

Investigating Metabolomics Analytical Techniques to Analyze and Better Understand the Human  
Saliva Metabolome

Research Thesis

Presented in Partial Fulfillment of the Requirements for Graduation with Research Distinction in  
the Undergraduate Colleges of The Ohio State University

By

Jack Gothier

The Ohio State University College of Food Agriculture and Environmental Sciences

Department of Food Science and Technology

December 2019

Project Advisor: Dr. Emmanuel Hatzakis

## Abstract

As one of the most rapidly evolving fields in biomedical research, metabolomics offers small molecule compositional analysis of biological fluids, cells, and tissues. The metabolome itself can provide insights into biological mechanisms of disease. Chemical phenotyping using biospecimens and metabolomics can lead to the identification of biomarkers which may underlie diseases. Saliva, in particular, contains multiple biomarkers making it suitable for metabolomics. Despite the fact that saliva testing already exists and offers a non-invasive and rapid method of analysis, characterization of the human saliva metabolome and saliva-related metabolomics applications are not very extensive. Additionally, current nuclear magnetic resonance (NMR) saliva analysis often involves several time-consuming steps that increase the time of analysis and may have a negative impact on the analysis. This project aims to investigate more suitable metabolomic analysis techniques of NMR on saliva. A comparison between pulse sequences (CPMG and 1D NOESY) and magnetic fields (600 and 700 MHz), as well as evaluation of sample storage stability was conducted with 1:1 saliva, sodium phosphate buffer solution samples. Results indicated that metabolite peak shifts were not significantly affected by NMR sequences and instruments but the water peak suppression in CPMG pulse sequencing with 700 MHz magnetic field, deemed most efficient. In addition, the CPMG experiment produces spectra with slightly better baseline in some cases. Sample stability persisted in spectra analysis between tested samples based on storage conditions of 1 week in the freezer and 24 hours at room temperature. After method optimization, NMR spectra were collected and used to assess the influence of genetic and environmental factors on the saliva metabolome. More specifically, 119 saliva samples from smokers and non-smokers were analyzed in total. Finally, a systematic 2D NMR study was conducted to perform the unambiguous assignment of the NMR spectra of saliva. This is important for performing biomarker identification and pathway analysis which can further contribute to foods impact on the saliva metabolome.

## Introduction

Metabolomics is known as the systematic study of metabolites within biological species such as cells, tissues, biofluids, or even whole organisms.<sup>1</sup> Considering that metabolomics has been rapidly on the rise in the past decade, relatively new and improved analytical methods have been developed and are still developing over time. The advances and improvements of various analytical techniques give a more insightful evaluation of the subject at study. Two of the more well-known techniques used in metabolomics are Nuclear Magnetic Resonance (NMR) spectroscopy and Mass Spectrometry (MS). NMR is based on the energy absorption and re-emission of the atom nuclei due to variation in the external magnetic field.<sup>2</sup> The most common nuclei used in metabolomics is proton ( $^1\text{H}$ ), whereas other atoms, such as carbon ( $^{13}\text{C}$ ) and phosphorus ( $^{31}\text{P}$ ) are also used to provide additional information about the structure and the abundance of metabolites. In cases where there is an extensive signal overlapping in the 1D NMR spectrum, 2D NMR experiments such as correlation spectrometry (COSY), total correlation spectroscopy (TOCSY), and heteronuclear single quantum coherence may be applied allowing for better NMR peak separation.<sup>2</sup>

Biofluids, such as blood and urine are often used in metabolomics. Although human saliva has significant advantages compared to other biofluids, as it is easily accessible and straightforward to collect in an inexpensive and non-invasive manner without the introduction of stress or pain, it is still rarely used in metabolomics studies.<sup>1</sup> Saliva is an extremely important and abundant biofluid within the oral cavity and has a known composition of mainly water with several minor components such as mucus, digestive enzymes, antibacterial peptides, bacterial cells, and low molecular weight metabolites.<sup>3</sup> The salivary proteome provides a large source of useful biomarkers that can be important in disease diagnosis.<sup>4</sup> Aside from the proteome aspect of

saliva however, the metabolome composition is not completely known even though several metabolites have been identified.<sup>3</sup> In addition, a multistep analysis of the small molecules that includes centrifugation and filtration steps is usually adopted to remove the large size molecules.<sup>5</sup> However, these steps increase the experimental time and may induce additional variation to the samples, which potentially affects the reliability of the metabolomics analysis.

