.2	6.0
<i>Lactate</i>	236.4	242.5	237.7	4.8	0.0
<i>Formate</i>	279.4	279.7	281.0	0.6	0.2
<i>Acetate</i>	329.9	318.3	320.8	4.5	0.0

369 Note: Metabolites shown in bold present intra- and/or inter-day CV values close to (between 8 and
370 10%) or exceeding the 10% threshold.

371

DISCUSSION

In many previous articles, saliva samples for metabolomics analysis were mainly prepared using a simple centrifugation step to partially remove macromolecules from the sample matrix. Gardner et al.⁴ developed a ¹H NMR-based metabolomics protocol that includes such a centrifugation step. However, it seems that this type of preparation is not optimal since we demonstrated in this study that macromolecules signals are present in the samples. Thus, the use of a CPMG pulse sequence is therefore mandatory to suppress subsequent signal of large molecules. Nevertheless, this sequence is not ideally suited for absolute quantification of metabolites, as it only partially removes broad signals from macromolecules, leading to a less flat baseline compared to the NOESY-presat sequence. Moreover, an accurate quantification is a critical consideration in clinical and precision medicine applications.

The increased inter-replicate variability observed for protocol 1 compared with the other methods further supports these observations. In particular, the spectral region between 1.98 and 3.02 ppm was identified as a major contributor to the variability of this protocol. This region is dominated by residual macromolecular signals whose attenuation by the CPMG T₂ filter can vary between replicates, depending on macromolecular content and relaxation properties³⁵, leading to baseline fluctuations. In addition, the spectral region between 4.18 and 4.50 ppm, located close to the residual water signal, was also found to be more variable between replicates. While the NOESY-presat sequence has been specifically developed to achieve efficient water suppression, the use of water presaturation in combination with a CPMG T₂ filter can be more sensitive to sample composition, potentially resulting in increased baseline distortions in regions adjacent to the water signal. Taken together, these effects contribute to the reduced intra-day repeatability

observed for protocol 1 compared with protocols including an ultrafiltration step and the use of a NOESY-presat sequence.

This lack of quality has led various researchers to complement centrifugation with an additional ultrafiltration step^{1,10,26}, following protocols commonly applied in blood sample preparation. However, despite the use of filters in some studies, comparison of saliva only centrifuged or centrifuged with an ultrafiltration has rarely been discussed. To our knowledge, the only study comparing these two methods has been done by Quartieri et al.²⁸, who reported that removing proteins signals improved spectral quality.

In our study, ultrafiltration also improved spectral quality by eliminating broad signals from mucus, proteins and other particulates, resulting in a flatter baseline. As detailed in the Results section, the appearance of spectra and comparison of overall integral values in the aliphatic region between the protocols 1 (centrifugation) and 2 (ultrafiltration) revealed that centrifuged samples retained, as expected, residual macromolecular signals. Spectra from these samples exhibited nearly threefold higher overall integrals compared to those prepared with an additional ultrafiltration step. Although partial retention of metabolites on the filter cannot be excluded, the increased global integral observed in centrifuged samples could reflect the residual macromolecular signals that affect the baseline and were therefore included in the integral values. Moreover, although we identified and quantified a similar number of metabolites in both protocols (40 metabolites in protocol 1 and 39 metabolites in protocol 2), the metabolic profiles differed, with certain metabolites detected exclusively in one protocol or the other. Additionally, it is worth noting that metabolite concentrations are generally lower in ultrafiltered samples, particularly in the aromatic region. While partial retention of metabolites by the filter cannot be entirely excluded, previous works have shown that aromatic compounds, including histidine, tyrosine, and

phenylalanine, strongly bind to proteins ³⁶. These metabolites could therefore be retained on the filter due to these interactions, resulting in decreased concentrations of aromatic compounds in the filtrate.

