

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,800

Open access books available

142,000

International authors and editors

180M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Metabolomics Distinction of Cigarette Smokers from Non-Smokers Using Non-Stationary Benchtop Nuclear Magnetic Resonance (NMR) Analysis of Human Saliva

Benita C. Percival, Angela Wann, Sophie Taylor, Mark Edgar, Miles Gibson and Martin Grootveld

Abstract

Implementations of high-field nuclear magnetic resonance (NMR) facilities into metabolomics studies are unfortunately restricted by their large dimensions, high costings, and specialist technical staff requirements. Therefore, here the application and practical advantages offered by low-field (60 MHz), compact NMR spectrometers for probing the metabolic profiles of human saliva was explored, as was their value in salivary metabolomics studies. Saliva samples were collected from cigarette smoking ($n = 11$) and non-smoking ($n = 31$) human participants. ^1H NMR spectra were acquired on both low-field (60 MHz) and medium-field (400 MHz) spectrometers. Metabolomics analyses were employed to evaluate the consistencies of salivary metabolite levels determined, and their abilities to distinguish between smokers and non-smokers. Low-field ^1H NMR analysis detected up to 15, albeit permitted the reliable quantification of 5, potentially key diagnostic biomolecules simultaneously (LLOQ values 250–400 $\mu\text{mol/L}$), although these were limited to those with the most prominent resonances. Such low-field profiles were also found to be suitable for salivary metabolomics investigations, which confirmed the successful discrimination between smoking and non-smoking participant sample donors. Differences observed between these groups were largely ascribable to upregulated salivary levels of methanol, and its metabolite formate, in the smoking group, but higher smoking-mediated concentrations of acetate, propionate and glycine may arise from a diminished salivary flow-rate in these participants. In conclusion, determination of salivary biomolecules using low-field, benchtop ^1H NMR analysis techniques were found to be valuable for bioanalytical and metabolomics investigations. Future perspectives for the applications of this non-stationary NMR technique, for example for the on-site ‘point-of-care’ testing of saliva samples for diagnostic oral disease screening purposes at dental surgeries and community pharmacies, are considered.

Keywords: compact low-field NMR analysis, NMR spectroscopy, NMR-linked metabolomics, bioanalytical chemistry, saliva, salivary biomarkers, oral diseases, tobacco smoking, salivary biomarkers for smoking, methanol

1. Introduction

'State-of-the-art' developments of novel devices and facilities for the metabolic screening of biofluids through 'omics' strategies provide an encouraging and thoroughly emerging outlook for future healthcare management prospects, including those focused on the diagnosis of diseases and/or their prognostic stratification [1]. Indeed, high-field (HF) nuclear magnetic resonance (NMR) facilities are routinely used for metabolomics investigations in order to rapidly recognize unusual metabolic patterns in patients suffering from a wide range of diseases. However, deployments in healthcare settings have been prohibitively impacted by the large sizes and costs of the instrumentation required for these purposes [2].

Imbalances in the human metabolome have long been associated with health-mediated disturbances, with ancient societies employing very crude methods to assess biofluids [3], including the smell and taste of urine samples to detect urinary ketone bodies and glucose, respectively, in cases of diabetes, for example. However, more recently there have been many notable developments in this research area, and these have given rise to the development of multicomponent metabolic profile analysis by many researchers since the early 1970s [4–7]. Whilst tandem liquid-chromatographic- and gas-chromatographic-mass spectrometric methods (LC-MS and GC-MS, respectively) can be employed for the reliable and sensitive determination of low concentrations of metabolites in biofluids, these approaches suffer from issues associated with matrix effects, and both are destructive techniques, which usually require a substantial knowledge of sample composition, and the likely identities of key biomarker analytes, prior to analysis [8]. Conversely, high-resolution NMR-based metabolomics analysis offers a highly selective means of simultaneously identifying a wide range of biomolecules present in complex biofluids at a minimal detectable concentration of $\leq 5 \mu\text{M}$, providing an untargeted methodology ideal for the analysis of a very large number of biomolecules simultaneously [9] (up to 120 or so and *ca.* 80 in human urine and saliva, respectively at operating frequencies of ≥ 600 MHz). Therefore, the multicomponent ^1H NMR analysis of biofluids such as blood plasma, urine and saliva, and tissue biopsies, offers a high level of potential regarding the investigation of metabolic processes, and when coupled with conventional and/or newly-developed multivariate (MV) data analysis techniques, serves as an extremely powerful means of probing the biochemical basis of human disease aetiology [1–5]. Indeed, this form of combined multianalyte-MV analysis is generally classified as metabolomics, and has been extensively applied in a very wide range of biomedical and clinical investigations, including the identification of diagnostic or prognostic biomarkers for a very wide range of diseases.

Although previous applications of low-field (LF) rather than HF NMR spectroscopy to the multicomponent analysis of intact or near-intact biofluids, or other biological media, have been severely limited, our laboratory has paved the way forward for the performance of such studies. Indeed, one of our recent investigations explored the ability of a LF (60 MHz) NMR spectrometer to provide valuable urinary metabolite data for the monitoring of type 2 diabetes (T2D) in humans [10]. Indeed, this application displayed a high level of chemopathological classification success, although this is perhaps not completely unexpected, since uncontrolled or poorly controlled T2D samples all contain quite high levels of glucose (both ^1H

NMR-distinguishable α - and β -anomers), along with the ketone bodies acetoacetate, acetone and 3-D-hydroxybutyrate, whereas little or none of these biomolecules are normally detectable in healthy control samples. However, one of the major advantages of the ^1H NMR technique is that it can simultaneously detect and monitor abnormal levels of a range of further metabolites involved in human disease pathology and associated co-morbidities, for example excessive urinary creatinine concentrations arising from kidney dysfunction and damage in T2D, together with hypoglycaemic drugs such as metformin [11].

Saliva serves as a multifunctional biofluid which plays important roles in facilitating the chewing, swallowing and tasting of foods, and also the regulation of oral flora; hence, it is of much importance for the maintenance of overall health in humans. Indeed, human saliva comprises an agglomerate hypertonic 'solution' which contains oral mucosal exudates, salivary acini and gingival crevicular fluid (GCF) [12]. Since this biofluid is readily accessible, it offers much potential as a medium for the identification and monitoring of established or potential biomarkers for human diseases, particularly oral health conditions, but not exclusively so. Indeed, its collection can be self-performed by participants with minimal training and without clinical supervision. Therefore, here we have employed optimized LF NMR-based metabolic profiling strategies for the global analysis of human saliva in order to assess the viabilities of these compact instruments for biomedical applications in metabolic profiling, metabolomics, oral health assessments, and potentially future screenings of the efficacies of oral healthcare products. Coupling of this LF salivary analysis technique to MV analysis strategies may indeed serve to facilitate the development of favourable outcome strategies for such investigations, and therefore here we also demonstrate, for the first time, the applications of LF NMR-based metabolomics protocols to the distinction of saliva samples collected from non-smoking and tobacco cigarette-smoking subjects. The future clinical monitoring applications of this novel technique are discussed.

2. Materials and methods

2.1 Saliva sample collection from human participants

Whole mouth saliva samples ($n = 61$) were collected from healthy human participants ($n = 42$, age range 21–65 years, 14 male/28 female), of whom 31 were non-smokers, and 11 were regular 'mild-to-heavy' smokers of tobacco cigarettes (an average of 3 to ≥ 20 cigarettes per day, with 44% of these smoking ≥ 20 per day). These non-smoking and tobacco-smoking participant groups were age-matched, with their mean \pm SEM ages being 42.52 ± 2.25 and 41.91 ± 2.70 years respectively. All ethical considerations were in accord with those of the Declaration of Helsinki 1975 (7th amendment made in 2013). All samples were collected with informed consent and approved by the Faculty of Health and Life Sciences Research Ethics Committee, De Montfort University, Leicester, UK (reference no. 1082). Participants were fasted for a 12-h period prior to providing saliva specimens. All participants were requested to refrain from all oral activities, including eating, drinking, tooth-brushing and smoking, etc. throughout this period, including the short, *ca.* 5 min. duration between awakening and sample donation. They were also requested not to consume any alcoholic beverages 24 h prior to the sample collection time-point. A range of 1–3 samples were collected from each participant, and those donating >1 sample provided these on separate daily a.m. 'wake-up' episodes. All samples were collected in sterile plastic universal containers and were transferred to the laboratory on ice. These were then immediately centrifuged at 3500 rpm at 4°C

for a period of 15 min, and following sample preparation as outlined below, the clear human salivary supernatants (HSSs) arising therefrom were then stored at -80°C for a maximal duration of 72 h until ready for NMR analysis.