In this study, we aimed to determine the optimum conditions for the analysis of saliva in order to further understand its metabolome. More specifically, we tested various NMR-based approaches, such as the first increment of NOESY pulse sequence with presaturation and the Carr-Purcell-Meiboom-Gill sequence (CPMG) in combination with common analytical practices such as centrifugation/filtration. After evaluating the impact of these approaches to the NMR spectrum, we performed a comparative metabolomic analysis using saliva samples from smokers and non-smokers to investigate the effect of smoking in saliva metabolome.

### **Problem Justification**

This study was conducted to further enhance the analysis of saliva metabolome through NMR. Because the experimental times in high throughput metabolomics studies can be long due to the large number of samples, it is crucial to understand how the stability and biochemical composition of samples can be impacted overtime. The way in which the samples are stored during other analysis, may impact the outcome of the metabolites appearance in saliva. Understanding the affects storage has on the overall stability of the saliva samples will allow for better planning and the potential to set up longer experiments. In addition, applying optimum experimental conditions is crucial in order to get the best metabolomic analysis of the sample. This includes determining if certain pulse sequences have any effect on the overall spectra

quality. Ideally, a particular pulse sequence and NMR instrument are expected to generate spectra with high resolution, sufficient water peak suppression, and smooth baseline.

### **Hypothesis/Objective**

It is hypothesized that the combination of NMR pulse sequences with analytical and chemical methods, will lead to the development of experimental protocols that enable the fast and reliable analysis of saliva. These protocols are expected to overcome problems related to current methods of analysis, such as signal broadening, long experimental times and aggregation of biochemically important metabolites. Furthermore, it is hypothesized that saliva samples from smokers and non-smokers can be effectively analyzed using NMR.

This project aims to investigate the impact of sample storage and other experimental parameters on the metabolome composition of saliva and the quality of the NMR spectra. This project aim can be carried out through three objectives:

*Objective 1:*

Determine the impact storage conditions have on the stability of metabolites appearance in saliva.

*Objective 2:*

Determine the effect of pulse sequence and magnetic field on spectral quality.

*Objective 3:*

Apply the optimum experimental condition for the untargeted metabolomics analysis of a large number of saliva samples and perform 2D NMR-based assignment of individual metabolites.

## **Procedures/Methods**

### *Saliva samples:*

Saliva samples from smokers and non-smokers were provided by Dr. Purnima Kumar through The Ohio State College of Dentistry.

### *Sample preparation:*

A 1:1 ratio of phosphate buffer to saliva was mixed in samples.

Four different sample preparations were tested in this project.

*i)* Saliva samples were analyzed using the first increment of NOESY pulse sequence with presaturation; *ii)* Saliva samples were centrifuged (14,000 rpm for 30 min at 4 °C) and then analyzed using the first increment of NOESY pulse sequence with presaturation; *iii)* Saliva samples were analyzed using the CPMG pulse sequence; *iv)* Saliva stability was tested with immediate analysis, 24 hour room temperature analysis, and 1 week freezing analysis.

### *NMR-based metabolomics:*

The NMR experiments were performed in the NMR facility at OSU using a 600 and 700 MHz instrument equipped with a 5mm TCI cryoprobe. 2D NMR are applied on selected samples to confirm the identity of the specific metabolites.