Because ultrafiltration leads to a loss of sensitivity, adding a freeze-drying step may be required to improve sensitivity while preserving spectral quality. Thus, researchers and particularly Quartieri et al. ²⁸ implemented a lyophilization step of salivary samples in addition to ultrafiltration. Their study optimized the protocol by first centrifuging, ultrafiltering, and finally freeze-drying the samples. In our study, we further evaluated whether the order of ultrafiltration and lyophilization affects the quality of ¹H NMR spectra. Our results indicate that the sequence of preparation steps is indeed critical. When lyophilization is performed before ultrafiltration, the resulting spectra exhibit an overall more intense signal compared to the reverse order. Using this approach, 42 metabolites are identified and quantified (protocol 4) against 37 when ultrafiltration preceded lyophilization (protocol 3). Furthermore, some volatile metabolites are decreased or even absent in method 3 compared to method 4. We hypothesize that during freeze-drying, macromolecules retain certain small volatile metabolites by binding to them, thereby preventing their evaporation. In contrast, ultrafiltration before lyophilization removes these macromolecular complexes, resulting in the loss of volatile compounds.

Surprisingly, among the four tested protocols in the present work, the method 3 (ultrafiltration - freeze-drying) is the protocol leading to the least number of identified metabolites which contrasts with the findings reported by Quartieri et al. ²⁸. In their study, only 13 metabolites were identified in centrifuged samples (protocol 1) against 22 metabolites when ultrafiltration was applied (protocol 2) and up to 47 metabolites when a subsequent freeze-drying step was included (protocol 3). A critical detail in their methodology is the use of ¹H-NMR NOESY-presat sequences

for all the samples. This likely limited metabolite detection in method 1 (centrifuged samples) because of the large macromolecular signals in spectra. In contrast, our study employed a CPMG sequence for this method, which attenuates broad signals from macromolecules and therefore enabling the identification of 40 metabolites. Regarding the impact of the freeze-drying step following ultrafiltration (method 3), as highlighted by Quartieri et al.²⁸, our comparison of protocols 2 (“ultrafiltration”) and 3 (“ultrafiltration - freeze-drying”) similarly shows increased concentration for most of the metabolites in the “ultrafiltration-freeze-drying” protocol compared to ultrafiltration alone. Nevertheless, contrary to their results, the total number of metabolites identified in ultrafiltered and subsequently freeze-dried samples (protocol 3) was lower than in only ultrafiltered samples (protocol 2). This decrease can be essentially attributed to the loss of volatile metabolites during the lyophilization process.

Next, it is important to note that methodological divergences between the two studies may partly account for these differences, including the number of metabolites identified, with 47 reported in their study compared with 37 in the present work using the same protocol. While Quartieri et al.²⁸ analysed submandibular and sublingual saliva, a saliva type rarely investigated in the literature, our work focused on unstimulated whole saliva, which is more widely used in metabolomics studies. The use of a different type of saliva could therefore explain variations in metabolite profiles, concentrations, and the resulting conclusions. Furthermore, the work of Quartieri et al.²⁸ did not report inter-day repeatability and provided only limited information on intra-day repeatability, based on three samples and seven metabolites. Consequently, among the 47 metabolites they identified, it remains uncertain whether all yield sufficient analytical repeatability to support reliable application in a clinical context.

Finally, the repeatability and inter-day precision of our protocol of choice (freeze-drying - ultrafiltration; method 4) were evaluated, which, to our knowledge, have never been previously studied for salivary metabolomics using NMR. Results show that this in-house method provides good intra and inter-day precision. Only three metabolites, methylamine, phenylacetate and putrescine have intra-day CV slightly above 10%. For inter-day precision, only ornithine and putrescine exceed this threshold. First, the variability observed for methylamine is likely due to its low signal-to-noise ratio (approximately 5). Similarly, the higher variability of phenylacetate may result both from its low signal-to-noise ratio (around 5) and from the overlap of its signal with the one of phenylalanine, thus reducing quantification accuracy. On the other hand, the increased variability of ornithine and putrescine could be the result of a fluctuating bacterial activity. Indeed, putrescine is a metabolite produced by microorganisms from salivary microbiota via the decarboxylation of ornithine catalysed by ornithine decarboxylase³⁷. Residual microbial activity may persist in the samples despite -80°C storage, potentially leading to ongoing production of both metabolites. Moreover, it is well established that NMR repeatability decreases as the signal-to-noise ratio decreases³¹. Consequently, higher CV values are more likely to be observed for metabolites present at low concentrations, as presented in Results section. In this context, it is noteworthy that, aside from the metabolites discussed above with CVs exceeding 10%, other metabolites with higher variability still exhibit intra- and inter-day CVs below the 10% threshold. Despite these minor variations, the intra and inter-day repeatability of the optimized protocol are acceptable as long as these variabilities are considering when conducting a metabolomics study on salivary samples.