2.2 Sample preparation and ^1H NMR analysis

All reagents and chemicals were purchased from Sigma-Aldrich (Gillingham, UK) unless otherwise stated. Aliquots (500 μL) of HSS samples were treated with 60 μL of pH 7.00 phosphate buffer (1.00 mol/L) containing 0.04% (w/v) sodium azide, and 50 μL of $^2\text{H}_2\text{O}$ containing 0.05% (w/v) sodium 3-(trimethylsilyl)propionate-2,2,3,3- d_4 (TSP) (final added HSS concentration 238 $\mu\text{mol/L}$). TSP served as an internal chemical shift reference and quantitative calibration standard; sodium azide acted as a microbicidal preservative in order to protect against the artefactual generation and/or consumption of microbial catabolites during the sample transport and preparation stages; phosphate buffer served to control sample pH values; and $^2\text{H}_2\text{O}$ acted as a field frequency lock. Admixtures were then homogenized and transferred to 5-mm diameter NMR tubes (Norell, Morganton, NC, USA). LF ^1H NMR spectra were acquired on a 60 MHz Magritek Spinsolve Ultra Benchtop spectrometer (Magritek GmbH, Philipsstr. 852068, Aachen, Nordrhein-Westfalen, Germany) with 64 and/or 384 scans, acquisition and repetition times of 6.4, and 10 or 15 s respectively, and a pulse angle of 90° ; the $\text{H}_2\text{O}/\text{HOD}$ presaturation frequency was optimized at $\delta = 4.80$ ppm using the programmed 1D PRESAT function. For the calibration and metabolomics studies described here, scan number and repetition time were standardized at 64 and 10 s respectively. Notwithstanding, total analysis times of <15 min per sample were possible. These samples also underwent medium-field (MF) ^1H NMR analysis at an operating frequency of 400 MHz (Bruker Avance-I 400 spectrometer (Bruker AXS, GmbH, Östliche Rheinbrückenstr. 49 76187, Karlsruhe, Germany, Leicester School of Pharmacy facility, De Montfort University, Leicester, UK), operating at a frequency of 400.13 MHz, and using the noesygppr1d pulse sequence for water suppression (H_2O , $\delta = 4.80$ ppm); 32 k data points were acquired in 128 scans, with 2 dummy scans, a sweep width of 4844 Hz, and an automatically-adjusted receiver gain.

^1H NMR resonances present in each HSS spectrum acquired were routinely assigned by a consideration of chemical shift values, coupling patterns and coupling constants with reference to literature sources, and where required, two-dimensional ^1H - ^1H correlation and total correlation (COSY and TOCSY respectively) spectra were acquired to confirm these assignments. Median ^1H NMR signal-to-noise (STN) ratios were determined from the formula $\text{STN} = 2.50A/\text{Npp}$, where A represents resonance height, and Npp the highest peak-to-peak noise difference determined at each chemical shift region selected. Lower limits of detection and quantification (LLOD and LLOQ respectively) values were computed as 3- and 10-times these median STN values. HSS spectral resonances were manually-bucketed, and their intensities determined using *ACD/Spectrum Processor 2019* software; that of residual $\text{H}_2\text{O}/\text{HOD}$ was removed prior to performing univariate (UV) or MV statistical analysis. Salivary biomolecule levels were determined from calibration plots of ratios of their pre-selected resonance intensities to that of internal TSP against their known concentrations in a series of analytical calibration standard solutions.

Calibration and Bland-Altman dominance plots of the ^1H NMR-determined concentrations of acetate, propionate, formate, glycine and methanol featured matched analysis sample datasets, with determinations made on these salivary metabolites at both 60 and 400 MHz operating frequencies. All determinations which were found to have none detectable (nd, specifically those with values $<$ LLOD) at both operating frequencies utilized, were removed from the datasets. As

recommended [13], corresponding ^1H NMR profiles of blank samples, which were prepared as outlined above, but with HPLC-grade water in place of HSSs, were acquired, and their 'noise' intensities at the appropriate δ values were included in these calibration plots. Spectra were acquired on replicate ($n = 3$) preparations of such blank samples for these purposes.

2.3 UV and MV statistical and metabolomics analyses

2.3.1 UV analysis

A paired sample t-test was applied to test for any differences between LF 60 MHz spectra acquired with relaxation delays of 10 or 15 s, and an *XLSTAT2020* software module was employed for this purpose.

An analysis-of-covariance (ANCOVA) experimental design was employed to test the statistical significance of the 'between-smoking status' source of variation, along with the potential effects of essential demographic variables recorded on participant sample donors, on salivary acetate, formate, propionate, glycine and methanol concentrations (Eq. (1)). Overall, this model evaluated the influences of the 'between-participant' (random) effect $P_{(k)l}$, and the 'between-participant ages' (A_i), 'between-participant-genders' (G_j), and 'between-smoking status' (S_k) sources of variation (all fixed) on these five sets of ^1H NMR-determined levels. Moreover, the statistical significance of the age \times gender, age \times smoking status, and gender \times smoking status first-order interaction effects (AG_{ij} , AS_{ik} and GS_{jk} respectively) were also assessed. ANCOVA was performed using *XLSTAT2014* and *2020* software modules.

$$Y_{ijklm} = \mu + A_i + G_j + S_k + P_{(k)l} + AG_{ij} + AS_{ik} + GS_{jk} + e_{ijklm} \quad (1)$$

An additional ANCOVA model explored the significance of any differences between the two operating frequencies in the non-smoking group only, and in this experimental design, the 'between-ages', 'between-genders' and 'between-participants' sources of variation were also evaluated, as was the age \times gender first-order interaction effect (Eq. (2)). In this design, O_k represents the 'between-spectrometer operating frequencies' effect (fixed).

$$Y_{ijklm} = \mu + A_i + G_j + O_k + P_{(k)l} + AG_{ij} + e_{ijklm} \quad (2)$$

2.3.2 MV metabolomics analysis

Principal component analysis (PCA) was primarily employed to identify any possible outlier samples present in the 60 MHz operating frequency ^1H NMR dataset, and in total 8 of these were found and classified as such, and then subsequently removed prior to the performance of further MV analysis. PCA was then employed to determine the reproducibility of replicate salivary metabolite determinations made on the LF benchtop spectrometer, and this check was performed with $n = 9$ duplicated salivary samples randomly selected from the smoking group of participants, with all the above five metabolite variables included, and not non-smokers in view of the restricted availability of data on salivary methanol and, to a lesser extent, formate concentrations above their LLOQ indices. For this PCA analysis, salivary metabolite concentrations were not constant sum-normalized (CSN), nor transformed, and nor auto- or Pareto-scaled.

For the major objective of this study, a LF benchtop ^1H NMR-based metabolomics investigation featured a comparison of saliva specimens collected from the non-

smoking and tobacco-smoking participants, and for this purpose all the above 5 potential predictor variables, determined via TSP-normalization as described above, were incorporated. For these purposes, the dataset was product quotient normalized (PQN), generalised \log_{10} (glog)-transformed, and Pareto-scaled prior to MV analysis, which involved PCA, partial least squares-discriminatory analysis (PLS-DA), orthogonal partial least squares-discriminatory analysis (OPLS-DA), random forest (RF) and agglomerative hierarchical clustering (AHC) techniques (*MetaboAnalyst 5.0*, University of Alberta and National Research Council, National Institute for Nanotechnology, (NINT), Edmonton, AB, Canada). Distinctions found between the two groups with the above PLS-DA and OPLS-DA strategies were cross-validated with determination of Q^2 statistics, and also permutation tests with 2000 permutations. A Q^2 value of ≥ 0.50 was considered as a significant discriminatory cut-off threshold [14].

PQN converts ^1H NMR metabolomics profiles according to an overall estimate of the most probable ‘dilution’ influence [15], and for saliva this includes reductions in salivary flow-rate (SFR), which has been reported to be significantly reduced in cigarette smokers [16]. This strategy usually involves the subtraction of a mean or median column-bucketed reference spectral profile from those of either all or a subset of study samples, and for this investigation the non-smoking control group was employed for this purpose.

At an operating frequency of 60 MHz, some missing data in the MV datasets were unavoidable in view of resonance overlap complications. Therefore, for this metabolomics model, these randomly missing values, *ca.* 9% of the total available ^1H NMR-determined concentrations, were estimated and imputed using the non-linear iterative partial least squares (NIPALS) approach (*XLSTAT2020*) [17], since this method is considered appropriate for MV metabolomics datasets such as that analyzed in the current study. For UV data analysis, these imputations were accompanied by a corresponding decrease in degrees of freedom available for the parametric statistical evaluations conducted. Metabolite concentration values below the detection limit (i.e., zero analyte values or ‘less-thans’) were replaced using the simple multiplicative replacement approach described in Ref. [18], specifically as 65% of their metabolite LLOD values.

Multivariate ANOVA (MANOVA) was performed using *XLSTAT2014* software. The non-parametric RF analysis featured 1000 trees and 2 predictor variables selected at each node following tuning. The dataset was randomly split into training and test sets comprising *ca.* two-thirds and one-third of the samples respectively. The training set was used to construct the RFs model and determine an out-of-the-bag (OOB) error value in order to assess classification performance.