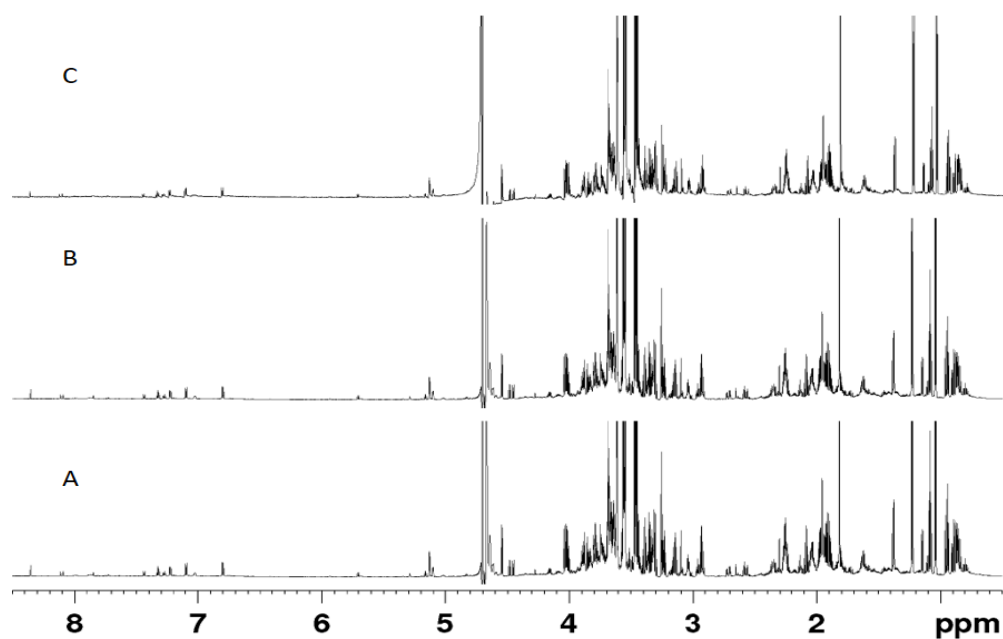
### *General Methods:*

After saliva samples were obtained, a sample was prepared with 300 µl of centrifuged saliva and 300 µl phosphate buffer at pH 7.16. The saliva sample was centrifuged at 0.8G for 10 minutes at 4°C and the supernatant was transferred into an NMR tube followed by the buffer addition. Saliva samples were analyzed immediately, after 24 hours of room temperature storage, and after 1 week of freezing storage. NMR analysis was conducted with a large number of

samples using pulse sequences of CPMG with a 700MHz magnetic field as well as a pulse sequence of 1D-NOESY under 600 and 700 MHz magnetic fields.