Among the tested protocols, including those previously described in the literature, our in-house method yielded the highest number of identified and quantified metabolites, most of which

also exhibited the highest signal intensities. Only two metabolites, ethanol and trimethylamine were not detected in this method compared to protocols without a freeze-drying step. This observation is important if a targeted metabolomics study focusing on one of these metabolites must be conducted. Despite this limitation, this protocol enabled the detection of two metabolites, 3-phenylpropionate and glycine, not observed in the methods without lyophilization. This improvement is largely attributable to the freeze-drying step, which concentrated the samples and enhanced their signal-to-noise ratio, allowing the detection of metabolites previously below the 1D NMR detection threshold. Additionally, the insertion of lyophilization helped mitigate the loss of aromatic amino acids typically associated with ultrafiltration. Moreover, by reversing the order of ultrafiltration and lyophilization compared to the protocol proposed by Quartier et al.²⁸, we were able to limit the loss of volatile metabolites during freeze-drying as observed in protocol 3.

For high-throughput applications, protocols involving freeze-drying are limited by the long lyophilization time (22h), making them less suitable for rapid and large-scale analyses. In this context, protocol 1 represents the most time-efficient option. However, good analytical repeatability is essential for clinical implementation, and this criterion was not fulfilled by protocol 1 in this study. Although further optimization could be considered, such as modifying CPMG parameters, using alternative pulse sequences, or applying computational approaches to reduce macromolecular contributions such as SMOIESY³⁸, variability related to residual macromolecular content after centrifugation step could persist. Consequently, protocol 1 may be suitable for high-throughput analyses targeting metabolites located in reproducible spectral regions, but alternative approaches should be considered when metabolites of interest lie in more variable regions. Protocol 2 may represent an intermediate compromise, offering improved repeatability. This approach, however, requires additional time (approximately 1.5h), incurs higher costs due to filtration

509 devices, and may lead to the loss of certain metabolites. Nevertheless, for high-throughput
510 applications requiring robust quantification of metabolites located in spectral regions affected by
511 macromolecular variability in protocol 1, protocol 2 may represent a more appropriate choice.
512 Finally, when rapid analysis is not a strict clinical requirement, protocol 4 offers the best overall
513 compromise, combining good repeatability, broad metabolome coverage and high spectral quality.
514 In contrast, protocol 3 was not retained, as it involves similar time and resource constraints as
515 protocol 4 while presenting additional disadvantages, as discussed above. In addition to the four
516 protocols addressed here, specific analytical protocols could also be developed for selected
517 metabolites in targeted metabolomic analyses, considering analysis time constraints.

518 In conclusion, saliva is a highly attractive biofluid for clinical use thanks to its non-invasive,
519 cost-effective and patient-friendly collection method. Moreover, for certain metabolites, salivary
520 levels may reflect changes observed in blood, making saliva a valuable and less invasive alternative
521 matrix for monitoring metabolic dynamics. Through biomarker discovery, saliva could be used as
522 a diagnostic and prognostic tool for disease apparition and progression. Moreover, it could play a
523 key role in personalized medicine by allowing treatment adjustments based on a patient's metabolic
524 profile. To date, numerous studies have highlighted the potential of salivary metabolomics and
525 reported promising results. However, for translation into clinical practice, it is essential to
526 standardize collection procedure, saliva type and sample preparation protocols. Indeed, to obtain
527 clinically reliable and reproducible results, monitoring of these pre-analytical steps is essential, as
528 well as ensuring a good repeatability and a good inter-day precision of the selected method. On
529 this basis, the present work represents one of the few studies that systematically compare multiple
530 sample preparation protocols for salivary NMR-based metabolomics. It focuses on selecting an
531 optimized method ensuring high-quality spectra, broad coverage of metabolome and good intra-

and inter-day precision. As expected, the set of metabolites detected was shown to depend on the protocol used, an important consideration for selecting the most appropriate method according to the needs of future metabolomics studies. In addition, the optimized protocol is fully compatible with MS-based approaches, representing an advantage for combined NMR-MS workflows and expanded metabolome coverage. Ultimately, establishing robust and standardized protocols for salivary metabolomics will be a key step toward integrating this accessible biofluid into clinical practice and contributing to the advancement of preventive and personalized medicine.