3. Results

3.1 Outline of ^1H NMR analysis results: detection and quantitative determination of salivary biomolecules at 60 MHz operating frequency

Figure 1 shows the 60 MHz ^1H NMR profile of a typical human salivary supernatant (HSS) sample obtained with 384 scans, a process which took 60 min to acquire. This spectrum contains clear ^1H resonances assignable to acetate (signal 7) and methanol (signal 14), which are both ascribable to their $-\text{CH}_3$ groups. Further prominent resonances in the spectra acquired were those of propionate (both $-\text{CH}_3$ and $-\text{CH}_2$ functions, signals 1 and 8 respectively) and glycine ($\alpha\text{-CH}_2$ protons, signal 15), together with that of the single ^1H NMR-detectable proton (H-CO_2^-) of formate (signal 20). Further, albeit weak signals were those assignable

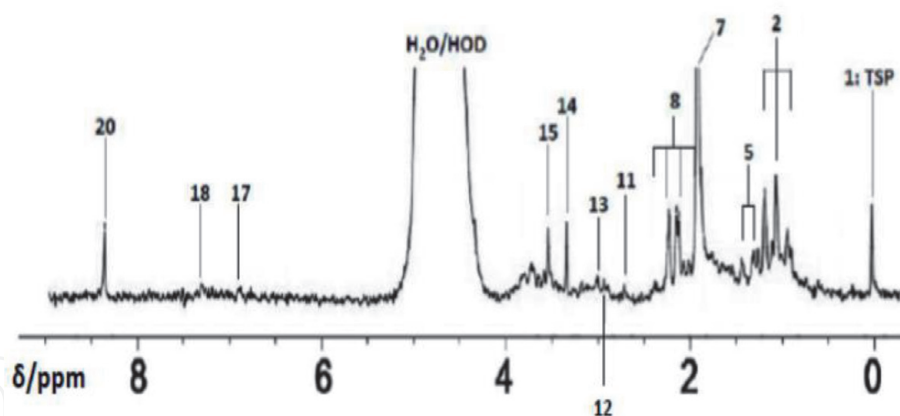


Figure 1. Experimental 60 MHz ^1H NMR profile of a HSS sample acquired with 384 scans. Numerical assignment codes correspond to those in **Table 1**. TSP represents the 3-(trimethylsilyl)propionate-2,2,3,3- d_4 chemical shift reference and internal quantitative ^1H NMR standard, and $\text{H}_2\text{O}/\text{HOD}$ the residual water signal.

to dimethyl- and trimethylamine, the terminal- CH_2 groups of amine species such as lysine and 5-aminovalerate, and the $>\text{N}-\text{CH}_3$ group of creatinine/creatine, together with two aromatic resonances (one a composite phenylalanine/tyrosine one), although all these biomolecules predominantly had salivary concentrations below their LLOQ values. A full list of all resonances identified in the 60 MHz ^1H (superscript 1) NMR profiles of HSS samples analysed is provided in **Table 1**.

Acceptable benchtop 60 MHz spectra were also obtained with only 64 scans, which involved a 10 min acquisition time, although all ^1H NMR resonances therein were, however, notably affected by significantly lower signal-to-noise (STN) parameters, as expected. All carboxylic acid anions detectable represent oral microbial catabolites, and the simultaneous ^1H NMR measurement of their salivary concentrations in this manner may offer significant potential regarding the provision of valuable diagnostic and prognostic screening information for dental surgeons and oral healthcare specialists alike, especially those regarding conditions such as dental caries and periodontal diseases, as previously noted in studies performed with conventional specialist NMR laboratory-based medium- and high-resolution 400 and 600 MHz operating frequency spectrometers respectively [19, 20]. Moreover, amino acids detectable such as glycine are potentially derived from the actions of proteolytic bacteria on salivary proteins, although there are, of course, also host sources of this metabolite [21], as indeed there are for some of the organic acid anions, such as acetate.

Although the 60 MHz spectral profiles of HSS specimens investigated are largely dominated by the highest intensity resonances therein (i.e. those assigned to biomolecules of the highest salivary concentrations such as acetate and propionate), this technique also lent itself to the assignment of lower intensity signals, and the quantification of salivary metabolites present at significantly lower concentrations. Indeed, the singlet resonances of formate, methanol and glycine were clearly resolved from potential overlapping signals, and therefore appeared suitable for quantification purposes. As previously documented [23], one major source of salivary methanol in humans is the ingestion of cigarette smoke. However, in some HSS specimens, the lactate- CH_3 function doublet resonance at $\delta = 1.33$ ppm was also clearly visible in 60 MHz spectra, most especially those with quite high millimolar salivary levels. Although highly variable, reported mean levels for salivary lactate in adults are 0.1–20.3 mmol/L [24]. However, our previously reported overall mean lactate concentration in saliva samples collected from a pre-fasting healthy control population was found to be 13.3 mmol/L, although those of replicate samples from $n = 20$ participants also varied substantially, i.e. from 0.08 to 100.9 mmol/L [18].

Resonance code	Chemical shift (multiplicity)	Assignment
1	0.00 (<i>s</i>)	3-Trimethylsilyl-(2,2,3,3- ² H ₄)-1-propionate (TSP)-Si(CH ₃) ₃ : internal chemical shift reference and quantitative integration intensity calibration standard
2	1.06 (<i>t</i>)	Propionate-CH ₃
3	1.13 (<i>d</i>)	Propane-1,2-diol-CH ₃ *
4	1.18 (<i>t</i>)	Ethanol-CH ₃ *
5	1.33 (<i>d</i>)	Lactate-CH ₃
6	1.48 (<i>d</i>)	Alanine-CH ₃
7	1.92 (<i>s</i>)	Acetate-CH ₃
8	2.18 (<i>q</i>)	Propionate-CH ₂
9	2.23 (<i>t</i>)	5-AV-α-CH ₂
10	2.41 (<i>s</i>)	Succinate-CH ₂
11	2.75 (<i>s</i>)	Dimethylamine-N(CH ₃) ₂
12	2.90 (<i>s</i>)	Trimethylamine-N(CH ₃) ₃
13	3.04 (<i>m</i>)	Creatinine-N(CH ₃)/Creatine-N(CH ₃)/Lysine-ε-CH ₂ /5-AV-δ-CH ₂
14	3.37 (<i>s</i>)	Methanol-CH ₃ *
15	3.57 (<i>s</i>)	Glycine-α-CH ₂
16	3.67 (<i>m</i>)	Ethanol-CH ₂ *
17	6.91 (<i>d</i>)	Tyrosine aromatic ring protons-CH
18	7.22 (<i>d</i>)/7.30–7.45 (<i>m</i>)	Tyrosine/phenylalanine aromatic ring protons-CH
19	7.91 (<i>s</i>)	Histidine imidazole ring-CH
20	8.45 (<i>s</i>)	Formate-CH

Abbreviation: 5-AV, 5-aminovaleate.

*Indicates resonances of molecules which may also arise from exogenous sources, e.g., propane-1,2-diol and methanol from tobacco smoking [23], and ethanol from alcoholic beverage ingestion [19].

Table 1.

Assignments for resonances present in the 60 MHz ¹H NMR spectra shown in **Figure 1** (coupling patterns for these are also provided). Resonances highlighted in red are visible in both LF (60 MHz) and HF (400 MHz) ¹H NMR profiles, whereas those in blue are observable but not readily quantifiable at the lower operating frequency in view of overlap with or close localization to other biomolecule signals, or being below the lower limits of quantification (LLOQ) values for their assigned metabolites. The identities of selected signals were confirmed via reference to [22].

TSP was present in HSS analyte solutions at an added level of only 238 μmol/L (i.e. a 238/9 = 26.44 μmol/L single ¹H nucleus equivalent value), and because this was also one of the most intense signals present in the spectra acquired (*s*, δ = 0.00 ppm), concentrations of <60 μmol/L were also readily visible in spectra of chemical model systems containing this internal standard. However, since this intense resonance arises from a total of 9 protons (i.e. 3 equivalent Si-CH₃ units), it is conceivable that many salivary and perhaps other biofluid metabolites containing only single, or one or more magnetically-distinguishable —CH₃ functions with singlet resonances (for example, acetate) are clearly detectable and potentially also quantifiable at concentrations of ≤150 μmol/L in this complex, multicomponent biofluid matrix, although further investigations are required to explore this. Moreover, for such —CH₃ function-containing analytes, a LLOD value of *ca.* 100 μmol/L

was estimated for HSS samples (i.e. $3 \times$ the mean spectral noise intensities at their specified chemical shift values in water blank solution spectra acquired under the same experimental conditions, or a SNR value of 3). Such LLOD values will also be influenced by further factors, such as T_2 -dependent resonance line-widths (the potential influence of which will be much greater at LFs), the dependence of noise on chemical shift values, saturation effects, digital resolution and baseline slants, etc. Notwithstanding, further key experiments are required to determine these LLOQ values (with SNR = 10) for key biomolecules present within LF salivary spectra, particularly oral disease-linked biomarkers of interest.

The relaxation delay employed for acquisition of LF 60 MHz ^1H NMR spectra, i.e. 10 or 15 s, did not give rise to any differences in the estimated salivary concentrations of acetate, propionate, formate, glycine and methanol (paired sample t-tests performed on untransformed datasets, $n \geq 10$ matched spectra for each biomolecule). These data demonstrated that a relaxation delay of 10 s was sufficient for full T_1 relaxation of the ^1H environments involved.