## Results

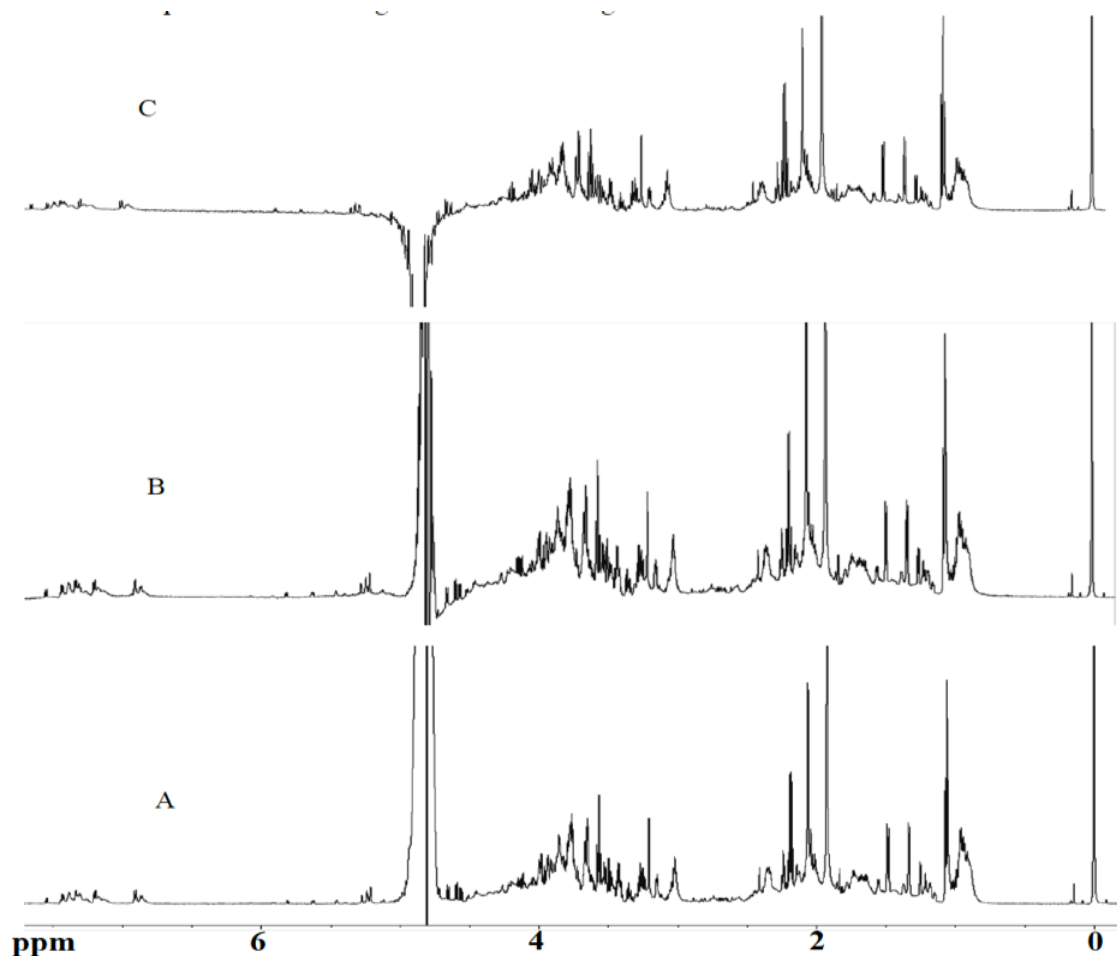
For the saliva stability analysis, three samples were used, one that was analyzed immediately after preparation, one that was stored at room temperature for 24 hours, and one that was frozen for a week. The results of the spectra are shown below in Figure 1. Other storage conditions such as, several days of refrigeration were also examined and no differences in the biochemical composition, as determined by NMR, were observed.



**Figure 1.** Saliva  $^1\text{H}$  NMR analysis at 800 Hz. Original sample (A), after 24 hours at room temperature (B) and after one week being frozen (C).

To determine the potential impact of different pulse sequences and magnetic fields on the NMR analysis, experiments were conducted by analyzing the same sample using two pulse sequences,

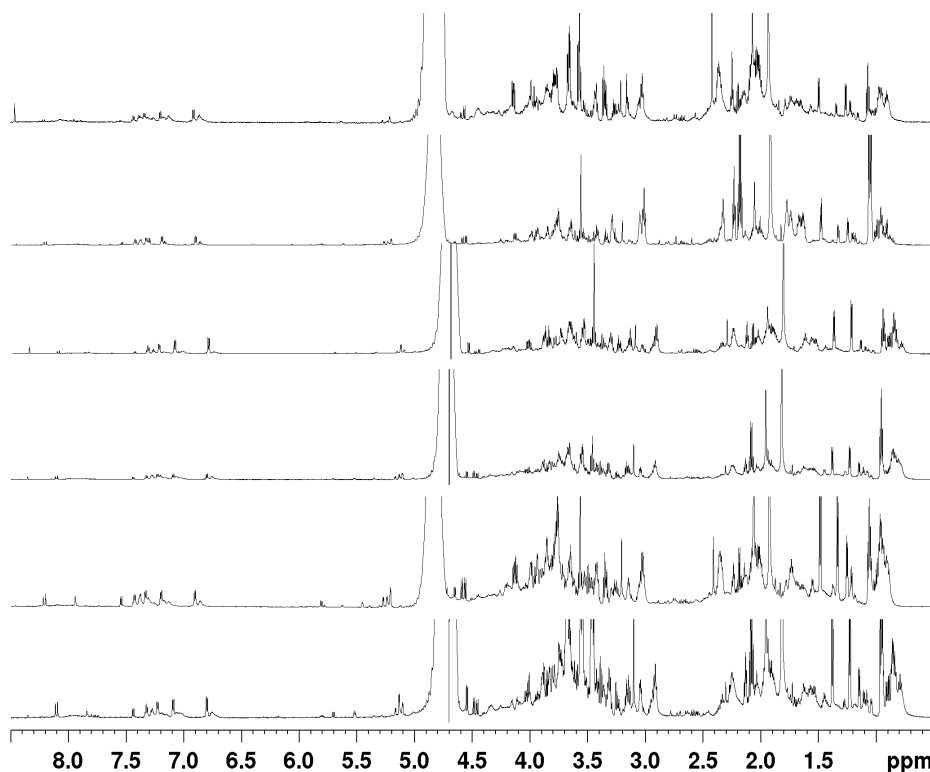
namely CPMG and 1D-NOESY and two instruments, a 600 MHz and a 700 MHz. The acquired spectra are shown in Figure 2.