ASSOCIATED CONTENT

Data Availability Statement

The data underlying this study are openly available in Mendeley data at [https://data.mendeley.com/datasets/ \(DOI: 10.17632/z75ccv4cbz.1\)](https://data.mendeley.com/datasets/DOI:10.17632/z75ccv4cbz.1).

AUTHOR INFORMATION

Corresponding author

*Address: Quartier Hôpital, B36, Tower 4, +5, Avenue Hippocrate 15, 4000 Liège, Belgium. E-mail: mcampas@uliege.be. Phone: +32 4366 4381

Present Addresses

⁴Institut de Recherches Cliniques de Montréal and Department of Anatomy and Cell Biology, McGill University, H3A 0C7 Montreal, Canada

⁵Department of Chemistry and Biochemistry, Concordia University, H4B 1R6 Montreal, Canada

⁶Département de pharmacie hospitalière, CHU of Liege, 4000 Liege, Belgium

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

During this work, Constance Germay was a pharmacy master's student and her master's thesis contributed to part of the reported results.

SUPPORTING INFORMATION

Research of endogenous formate in an unspiked pooled saliva sample prepared according to protocol 1 (methods, results with Figure S1 and discussion); Figure S2: superimposed replicate ¹H-NMR spectra obtained with protocol 1 (DOC).

ACKNOWLEDGMENTS

We extend our appreciation to all participants for their valuable donation of saliva samples. Pascal de Tullio is Research director of the Fonds de la Recherche Scientifique-FNRS. **Table of Contents** created in BioRender. Cavalier, E. (2026) <https://BioRender.com/i2kf3ok>. **Figure 1** created in BioRender. Cavalier, E. (2026) <https://BioRender.com/8ejgq9m>.

ABBREVIATIONS

MS, mass spectrometry; LC, liquid chromatography; GC, gas chromatography; NMR, nuclear resonance magnetic; DSS, 4,4-dimethyl-4-silapentane-1-sulfonic acid; TSP, sodium trimethylsilyl-[2,2,3,3-2H₄]-propionate; CPMG, Carr-Purcell-Meiboom-Gill; NOESY, nuclear Overhauser effect spectroscopy; RD, relaxation delay; SD, standard deviation; COSY, correlation

573 spectroscopy; HSQC, heteronuclear single quantum coherence; HMDB, human metabolome
574 database; UV scaling, unit variance scaling; PCA, principal component analysis; ANOVA, analysis
575 of variance; CV, coefficient of variation