3.2 Comparative evaluations of the ^1H NMR profiles of human salivary supernatants at 60 and 400 MHz operating frequencies

All resonances detectable in the above 60 MHz ^1H NMR profiles were, of course, also readily detectable in corresponding 400 MHz spectra, and as expected, median resonance STN values were much greater, and hence corresponding LLOQ and LLOD parameters were substantially lower with this MF spectrometer. Indeed, signals arising from metabolites with detectable but not quantifiable signals in the 60 MHz profiles acquired, for example the malodorous amines DMA and TMA, and the amino acids phenylalanine and tyrosine, were readily quantifiable at the 400 MHz operating frequency. Additionally, in view of the much improved spectral resolution and quality, metabolites which were only detectable but unquantifiable, or completely undetectable at 60 MHz, were also detectable and predominantly quantifiable at MF, and these included leucine ($-\text{CH}_3$ (*t*), $\delta = 0.96$ ppm); valine ($-\text{CH}_3$ s (both *d*), $\delta = 0.98$ and 1.03 ppm); alanine ($-\text{CH}_3$ (*d*), $\delta = 1.48$ ppm); glutamate ($\gamma\text{-CH}_2$ (*m*), $\delta = 2.34$ ppm); glutamine ($\gamma\text{-CH}_2$ (*m*), $\delta = 2.44$ ppm); taurine ($-\text{CH}_2\text{NH}_3^+$ and $-\text{CH}_2\text{SO}_3^-$ (both *t*), $\delta = 3.23$ and 3.47 ppm respectively); *n*-butyrate ($-\text{CH}_3$, $-\text{CH}_2-$ and CH_2CO_2^- (*t*, *m* and *t*), $\delta = 0.94$, 1.55 and 2.15 ppm respectively); 2,2-dimethylsuccinate ($-\text{CH}_3$ s (*s*), $\delta = 1.22$); 3-D-hydroxybutyrate ($-\text{CH}_3$ (*d*), $\delta = 1.24$ ppm); lactate ($-\text{CH}_3$ and $-\text{CH}$ (*d* and *q*), $\delta = 1.33$ and 4.13 ppm respectively); 5-aminovalerate (β/γ -, α - and $\delta\text{-CH}_2$ s (*m*, *t* and *t*), $\delta = 1.63$, 2.23 and 3.05 ppm respectively); pyruvate ($-\text{CH}_3$ (*s*), $\delta = 2.39$ ppm); succinate ($-\text{CH}_2$ s (*s*), $\delta = 2.405$ ppm); choline ($-\text{N}(\text{CH}_3)_3^+$ head group (*s*), $\delta = 3.21$ ppm); ethanol ($-\text{CH}_3$ and $-\text{CH}_2\text{OH}$ (*t* and *q*), $\delta = 1.18$ and 3.66 ppm respectively); carbohydrates such as glucose and sucrose (C1H anomeric protons located at 4.66/5.25 for the former (both *ds*), and 5.41 ppm (*d*) for the glycosidic proton of the latter, respectively), where detectable; dihydroxyacetone ($-\text{CH}_2\text{OH}$ (*s*), $\delta = 4.46$ ppm), and N-acetylsugars and N-acetylamino acids, both high- and low-molecular mass (broad and narrow $-\text{NHCOCH}_3$ signals respectively (*s*), $\delta = 2.01$ –2.08 ppm); aromatic amino acids (aromatic ring resonances of tyrosine ($2 \times d$, $\delta = 6.88$ and 7.25 ppm), phenylalanine ($3 \times m$, $\delta = 7.32$, 7.38 and 7.43 ppm) and histidine ($2 \times s$, $\delta = 7.07$ and 7.81 ppm); and those assignable to a number of pyrimidine, or nicotinate and nicotinamide pathway metabolite(s).

Plots of salivary acetate, glycine and methanol concentrations determined on the LF 60 MHz NMR facility *versus* those obtained on the HF 400 MHz instrument were all linear ($r = 0.990$, 0.987 and 0.995 respectively), and these results confirmed good-to-excellent correlations and agreements between these two methods of ^1H

NMR analysis for these biomolecules. However, this correlation was found to be less strong for formate ($r = 0.927$). Moreover, for propionate, despite a strong linear correlation between these two forms of NMR analyses ($r = 0.973$), 95% confidence intervals (CIs) for its regression coefficient were found to be significantly greater than the 1.00 value expected for good agreement between these values (i.e. 1.22–1.41). This observation is explicable by potential interferences arising from further ^1H NMR resonances located within the $\delta = 0.92$ – 1.18 ppm chemical shift range spanned by the propionate- CH_3 signal at 60 MHz (15.3 Hz in total: $J = 7.67$ Hz for this signal). These potential interfering signals are those assignable to the terminal- CH_3 functions of both long- and alternative short-chain fatty acids (including that of *n*-butyrate at $\delta = 0.94$ ppm and those of branched-chain amino acids such as valine). Therefore, the apparent propionate concentration determined at 60 MHz was significantly inflated by a mean value of approximately 1.32-fold over those determined at the more conventional 400 MHz operating frequency.

In order to explore the significant LF spectral overestimations of propionate further, the complete chemical shift span of its $-\text{CH}_3$ function triplet at 60 MHz (0.92–1.18 ppm) was integrated in the corresponding 400 MHz spectra acquired, and apparent propionate concentrations obtained in this manner were then plotted against those determined at 60 MHz (**Figure 2**). As expected, this plot was found to display a much-improved agreement between the two estimated concentration values, with 95% CI values for the regression coefficient and y -intercept both covering unity and 0.00 respectively ($r = 0.985$). However, a comparison of these bioanalytical calibration plots indicated only a relatively marginal interference from the above potentially overlapping, albeit minor signals for the direct determinations of propionate at 60 MHz when its salivary level was *ca.* ≤ 1.0 mmol/L. Therefore, for the LF NMR determination of salivary propionate, one possible solution is that

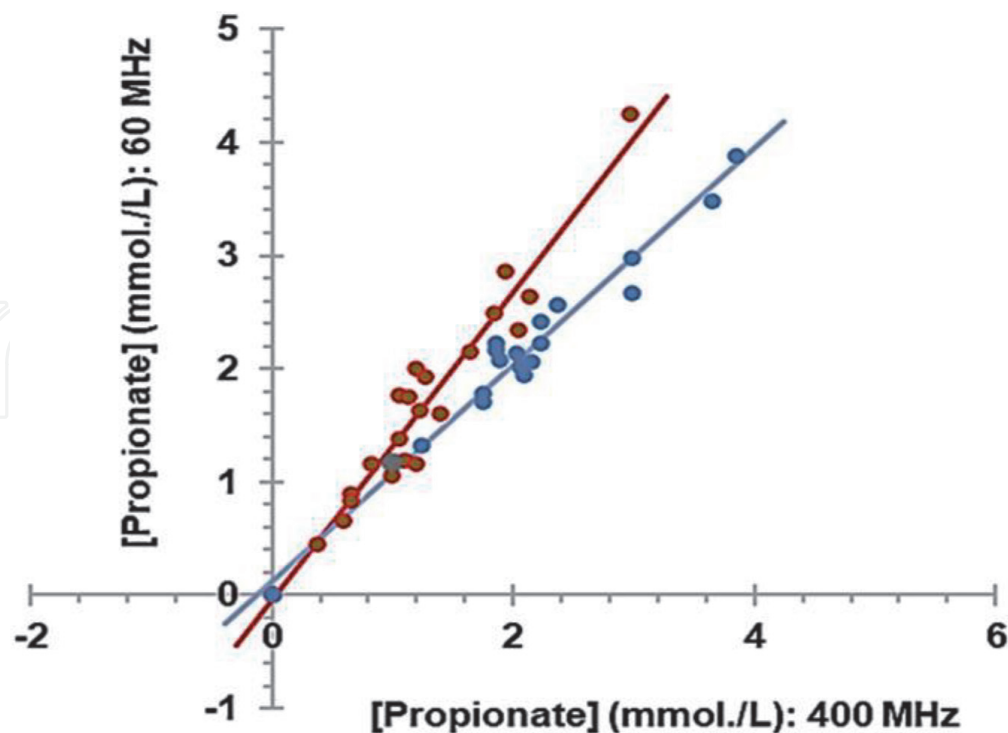


Figure 2.

Comparison of a plot of the estimated salivary concentrations of propionate determined at 60 MHz with that from spectra acquired at 400 MHz for its $-\text{CH}_3$ function resonance encompassing its chemical shift span i.e. $2 \times 7.67 = 15.34$ Hz at both operating frequencies, and equivalent to 0.256 and 0.038 ppm at 60 and 400 MHz respectively (red data-points and regression line), with that obtained from employment of the corresponding 60 MHz 0.256 ppm bucket span for integration purposes at 400 MHz, i.e. 0.92–1.18 ppm (blue data-points and regression line).

samples with concentrations greater than this value are first diluted to levels close to or below this limit in order to facilitate its quantification.

Nevertheless, a novel statistical approach was employed for determining the maximal concentration determinable at this operating frequency. Firstly, an alternative ANCOVA model (model 2) was designed and employed for statistical analysis of the non-smoking participant group only (Eq. (2)), and this included a consideration of variance contributions from differential ages, genders, participants and spectrometer operating frequencies, plus the age \times gender first-order interaction effect. Secondly, p values for the statistical significance of the fixed 'between-operating frequencies' (O_k) effects were isolated for a series of these ANCOVA models arising from sequential removal of the highest or highest remaining estimated salivary propionate concentration, i.e. $[\text{propionate}]_{\text{max}}$ (at 60 MHz), starting from a prior sample size of $n = 30$ observations (i.e. with coupled 60 and 400 MHz ^1H NMR determinations for each of $n = 15$ participants), down to only $n = 6$ (with coupled measurements made for only $n = 3$ participants); as expected, these p values increased with decreasing sample size, i.e. the degree of statistical significance between the two operating frequencies tested decreased. Thirdly, $-\log_{10} p$ transformations of these p values were then plotted as a function of $[\text{propionate}]_{\text{max}}$ value (Figure 3), and the ordinate axis value of this curve set at $p = 0.05$ served to