**Figure 2.** Analysis of the same saliva sample using different 1D  $^1\text{H}$ -NMR pulse sequences and magnetic fields. CPMG - 700 MHz (A); NOESY - 700 MHz (B); NOESY - 600 MHz (C).

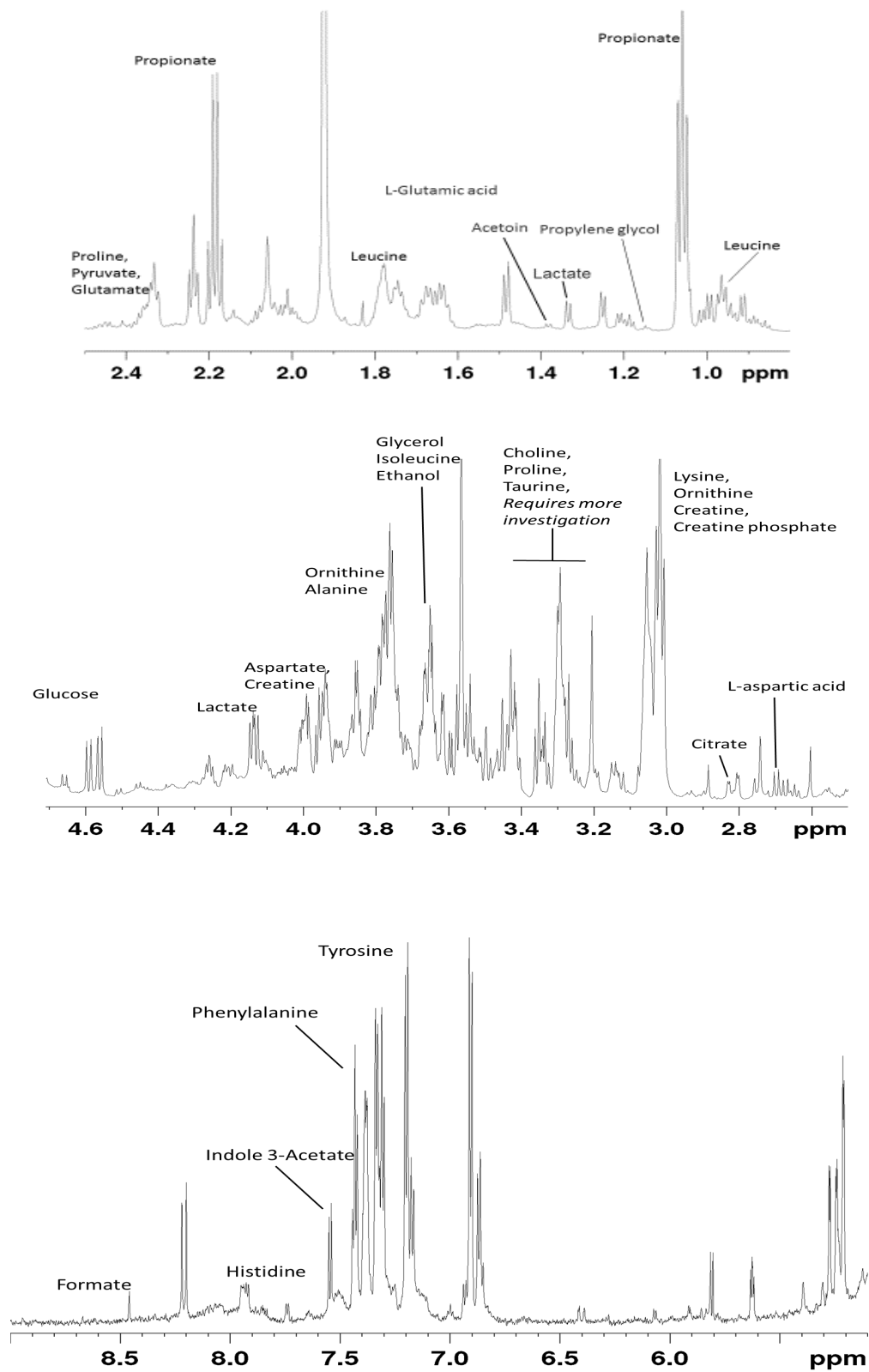
Once analysis parameters were set, a large set of samples were run under the optimized conditions as seen in Figure 3 below. The difference in samples were between smokers and non-smokers which undergoes further investigation in future studies, by applying multivariate statistical analysis.