REFERENCES

- (1) Dame, Z. T.; Aziat, F.; Mandal, R.; Krishnamurthy, R.; Bouatra, S.; Borzouie, S.; Guo, A. C.; Sajed, T.; Deng, L.; Lin, H.; Liu, P.; Dong, E.; Wishart, D. S. The Human Saliva Metabolome. *Metabolomics* **2015**, *11* (6), 1864–1883. <https://doi.org/10.1007/s11306-015-0840-5>.
- (2) Takeda, I.; Stretch, C.; Barnaby, P.; Bhatnager, K.; Rankin, K.; Fu, H.; Weljie, A.; Jha, N.; Slupsky, C. Understanding the Human Salivary Metabolome. *NMR Biomed.* **2009**, *22* (6), 577–584. <https://doi.org/10.1002/nbm.1369>.
- (3) Nam, M.; Jo, S. R.; Park, J. H.; Kim, M.-S. Evaluation of Critical Factors in the Preparation of Saliva Sample from Healthy Subjects for Metabolomics. *J. Pharm. Biomed. Anal.* **2023**, *223*, 115145. <https://doi.org/10.1016/j.jpba.2022.115145>.
- (4) Gardner, A.; Parkes, H. G.; Carpenter, G. H.; So, P.-W. Developing and Standardizing a Protocol for Quantitative Proton Nuclear Magnetic Resonance (¹H NMR) Spectroscopy of Saliva. *J. Proteome Res.* **2018**, *17* (4), 1521–1531. <https://doi.org/10.1021/acs.jproteome.7b00847>.
- (5) Gardner, A.; Carpenter, G.; So, P.-W. Salivary Metabolomics: From Diagnostic Biomarker Discovery to Investigating Biological Function. *Metabolites* **2020**, *10* (2), 47. <https://doi.org/10.3390/metabo10020047>.
- (6) Javaid, M. A.; Ahmed, A. S.; Durand, R.; Tran, S. D. Saliva as a Diagnostic Tool for Oral and Systemic Diseases. *J. Oral Biol. Craniofacial Res.* **2016**, *6* (1), 67–76. <https://doi.org/10.1016/j.jobcr.2015.08.006>.
- (7) Soares Nunes, L. A.; Mussavira, S.; Sukumaran Bindhu, O. Clinical and Diagnostic Utility of Saliva as a Non-Invasive Diagnostic Fluid: A Systematic Review. *Biochem. Medica* **2015**, *25* (2), 177–192. <https://doi.org/10.11613/BM.2015.018>.
- (8) Hyvärinen, E.; Savolainen, M.; Mikkonen, J. J. W.; Kullaa, A. M. Salivary Metabolomics for Diagnosis and Monitoring Diseases: Challenges and Possibilities. *Metabolites* **2021**, *11* (9), 587. <https://doi.org/10.3390/metabo11090587>.
- (9) Zhang, A.; Sun, H.; Wang, X. Saliva Metabolomics Opens Door to Biomarker Discovery, Disease Diagnosis, and Treatment. *Appl. Biochem. Biotechnol.* **2012**, *168* (6), 1718–1727. <https://doi.org/10.1007/s12010-012-9891-5>.
- (10) Ishikawa, S.; Sugimoto, M.; Kitabatake, K.; Sugano, A.; Nakamura, M.; Kaneko, M.; Ota, S.; Hiwatari, K.; Enomoto, A.; Soga, T.; Tomita, M.; Iino, M. Identification of Salivary Metabolomic Biomarkers for Oral Cancer Screening. *Sci. Rep.* **2016**, *6* (1), 31520. <https://doi.org/10.1038/srep31520>.
- (11) Panneerselvam, K.; Ishikawa, S.; Krishnan, R.; Sugimoto, M. Salivary Metabolomics for Oral Cancer Detection: A Narrative Review. *Metabolites* **2022**, *12* (5), 436. <https://doi.org/10.3390/metabo12050436>.
- (12) Nijakowski, K.; Gruszczyński, D.; Kopala, D.; Surdacka, A. Salivary Metabolomics for Oral Squamous Cell Carcinoma Diagnosis: A Systematic Review. *Metabolites* **2022**, *12* (4), 294. <https://doi.org/10.3390/metabo12040294>.
- (13) Suragimath, G.; Patil, S.; Suragimath, D. G.; Sr, A. Salivaomics: A Revolutionary Non-Invasive Approach for Oral Cancer Detection. *Cureus*, **2024**, *16* (11): e74381. <https://doi.org/10.7759/cureus.74381>.