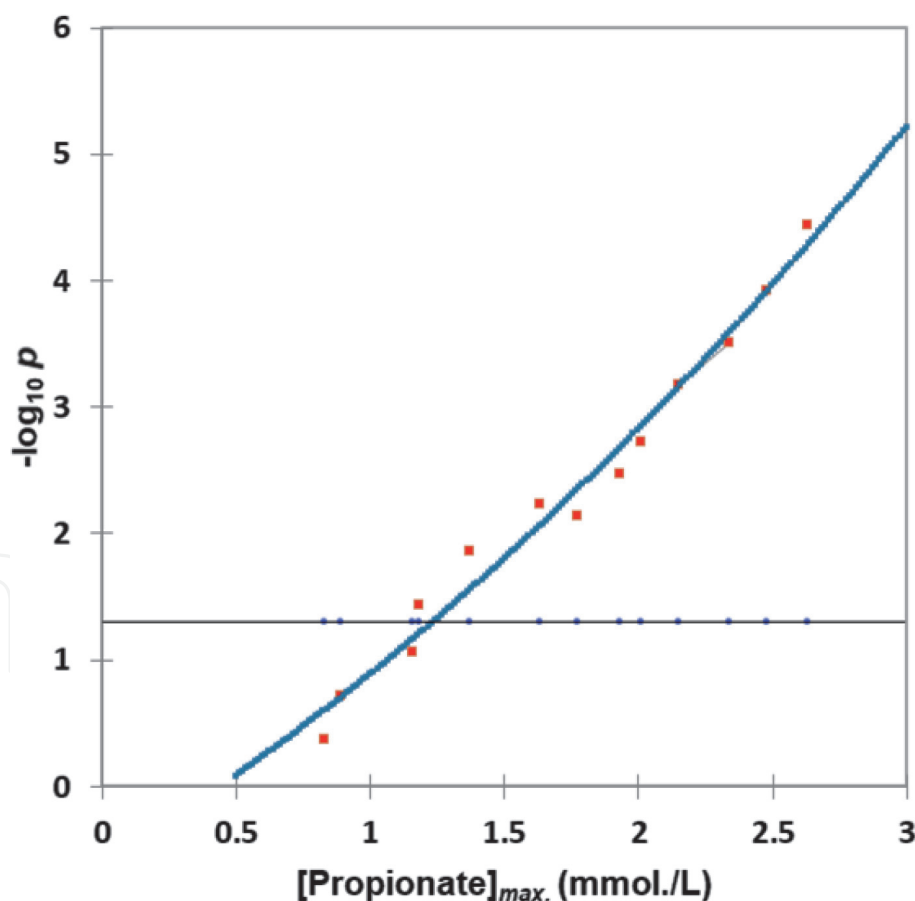


Figure 3.

(a) Polynomial plot of $-\log_{10} p$ value obtained for the significance of the 'between-operating frequencies' mean square of the ANCOVA model of Eq. (2) as a function of decreasing maximal salivary propionate concentration ($[\text{propionate}]_{\text{max}}$) with sample size, the latter diminishing via the sequential removal of the $[\text{propionate}]_{\text{max}}$ value (from $n = 15$ to $n = 3$ matched duplicate samples, one determination made at 60 MHz, one at 400 MHz ^1H NMR operating frequencies). The horizontal black line represents the $-\log_{10} p$ index arising from a p value of 0.05, i.e. the minimal level required for statistical significance; its crossing with blue polynomial plot yields a $[\text{propionate}]_{\text{max}}$ value of 1.2 mmol/L. The quadratic equation fitted to the experimental data was $-\log_{10} p = 0.603 + 1.261[\text{Propionate}]_{\text{max}} + 0.226[\text{Propionate}]_{\text{max}}^2$ ($R^2 = 0.9795$), which was an improved fit over that obtained with a standard linear relationship.

provide an estimate for the latter value's limit for bioanalytical purposes, i.e. that at which the difference observed between salivary concentrations determined at 60 and 400 MHz operating frequencies became statistically insignificant at the 5% level. This limit was therefore estimated as 1.2 mmol/L for salivary propionate concentrations determined at 60 MHz, a value similar to the 1.00 mmol/L limit proposed above.

3.3 PCA and MANOVA assessments of the bioanalytical precision of metabolite determinations at 60 MHz operating frequency

Subsequently, principal component analysis (PCA) was utilized to monitor the precision of duplicate, 'between-assay' sample analyses with this facility in a model containing 5 LF NMR-detectable salivary metabolites with quantifiable concentrations (acetate, propionate, formate, glycine and methanol). For this purpose, $n = 9$ duplicate sets of samples were randomly drawn from the tobacco-smoking group, and analyzed in different assay batches conducted with LF 60 MHz spectral acquisitions made on separate work-days. With the exception of two sets of matched analyses, there was a good agreement of both PC1 and PC2 scores obtained for all duplicate samples analyzed (**Figure 4**). Moreover, MV analysis-of-variance (MANOVA) of these experimental data found that the 'between-replicates' and the participant \times replicate interaction effects were both statistically insignificant ($p = 0.80$ and 0.97 respectively), although there was a very highly significant difference observed 'between-participants' ($p = 5.50 \times 10^{-4}$), as expected.

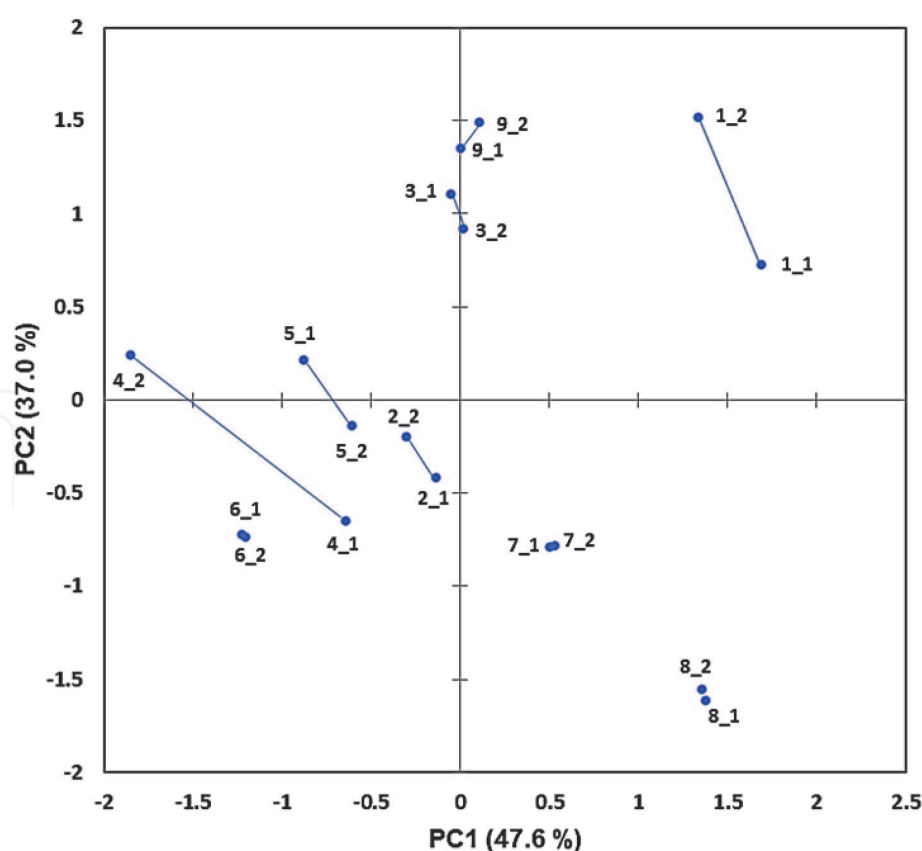


Figure 4.

PCA scores plot of PC2 versus PC1 for duplicate LF NMR determinations of 5 metabolites in saliva specimens collected from $n = 9$ tobacco smoking participants at an operating frequency of 60 MHz. This plot demonstrates that, with the exception of samples 1 and 4, there was a satisfactory agreement between the repeated determinations. Abbreviations: n_1 and n_2 refer to the first and second duplicate determinations made on sample n , and so on.

3.4 UV evaluations of metabolic differences between smoking and non-smoking participants

Mean \pm SEM salivary levels of acetate, propionate, formate, glycine and methanol determined using the LF 60 MHz benchtop spectrometer for both the non-smoking and smoking groups are provided in **Table 2**. Univariate statistical analyses of these data revealed that the salivary concentrations of methanol, glycine and acetate were all significantly higher in smokers, the substantial upregulation in methanol observed being concordant with results previously acquired by Percival et al. [23].

The above model 1 experimental design featured the ‘between-age (A_i), -gender (G_j) and -smoking status (S_k)’ main factors (fixed effects), the ‘between-participants’ random effect ($P_{(k)l}$), and the AG_{ij} , AS_{ik} and GS_{jk} first-order interactions effects. Results from this factorial analysis are shown in **Table 3**. Overall, the participant age factor was not significant for any metabolite; the gender predictor

Metabolite	Non-smokers (mmol/L)	Smokers (mmol/L)
Acetate	5.54 \pm 0.95 (range 0–30.92)	11.20 \pm 2.00 (range 0–30.20)*
Propionate [†]	1.89 \pm 0.18 (range 0–5.59)	2.43 \pm 0.35 (range 0–4.95)
Formate	0.14 \pm 0.04 (range 0–0.95)	0.37 \pm 0.09 (range 0–1.27)
Glycine	0.34 \pm 0.04 (range 0–1.04)	0.80 \pm 0.17 (0–2.49)*
Methanol ^{††}	0.00 \pm 0.00 (range 0–0.20)	0.48 \pm 0.11 (0–1.58)**

[†]Propionate levels were determined at an operating frequency of 400 MHz in view of resonance overlap complications observed at 60 MHz.