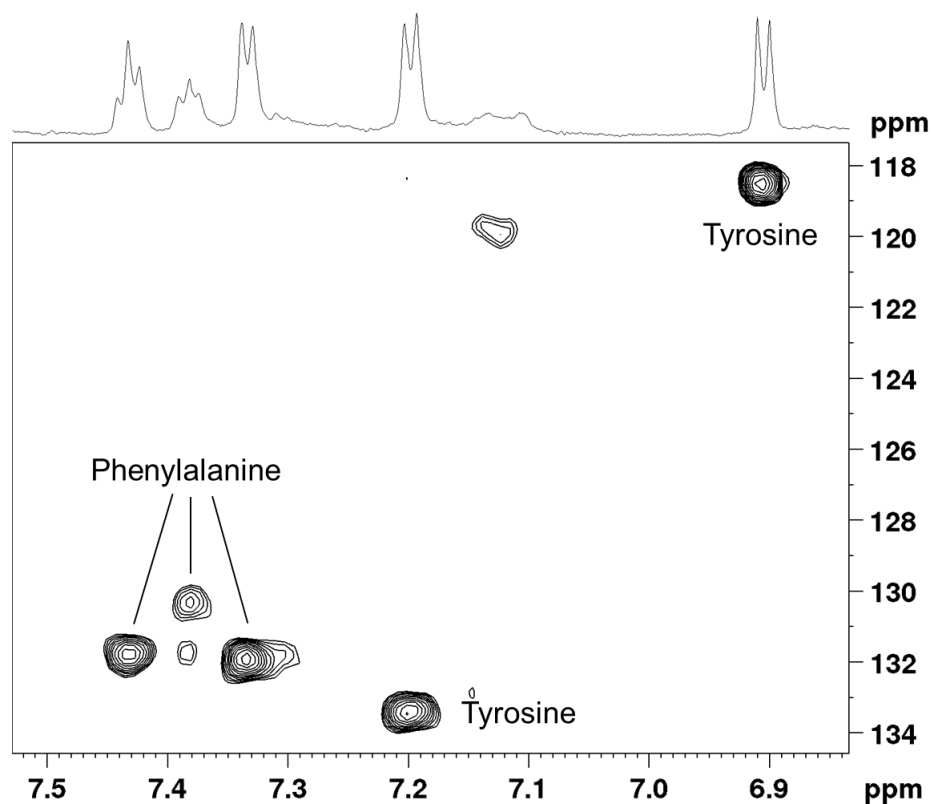
**Figure 3.** NMR analysis of a large number of samples with optimized conditions on a 700 MHz instrument. As an example, six saliva samples of smokers and non-smokers are shown.

A crucial factor for any compositional analysis and metabolomics study is the correct assignment of the signals in the  $^1\text{H}$  NMR spectra. For that purpose, we performed a 2D NMR analysis, which in combination with literature and databases such as the Human Metabolome Database (HMDB) and tools such as COLMAR (Complex Mixture Analysis by NMR)<sup>6</sup> allowed the assignment of several saliva metabolites in the  $^1\text{H}$  NMR spectra. A typical  $^1\text{H}$  NMR spectrum of a saliva sample with the assignments in three regions is shown in Figure 4.