- (14) Vignoli, A.; Tenori, L. NMR-Based Metabolomics in Alzheimer's Disease Research: A Review. *Front. Mol. Biosci.* **2023**, *10*, 1308500. <https://doi.org/10.3389/fmolb.2023.1308500>.
- (15) Meng, L.; Jin, H.; Yulug, B.; Altay, O.; Li, X.; Hanoglu, L.; Cankaya, S.; Coskun, E.; Idil, E.; Nogaylar, R.; Ozsimsek, A.; Shoaie, S.; Turkez, H.; Nielsen, J.; Zhang, C.; Borén, J.; Uhlén, M.; Mardinoglu, A. Multi-Omics Analysis Reveals the Key Factors Involved in the Severity of the Alzheimer's Disease. *Alzheimers Res. Ther.* **2024**, *16* (1), 213. <https://doi.org/10.1186/s13195-024-01578-6>.
- (16) Piacenza Florezi, G.; Pereira Barone, F.; Izidoro, M. A.; Soares-Jr, J. M.; Coutinho-Camillo, C. M.; Lourenço, S. V. Targeted Saliva Metabolomics in Sjögren's Syndrome. *Clinics* **2024**, *79*, 100459. <https://doi.org/10.1016/j.clinsp.2024.100459>.
- (17) Setti, G.; Righi, V.; Mucci, A.; Panari, L.; Bernardelli, G.; Tarentini, E.; Gambini, A.; Consolo, U.; Generali, L.; Magnoni, C.; Meleti, M.; Sandri, G.; Bellini, P. Metabolic Profile of Whole Unstimulated Saliva in Patients with Sjögren's Syndrome. *Metabolites* **2023**, *13* (3), 348. <https://doi.org/10.3390/metabo13030348>.
- (18) Alt-Holland, A.; Huang, X.; Mendez, T.; Singh, M. L.; Papas, A. S.; Cimmino, J.; Bairos, T.; Tzavaras, E.; Foley, E.; Pagni, S. E.; Baleja, J. D. Identification of Salivary Metabolic Signatures Associated with Primary Sjögren's Disease. *Molecules* **2023**, *28* (15), 5891. <https://doi.org/10.3390/molecules28155891>.
- (19) Condor, A.; Kui, A. I.; Buduru, S. D.; Negucioiu, M.; Condor, D. C.; Lucaciu, P. Metabolomics Analysis as a Tool in Periodontitis Diagnosis: A Systematic Review. *Clin. Exp. Dent. Res.* **2025**, *11* (1), e70095. <https://doi.org/10.1002/cre2.70095>.
- (20) Garcia, P. N.; De Souza, M. M.; Izidoro, M. A.; Juliano, L.; Lourenço, S. V.; Camillo, C. M. C. Saliva Metabolomics: Concepts and Applications in Oral Disorders. *Clin. Oral Investig.* **2024**, *28* (11), 579. <https://doi.org/10.1007/s00784-024-05990-y>.
- (21) Gardner, A.; Parkes, H. G.; So, P.-W.; Carpenter, G. H. Determining Bacterial and Host Contributions to the Human Salivary Metabolome. *J. Oral Microbiol.* **2019**, *11* (1), 1617014. <https://doi.org/10.1080/20002297.2019.1617014>.
- (22) Figueira, J.; Gouveia-Figueira, S.; Öhman, C.; Lif Holgersson, P.; Nording, M. L.; Öhman, A. Metabolite Quantification by NMR and LC-MS/MS Reveals Differences between Unstimulated, Stimulated, and Pure Parotid Saliva. *J. Pharm. Biomed. Anal.* **2017**, *140*, 295–300. <https://doi.org/10.1016/j.jpba.2017.03.037>.
- (23) Frédérich, M.; Pirotte, B.; Fillet, M.; De Tullio, P. Metabolomics as a Challenging Approach for Medicinal Chemistry and Personalized Medicine. *J. Med. Chem.* **2016**, *59* (19), 8649–8666. <https://doi.org/10.1021/acs.jmedchem.5b01335>.
- (24) Kohler, I.; Hankemeier, T.; Van Der Graaf, P. H.; Knibbe, C. A. J.; Van Hasselt, J. G. C. Integrating Clinical Metabolomics-Based Biomarker Discovery and Clinical Pharmacology to Enable Precision Medicine. *Eur. J. Pharm. Sci.* **2017**, *109*, S15–S21. <https://doi.org/10.1016/j.ejps.2017.05.018>.
- (25) Letertre, M. P. M.; Giraudeau, P.; De Tullio, P. Nuclear Magnetic Resonance Spectroscopy in Clinical Metabolomics and Personalized Medicine: Current Challenges and Perspectives. *Front. Mol. Biosci.* **2021**, *8*, 698337. <https://doi.org/10.3389/fmolb.2021.698337>.
- (26) Santone, C.; Dinallo, V.; Paci, M.; D'Ottavio, S.; Barbato, G.; Bernardini, S. Saliva Metabolomics by NMR for the Evaluation of Sport Performance. *J. Pharm. Biomed. Anal.* **2014**, *88*, 441–446. <https://doi.org/10.1016/j.jpba.2013.09.021>.
- (27) Psychogios, N.; Hau, D. D.; Peng, J.; Guo, A. C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.; Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Xia, J.; Knox, C.; Dong, E.; Huang,