^{††}At an operating frequency of 60 MHz, this agent was ¹H NMR-quantifiable in $n = 3$ samples only in the non-smoking participant cohort.

* $p < 0.05$ for differences between the mean values of non-smokers and smokers; ** $p < 10^{-3}$ (test performed for determining the significance of the ‘between-smoking status’ mean square in the model 1 ANCOVA model design delineated in (Eq. (1))).

Table 2.

Mean \pm SEM concentrations of salivary biomolecules determined by LF 60 MHz ¹H NMR analysis of HSS samples in non-smoking and smoking sampling groups (sample sizes were $n = 33$ and 19 for these groups respectively; ranges are provided in brackets).

Effect	Status	Metabolite				
		Acetate	Propionate	Formate	Glycine	Methanol
Age	Fixed	ns	0.114	ns	ns	ns
Smoking Status	Fixed	0.044	0.127	0.086	0.041	0.00093
Gender	Fixed	ns	ns	ns	0.057	0.094
Age \times gender	Fixed	ns	ns	ns	ns	ns
Age \times Smoking Status	Fixed	ns	ns	ns	ns	ns
Gender \times Smoking Status	Fixed	<0.05	ns	ns	<0.01	ns
s_p^2	Random	0.00032	0.00014	0.00072	0.00010	0.0000039*

Abbreviations: s_p^2 , ‘between-participant’ component of variance; ns, not significant.

*As expected, this difference was statistically significant for only the smoking participants and not the non-smoking group, and this p value corresponds to that cohort only.

Table 3.

Statistical significance (p values) of all main sources of variation (both fixed and random), and first-order interaction effects, from ANCOVA analysis of the model 1 salivary metabolite dataset.

variable was only close to statistical significance for glycine, with males having higher concentrations than females; cigarette smoking exerted highly significant effects on acetate, glycine and, of course, methanol. For the first-order interactions investigated, only the GS_{jk} effect was statistically significant, but only for acetate and glycine, i.e., non-additive responses to the four differential gender-smoking status combinations were observed. The random s_p^2 effect confirmed very highly significant differences ‘between-participants’ for all metabolites investigated.

3.5 Distinction between non-smoking and smoking participants using LF NMR-based MV metabolomics analysis

^1H NMR-based metabolomics analysis was, for the first time, utilized to determine the value of LF benchtop NMR analysis to discriminate between the salivary ^1H NMR profiles of the cigarette smoking *versus* non-smoking sampling groups. Therefore, the above ^1H NMR-detectable and validated 5 biomolecule variables, i.e. acetate, propionate, formate, methanol and glycine concentrations, as determined on the LF 60 MHz spectrometer, were employed to explore any MV differences between these. PCA, partial-least squares discriminatory analysis (PLS-DA), orthogonal partial least squares-discriminatory analysis (OPLS-DA), and agglomerative hierarchical clustering (AHC) techniques were used for this purpose, as was analysis using a RF model. **Figure 5(a)** and **(b)** show three-dimensional (3D) PLS-DA, and a two-dimensional (2D) OPLS-DA scores plots, arising from these forms of MV analysis, and both revealed that it was possible to achieve a satisfactory level of distinction between these two groups. Cross-validating $Q^2(R^2Y)$ values for these analyses were found to be high (0.721(0.803) and 0.709(0.786) respectively), and permutation tests performed for these models with 2000 permutations were very highly significant indeed ($p < 5.0 \times 10^{-4}$ in both cases). These analyses revealed that only methanol and its potential *in vivo* metabolite formate were important discriminatory variables for this comparison, which were both significantly upregulated in the smoking group (PLS-DA variable importance parameter (VIP) values of 1.88 and 0.81 respectively), with corresponding values for acetate, propionate and glycine being only 0.06, 0.64 and 0.63 respectively, i.e. acetate offered no discriminatory potential whatsoever. These values were reflected by the PCA strategy applied, which had very strong PC1 positive loadings for methanol and formate

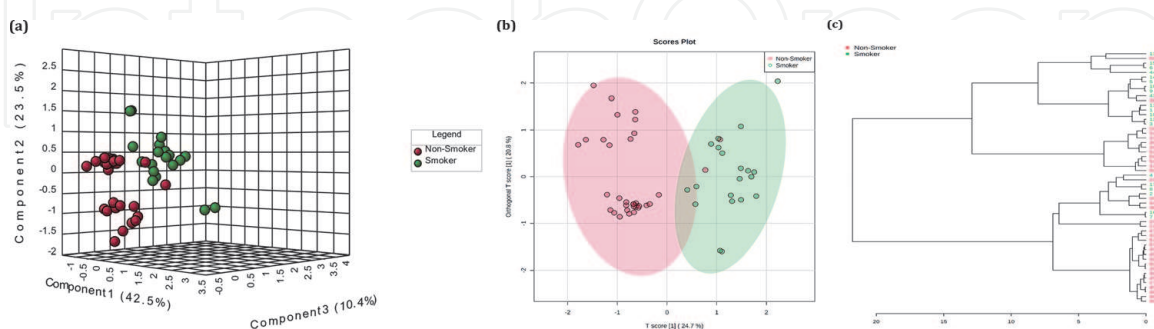


Figure 5.

(a) 3D PLS-DA scores plot of PC3 versus PC2 versus PC1 showing evidence for distinctive clusterings of tobacco cigarette-smoking and non-smoking participant classifications (green- and red-coded respectively). For this model, PC1, PC2 and PC3 accounted for 42.5, 23.5 and 10.4% of the total model variance respectively. (b) OPLS-DA plot of orthogonal T score {1} versus T score {1} also revealing a high level of distinction between the tobacco-smoking and non-smoking groups. (c) AHC analysis dendrogram of this dataset, revealing an at least moderate level of differentiation between the smoking and non-smoking groups, with notable sub-clusterings within each classification. Two misclassified non-smoking participant sample donors may arise from a self-reporting smoking bias, whereas the five misclassified smoking participants may result from prolonged durations between sample collection and their last smoking episode. Eight potential outlier samples detected in a provisional PCA were removed from the dataset prior to analysis.

(0.52 and 0.79 respectively), whereas those for acetate, propionate and glycine were negative, but only weakly so (-0.02 , -0.08 and -0.32 respectively). These loadings vectors are fully consistent with PC1 being derived from a cigarette tobacco smoking source only: in addition to being an oral microbiome catabolite, formate is an important *in vivo* metabolite of cigarette smoke-containing methanol, the route proceeding through a toxic formaldehyde intermediate [33]. However, it also appears that 2 or more of the non-smoking group of participants are classifying or clustering as cigarette smokers. In view of this observation, it remains a possibility that self-reporting bias may be involved in such cases, as further discussed in Section 4 below.

Similarly, AHC analysis confirmed an at least partial distinction between these two HSS sample classifications (**Figure 5(c)**). However, despite a reasonable-to-good level of discrimination, 2 of the samples donated by non-smoking participants appeared within the smoking group cluster, and 5 of the smoking ones are clustered with the non-smoking cohort. This is possibly explicable by self-reporting bias in the former case, but for the latter, it is possible that the participants concerned smoked their last cigarette some considerable period of time prior to sample collection.

Finally, an RF model demonstrated that 88 and 95% of the non-smoking and smoking participant donor samples were correctly classified (91% classification success rate overall). Hence, these MV comparisons demonstrate, for the first time, an important ^1H NMR-based metabolomics application which employs a non-stationary LF benchtop spectrometer. Four of the non-smoking participants were misclassified as smokers, whereas only one of the smokers was classified as a non-smoker with this analysis strategy.

3.6 Potential clinical and diagnostic significance of salivary metabolite tracking with LF benchtop NMR devices

In this study we have demonstrated the rapid, virtually non-invasive analysis of human saliva using a compact, LF 60 MHz benchtop NMR spectrometer. The major aim of the pilot investigations described here was to establish the abilities of LF NMR spectrometers to effectively perform the simultaneous quantitative analysis of a series of biomolecules in human saliva, and to consider their potential future value as trackable agents for the monitoring of selected oral diseases. We also critically examined limitations of the applications of this technique and their potential outcomes. Moreover, for the first time we have also applied 'state-of-the-art' NMR-linked metabolomics techniques to distinguish between saliva samples donated by both non-smokers and tobacco cigarette-smoking participants in a case study.

Overall, we have shown that 60 MHz ^1H NMR measurements can be employed to reliably determine selected salivary metabolite concentrations, with potentially much scope for future diagnostic and prognostic applications to oral health conditions. The authors are fully aware of issues related to resonance overlap at low magnetic fields in view of the dependence of resonance frequencies on static magnetic field strength, in which ^1H NMR signals appear to be broader, and with a spectrally-wider chemical shift range for all multiplets, with decreasing spectrometer operating frequencies. Such studies are therefore highly challenging in view of these inherent analyte selectivity considerations, along with potential sensitivity issues expected at such lower magnetic field strengths. However, we found that complications arising from overlapping resonances in complex biofluid samples such as saliva were minimal, or were circumventable, for the most common prominent resonances, and also for those located within relatively interference-free spectral regions. The ability of this LF NMR technique to detect exogenous agents

present in this biofluid should also be considered, for example the detection of drugs and other xenobiotics in saliva within specified time zones following their oral ingestion by humans.