**Figure 4.** Example of the NMR assignment of saliva in the aliphatic region (A) sugar/acids region (B) and the aromatic region (C).

As an example of the 2D NMR analysis, the HSQC spectrum of saliva in the aromatic region is shown in Figure 5.



**Figure 5.** HSQC spectrum of saliva in the aromatic region.

## Discussion

As shown in Figure 1, the spectra are very similar for all saliva samples that were tested. Persistence of the saliva stability held for all three samples stored at different temperatures for different times. The 24 hours period under room temperature with no observed variation in peak allows for long-overnight experiments to be set up for metabolomics analysis. When looking at the different pulse sequences and magnetic fields that saliva is analyzed at, the CPMG sequence was found to be very efficient in terms of removing signals of macromolecules even when using short centrifuge times. More specifically, looking at Figure 2, the CPMG at 700 MHz resulted in

spectra with efficient water peak suppression, smooth baseline, and good spectral resolution. In further analysis, using 2D NMR, literature, databases (HMDB) and tools (COLMAR), NMR assignments of many important metabolites appearance in saliva is shown in Figure 4.

Further research should utilize the metabolomic analytical techniques found to be most effective in this study, to further investigate the saliva metabolomic profile of smokers, non-smokers, obese, and nonobese individuals.

## **Conclusion**

This project increased our understanding about the optimized experimental conditions that are required to achieve a rapid and reliable high-resolution NMR analysis of saliva. NMR is a robust and reliable method for the analysis of saliva. Results showed that overnight studies can be safely performed as the saliva metabolome was found to be stable after 24 hours of storage at room temperature. Lastly, CPMG pulse sequencing with 700 MHz magnetic field deemed most efficient and can be used to further investigate the saliva metabolome. Because our approaches will be applied in a set of saliva samples from smokers and non-smokers in the future, we anticipate to identify several discriminatory metabolites that may be linked to specific metabolic pathways that can be affected by smoking.

## **Acknowledgements**

Special thanks to Food Science Undergraduate Research Experience (FoodSURE) for providing funding and to Dr. Purnima Kumar for providing the saliva samples. Finally, many thanks to Fenfen Tang for her help with the NMR analysis.

## Literature Cited

- <sup>1</sup>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. 2015. The human saliva metabolome. *Metabolomics* DOI 10.1007/s11306-015-0840-5.
- <sup>2</sup>Bothwell J. H., Griffin J. L. 2011. An introduction to biological nuclear magnetic resonance spectroscopy. *Biol. Rev. Camb. Philos. Soc.* 86, 493–510.10.1111/j.1469-185X.2010.00157.
- <sup>3</sup>Bertram H.C., Eggers N., Eller N. 2009. Potential of Human Saliva for Nuclear Magnetic Resonance-Based Metabolomics and for Health-Related Biomarker Identification. *Crit Rev Anal. Chem.* 81, 9188–9193.
- <sup>4</sup>Lorenzo-Pouso, A.I., Perez-Sayans, M., Bravo, S.B., Lopez-Jorne, P., Garcia-Vence, M., Alonso-Sampedro, M., Carballo, J., Garcia-Garcia, A., 2018. Protein-Based Salivary Profiles as Novel Biomarkers for Oral Diseases. *Dis. Markers*. DOI 10.1155/2018/6141845.
- <sup>5</sup>Silwood, C. J., Lynch, E., Claxson, A. W., & Grootveld, M. C. (2002). <sup>1</sup>H and (<sup>13</sup>C NMR spectroscopic analysis of human saliva. *Journal of Dental Research*, 81(6), 422–427.
- <sup>6</sup>Bingol, K.; Li, DW.; Zhang, B.; Brüscheweiler, R.; Comprehensive metabolite identification strategy using multiple two-dimensional NMR spectra of a complex mixture implemented in the COLMARm web server; *Anal. Chem*; 2016, 88, 12411-12418.