- P.; Hollander, Z.; Pedersen, T. L.; Smith, S. R.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J. W.; Goodfriend, T.; Wishart, D. S. The Human Serum Metabolome. *PLoS ONE* **2011**, *6* (2), e16957. <https://doi.org/10.1371/journal.pone.0016957>.
- (28) Quartieri, E.; Casali, E.; Ferrari, E.; Ghezzi, B.; Gallo, M.; Spisni, A.; Meleti, M.; Pertinhez, T. A. Sample Optimization for Saliva ¹H-NMR Metabolic Profiling. *Anal. Biochem.* **2022**, *640*, 114412. <https://doi.org/10.1016/j.ab.2021.114412>.
- (29) Martias, C.; Baroukh, N.; Mavel, S.; Blasco, H.; Lefèvre, A.; Roch, L.; Montigny, F.; Gatien, J.; Schibler, L.; Dufour-Rainfray, D.; Nadal-Desbarats, L.; Emond, P. Optimization of Sample Preparation for Metabolomics Exploration of Urine, Feces, Blood and Saliva in Humans Using Combined NMR and UHPLC-HRMS Platforms. *Molecules* **2021**, *26* (14), 4111. <https://doi.org/10.3390/molecules26144111>.
- (30) Fang, L.; Zhai, Q.; Zhang, H.; Ji, P.; Chen, C.; Zhang, H. Comparisons of Different Extraction Methods and Solvents for Saliva Samples. *Metabolomics* **2024**, *20* (2), 38. <https://doi.org/10.1007/s11306-024-02105-2>.
- (31) Barding, G. A.; Salditos, R.; Larive, C. K. Quantitative NMR for Bioanalysis and Metabolomics. *Anal. Bioanal. Chem.* **2012**, *404* (4), 1165–1179. <https://doi.org/10.1007/s00216-012-6188-z>.
- (32) Malz, F. Quantitative NMR in the Solution State NMR. In *NMR Spectroscopy in Pharmaceutical Analysis*; Elsevier, 2008; pp 43–62. <https://doi.org/10.1016/B978-0-444-53173-5.00002-0>.
- (33) Marchand, J.; Martineau, E.; Guitton, Y.; Le Bizec, B.; Dervilly-Pinel, G.; Giraudeau, P. A Multidimensional ¹H NMR Lipidomics Workflow to Address Chemical Food Safety Issues. *Metabolomics* **2018**, *14* (5), 60. <https://doi.org/10.1007/s11306-018-1360-x>.
- (34) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M.-A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, David. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; MacInnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. HMDB: The Human Metabolome Database. *Nucleic Acids Res.* **2007**, *35* (Database), D521–D526. <https://doi.org/10.1093/nar/gkl923>.
- (35) Le Guennec, A.; Tayyari, F.; Edison, A. S. Alternatives to Nuclear Overhauser Enhancement Spectroscopy Presat and Carr–Purcell–Meiboom–Gill Presat for NMR-Based Metabolomics. *Anal. Chem.* **2017**, *89* (17), 8582–8588. <https://doi.org/10.1021/acs.analchem.7b02354>.
- (36) Nicholson, J. K.; Gartland, K. P. R. ¹H NMR Studies on Protein Binding of Histidine, Tyrosine and Phenylalanine in Blood Plasma. *NMR Biomed.* **1989**, *2* (2), 77–82. <https://doi.org/10.1002/nbm.1940020207>.
- (37) *Human Metabolome Database: Showing metabocard for Putrescine (HMDB0001414)*. <https://www.hmdb.ca/metabolites/HMDB0001414> (accessed 2025-08-26).
- (38) Takis, P. G.; Jiménez, B.; Sands, C. J.; Chekmeneva, E.; Lewis, M. R. SMoLESY: An Efficient and Quantitative Alternative to on-Instrument Macromolecular ¹H-NMR Signal Suppression. *Chem. Sci.* **2020**, *11* (23), 6000–6011. <https://doi.org/10.1039/D0SC01421D>.

For Table of Contents Only