Intriguingly, human saliva may afford a transference-dependent ‘*diluted picture*’ of chemopathological changes occurring throughout the human body, in addition to more concentrated, localized metabolic features within the oral environment itself, since a large number of biomolecules, and disease biomarkers (of a range of specificities) have the ability to transfer to this biofluid from blood via intra-, extra-, trans- and pericellular pathways, which highlight active transport or passive diffusion within the gingival sulcus and salivary glands [25]. Indeed, researchers are now increasingly promoting the employment of saliva as a clinical diagnostic medium [26], and such applications have significant widespread potential. Correspondingly, salivary (particularly parotid salivary) metabolomic and proteomic modifications appear to mirror those observed in human blood [27–29].

In 2002, the study reported by Silwood et al. [19], was described in Ref. [9] as the very first untargeted metabolomics investigation of human saliva. This unique investigation successfully identified a total of 63 biomolecules therein using 600 MHz ^1H NMR analysis, and quantified 11 key microbial-derived catabolites, 9 of which displayed very highly significant ‘between-participant’ components of variance. These markers included acetate and lactate, with excessive levels of their corresponding acids being viewed as primary end-point biomarkers involved in the aetiology of dental caries [30]. However, formic and pyruvic acids (both present at millimolar or near-millimolar concentrations in human saliva, the former being higher in smokers as found here) are stronger acids than lactic acid, and therefore may also exert pro-cariogenic activities. Hence, the LF 60 MHz NMR detection and quantification of salivary formate demonstrated in the current study may indeed offer some diagnostic and/or prognostic monitoring potential. However, propionate, along with *n*- and *iso*-butyrates, are considered to be primary microbial catabolites involved in periodontal disease progression [31, 32]. Correspondingly, salivary short-chain organic acids/anions serve as biomarkers for the growth, preponderance and catabolic activities of micro-organisms, and hence species-dependent patterns of these agents may serve as biomarkers of pathologically-mediated alterations to the salivary microbiome [19, 33].

Furthermore, whilst modifications in salivary formate concentrations have been previously linked to between-gender differences (i.e. elevated concentrations in males) [28], which we did not find here, our data provides evidence that one potential source for it is the oxidative metabolism of methanol as an ingested and/or inhaled environmental toxin. Indeed, salivary methanol levels are markedly upregulated via tobacco smoke inhalation [23], and/or alternative exogenous sources such as dietary ones. In our study, although salivary formate levels were *ca.* 2-fold greater in males, which may arise from the impaction of an increased smoking frequency for this gender in our smoking cohort, this difference was found not to be statistically significant (**Table 3**).

Interestingly, the study reported in [28] found that salivary citrate, lactate, pyruvate and sucrose levels were significantly upregulated in saliva samples collected from smokers over those of non-smokers, and formate was downregulated therein. Moreover, although salivary methanol concentrations were *ca.* 3-fold greater in smokers than in non-smokers in this study, as might be expected from the current one, this difference was found not to be statistically significant. Similarly, cigarette humectant-derived propane-1,2-diol levels were higher in samples collected from smoking participants in Ref. [28], although again this difference was found not to be significant. Additionally, that investigation detected and determined glucose and sucrose in saliva samples. However, in those collected following

the rigorous overnight fasting protocol involved in the current and our other studies, little or none of these carbohydrates are ^1H NMR-detectable in HSS samples collected from participant cohorts. Therefore, it appears that the quite limited pre-collection participant restrictions instigated in the study described in [28] was unsuccessful in completely precluding dietary-derived agents from the saliva samples analyzed. Indeed, participants were only required to not consume alcohol on the day of sample collection (and not also the evening before), and these samples were only collected at least 1 h following the last meal, which in our view is insufficient to remove interferences arising from dietary agents, along with those from alcoholic beverages consumed the previous day. Even with a protocol requesting that all participants refrain from the consumption of alcoholic drinks 24 h prior to the sample collection time-point (Section 2.1), traces of ethanol remain ^1H NMR-detectable in our saliva specimens when detected at operating frequencies of ≥ 400 MHz. Notwithstanding, generally ethanol consumed at time-points ≥ 24 h was not found in 60 MHz salivary ^1H NMR profiles in view of the lowered sensitivity of this approach. However, our pilot studies have also shown that if participants drank alcoholic beverages such as a beer, their salivary ethanol levels were indeed detectable and quantifiable using a 60 MHz benchtop NMR facility at least several hours or more thereafter. This observation clearly offers a high level of potential regarding future forensic investigations.

Likewise, citrate is only very rarely ^1H NMR-detectable in our HSS samples collected according to our rigorous overnight fasting protocol, and therefore its direct derivation from human diets (which serve as rich sources of this metabolite), and insufficient periods of fasting in Ref. [34], remains a strong possibility.

4. Limitations of the study

One major limitation of the application of LF salivary ^1H NMR analysis is inherent resonance overlap problems experienced at this operating frequency, which is much lower than those of more traditional MF or HF spectrometers coupled with restrictively-sized superconducting magnets (for example, those of 400–750 MHz operating frequencies). Hence, these resonance superimposition problems clearly give rise to major analytical limitations, notably in complex multianalyte biofluid spectra. Therefore, for future prospective studies involving the quantification of salivary biomolecules and/or xenobiotics, at least some level of caution should be applied when employing such devices. Unfortunately, these complications increase substantially when integrating resonances of higher first-order and more complex coupling patterns, which may markedly hinder such intensity determinations. Nevertheless, with the exception of the propionate- CH_3 resonance, these interference problems may be considered minimal for the determination of major, high concentration salivary metabolites, specifically those with prominent resonances in LF spectra obtained. Indeed, these signals have only low or negligible levels of superimposition with lower intensity signals, or appear in relatively ‘spectroscopically clean’ regions of the spectra acquired, for example formate. Therefore, although the LF 60 MHz ^1H NMR profiles of HSSs are largely commanded by resonances of the highest intensity and with simple coupling patterns and orders, and/or those arising from metabolites of high concentrations, in principle this novel NMR strategy potentially offers valuable quantitative information for those detectable at lower levels, most notably with the advent of higher-field compact benchtop instruments which operate at frequencies of 80 or 100 MHz.

As observed and further explored in Ref. [10], an additional limitation of ^1H NMR-based metabolomics studies featuring LF NMR spectrometers is the intensity-

diminishing effects of the H₂O/HOD signal presaturation protocol, notably for signals located close to its chemical shift value ($\delta = 4.8$ ppm). However, although such effects substantially influence the C1-H resonances of both the α - and β -glucose anomers ($\delta = 5.25$ and 4.63 ppm respectively), such hurdles may be surmounted by the use of rigorous calibration processes with biomolecule standard solutions, and by the possible integration of alternative resonances derived from the agents affected, namely those with δ values sufficiently distant from the H₂O/HOD secondary irradiation one. An additional limitation arises from some significant differences between intramolecular ¹H relaxation times for a number of salivary metabolites, and also some long-range coupling phenomena, results which will be reported in detail elsewhere.

Finally, for the *prima facie* metabolomics investigation conducted here, data available in **Figure 5** indicates that this study may involve a small but significant level of self-reporting bias, since two, or perhaps more, of the non-smoking participant samples appeared to co-cluster with the tobacco-smoking cohort in PCA, PLS-DA and OPLS-DA scores plots, and AHC dendrograms, as did those of five of the smoking cohort with the non-smoking group. These apparent erroneously-clustered non-smoking participants may represent those with only a limited or very limited smoking preference, but who preferred to report themselves as ‘non-smokers’ in this investigation in view of their low smoking incidences and/or smoking irregularities, for example those known as ‘closet smokers’. Further investigations to explore this are currently in progress in our laboratories.

5. Conclusions

In this study, we have evaluated the viability of low-field (LF) benchtop ¹H NMR analysis technologies for metabolomics investigations of human saliva. This novel, convenient and near-portable technique was used to detect and/or potentially quantify and hence monitor up to 15 potentially healthcare-impacting oral metabolites in healthy human saliva, a strategy prospectively offering much potential for the direct ‘on-site’ testing of biofluids from patients affected by oral health or related conditions at clinical locations.

We report the detection of typical salivary metabolites, including propionate, acetate, succinate, glycine, dimethylamine, trimethylamine, methanol, formate and aromatic amino acids all at an operating frequency of only 60 MHz. However, quantification of the salivary levels of biomolecules was limited to only five of those with the most prominent ¹H NMR signals, although succinate (singlet signal, $\delta = 2.405$ ppm) could also be considered if not significantly quantitatively impacted by salivary pyruvate (*s*, $\delta = 2.388$ ppm) and/or glutamine (*m*, $\delta = 2.42$ ppm) resonances. Indeed, since the singlet resonances of formate, methanol and glycine did not suffer from significant resonance overlap issues, and were therefore quite clearly resolved, the direct LF NMR determination of these biomolecules was possible. Excessive salivary levels of organic acid anion catabolites may serve as key biomarkers of the pathogenesis and development of dental caries [19, 20, 30] and periodontal diseases [19, 32], and herein it was found that both salivary acetate and propionate, which represent biomarkers of dental caries and periodontal diseases respectively, could be readily quantitated in this biofluid, despite some bioanalytical concentration limit/interference problems encountered with the latter.

This study also demonstrated for the first time that LF ¹H NMR-linked metabolomics analysis could be employed to discriminate between the salivary biomolecular profiles of tobacco-smoking and non-smoking participants. When detectable at sufficient salivary concentrations, these analyses may be transferable

to the detection and perhaps quantification of exogenous agents such as ingested drugs in this biofluid. Moreover, the approach outlined here may indeed offer some forensic applications involving the identification of illicit drugs in human saliva samples collected at crime scenes, provided that such analytes have resonances present in spectrally-clear regions unaffected by overlapping signal interferences arising from endogenous species.

Notwithstanding, currently the authors recommend that future LF NMR investigations of human saliva and oral diseases should focus on the more prominent resonances present in spectra acquired, most notably those with relatively simple first-order coupling patterns (i.e., singlets, doublets, triplets, etc.) for quantification purposes, with the exception of those in which the targeted analyte has resonances (s) located in 'spectroscopically-clear' regions. However, future developments in LF benchtop NMR technologies operating at higher frequencies may serve to provide effective solutions to these issues.

The authors also suggest that this technique may also serve valuable diagnostic and/or prognostic tracking purposes in a clinical context for a range of human diseases. In principle, this non-stationary multi-analytical technique may be employed as a sensitive means of monitoring the salivary metabolic status of patients suffering from oral diseases, and potentially also physiologically-remote conditions, directly at dental surgery and primary healthcare sites, hospitals and hospital laboratories, and perhaps also community pharmacies. In this manner, such patient-contact sites may offer significant diagnostic and monitoring potential for oral health practitioners.

Acknowledgements

The authors wish to express the highest level of thanks to all project participants for the kind donation of their saliva samples for ^1H NMR analysis. We are also very grateful to Magritek GmbH, Aachen, Germany for the academic loan of a 60 MHz Magritek Spinsolve Ultra Benchtop NMR Spectrometer for research purposes.

Conflict of interest


The authors declare no conflict of interest.

Author details

Benita C. Percival, Angela Wann, Sophie Taylor, Mark Edgar, Miles Gibson and Martin Grootveld*
Leicester School of Pharmacy, De Montfort University, Leicester, United Kingdom

*Address all correspondence to: mgrootveld@dmu.ac.uk

IntechOpen

© The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Grootveld M, Percival B, Gibson M, Osman Y, Edgar M, Molinari M, et al. Progress in low-field benchtop NMR spectroscopy in chemical and biochemical analysis. *Analytica Chimica Acta*. 2019;**1067**:11-30. DOI: 10.1016/j.aca.2019.02.026
- [2] Teng Q. NMR-based metabolomics. In: *Structural Biology*. Boston, MA, US: Springer; 2013. pp. 311-392. DOI: 10.1007/978-1-4614-3964-6_9
- [3] Trivedi DK, Hollywood KA, Goodacre R. Metabolomics for the masses: The future of metabolomics in a personalized world. *European Journal of Molecular Clinical Medicine*. 2017;**3**: 294-305
- [4] Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proceedings of the National Academy of Sciences*. 1971;**68**: 2374-2376
- [5] Nicholson JK, Buckingham MJ, Sadler PJ. High resolution ^1H NMR studies of vertebrate blood and plasma. *Biochemical Journal*. 1983;**211**:605-615
- [6] Nicholson JK, O'Flynn MP, Sadler PJ, MacLeod AF, Juul SM, Sonksen PH. Proton-nuclear-magnetic resonance studies of serum, plasma and urine from fasting normal and diabetic subjects. *Biochemical Journal*. 1984;**217**:365-375
- [7] Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery*. 2016;**15**:473-484
- [8] Meiss E, Werner P, John C, Scheja N, Herbach N, Heeren J, et al. Metabolite targeting: Development of a comprehensive targeted metabolomics platform for the assessment of diabetes and its complications. *Metabolomics*. 2016;**12**:1-15
- [9] Dame ZT, Aziat F, Mandal R, Krishnamurthy R, Bouatra S, Borzouie S. The human saliva metabolome. *Metabolomics*. 2015;**11**: 1864-1883. DOI: 10.1007/s11306-015-0840-5
- [10] Percival BC, Grootveld M, Gibson M, Molinari M, Jafari F, Sahota T, et al. Low-field, benchtop NMR spectroscopy as a potential tool for point-of-care diagnostics of conditions: Validation, protocols and computational models. *High-Throughput*. 2018;**8**(1):2. DOI: 10.3390/ht8010002
- [11] Edgar M, Percival BC, Gibson M, Jafari F, Grootveld M. Low-field benchtop NMR spectroscopy as a potential non-stationary tool for point-of-care urinary metabolite tracking in diabetic conditions. *Diabetes Research and Clinical Practice*. 2021;**171**:108554. ISSN 0168-8227. DOI: 10.1016/j.diabres.2020.108554
- [12] Gardner A, Carpenter G, So PW. Salivary metabolomics: From diagnostic biomarker discovery to investigating biological function. *Metabolites*. 2020; **10**(2):47 10.3390/metabo10020047
- [13] Miller JN, Miller JC. *Statistics and Chemometrics for Analytical Chemistry*. London, UK: Pearson Higher Education; 2018. ISBN-13: 9781292186719
- [14] Grootveld M. *Metabolic Profiling: Disease and Xenobiotics*. Cambridge, UK: Royal Society of Chemistry (Issues in Toxicology Series); 2014. ISBN: 1849731632 (395 p)
- [15] Ruhl S. The scientific exploration of saliva in the post-proteomic era: From database back to basic function. *Expert Reviews in Proteomics*. 2012;**9**:85-96. DOI: 10.1586/epr.11.80
- [16] Rad M, Kakoie S, Niliye Brojeni F, Pourdanghan N. Effect of long-term

- smoking on whole-mouth salivary flow rate and oral health. *Journal of Dental Research, Dentistry and Clinical Dental Prospects*. 2010;**4**(4):110-114. DOI: 0.5681/jodddd.2010.028. Epub 2010 Dec 21
- [17] Andrecut MP, GPU. Implementation of iterative PCA algorithms. *Journal of Computational Biology*. 2009;**16**(11):1593-1599. DOI: 10.1089/cmb.2008.0221
- [18] Palarea-Albaladejo J, Martín-Fernández JA. Values below detection limit in compositional chemical data. *Analytica Chimica Acta*. 2013;**764**:32-43. ISSN 0003-2670. DOI: 10.1016/j.aca.2012.12.029
- [19] Silwood CJL, Lynch E, Claxson AWD, Grootveld MC. ¹H and ¹³C NMR spectroscopic analysis of human saliva. *Journal of Dental Research*. 2002;**81**:422-427
- [20] Silwood CJL, Lynch E, Seddon S, Sheerin A, Claxson AWD, Grootveld MC. ¹H-NMR analysis of microbial-derived organic acids in primary root carious lesions and saliva. *NMR in Biomedicine*. 1999;**12**:345-356
- [21] Gardner A, Parkes HG, So P-W, Carpenter GH. Determining bacterial and host contributions to the human salivary metabolome. *Journal of Oral Microbiology*. 2019;**11**(1):1617014. DOI: 10.1080/20002297.2019.1617014
- [22] Pang Z, Chong J, Zhou G, de Lima A, Morais D, Chang L, et al. MetaboAnalyst 5.0: Narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research*. 2021;**49**:W388-W396
- [23] Percival BC, Wann A, Masania J, Sinclair J, Sullo N, Grootveld M. Detection and determination of methanol and further potential toxins in human saliva collected from cigarette smokers: A ¹H NMR investigation. *JSM Biotechnology Bioengineering*. 2018;**5**(1):1081-1088
- [24] Park YD, Jang JH, Oh YJ, Kwon HJ. Analyses of organic acids and inorganic anions and their relationship in human saliva before and after glucose intake. *Archives of Oral Biology*. 2014;**59**(1):1-11. DOI: 10.1016/j.archoralbio.2013.10.006. Epub 2013 Oct 24
- [25] Spielmann N, Wong D. Saliva: Diagnostics and therapeutic perspectives. *Oral Diseases*. 2011;**17**:345-354
- [26] Wong DT. Salivary diagnostics powered by nanotechnologies, proteomics and genomics. *Journal of the American Dental Association*. 2006;**137**:313-321
- [27] Sugimoto M, Wong DT, Hirayama A, Soga T, Tomita M. Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. *Metabolomics*. 2010;**6**:78-95
- [28] Takeda I, Stretchay C, Barnabya P, Bhatnagera K, Rankinb K, Fub H, et al. Understanding the human salivary metabolome. *NMR in Biomedicine*. 2009;**22**:577-584
- [29] Chiappin S, Antonelli G, Gatti R, De Palo EF. Saliva specimen: A new laboratory tool for diagnostic and basic investigation. *Clinica Chimica Acta*. 2007;**383**:30-40
- [30] Hoppenbrouwers PMM, Driessens FCM. The effect of lactic and acetic acid on the formation of artificial caries lesions. *Journal of Dental Research*. 1988;**67**:1466-1467
- [31] Niederman R, Zhang J, Kashket S. Short-chain carboxylic-acid stimulated, PMN-mediated gingival inflammation. *Critical Reviews in Oral Biology and Medicine*. 1997;**8**:269-290

[32] Niederman R, Buyle-Bodin Y, Lu BY, Naleway C, Robinson P, Kent R. The relationship of gingival crevicular fluid short chain carboxylic acid concentration to gingival inflammation. *Journal of Clinical Periodontology*. 1996; **23**:743-749

[33] Guerrant GO, Lambert MA, Moss CW. Analysis of short-chain acids from anaerobic bacteria by high-performance liquid chromatography. *Journal of Clinical Microbiology*. 1982;**16**:355-360

[34] Gardner A, Parkes HG, Carpenter GH, So P-W. Developing and standardising a protocol for quantitative proton nuclear magnetic resonance (1h nmr) spectroscopy of saliva. *Journal of Proteome Research*. 2018;**17**:1521-1531

IntechOpen