THE COMPOSITION OF QUANTITATIVE NUTRITIONAL ANALYSIS USING RAMAN SPECTROSCOPY

An Undergraduate Research Scholars Thesis

by

AXELL RODRIGUEZ

Submitted to the LAUNCH: Undergraduate Research office at
Texas A&M University
in partial fulfillment of requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by
Faculty Research Advisor: Dr. Dmitry Kurouski

May 2023

Major: Biochemistry

RESEARCH COMPLIANCE CERTIFICATION

Research activities involving the use of human subjects, vertebrate animals, and/or biohazards must be reviewed and approved by the appropriate Texas A&M University regulatory research committee (i.e., IRB, IACUC, IBC) before the activity can commence. This requirement applies to activities conducted at Texas A&M and to activities conducted at non-Texas A&M facilities or institutions. In both cases, students are responsible for working with the relevant Texas A&M research compliance program to ensure and document that all Texas A&M compliance obligations are met before the study begins.

I, Axell Rodriguez, certify that all research compliance requirements related to this Undergraduate Research Scholars thesis have been addressed with my Faculty Research Advisor prior to the collection of any data used in this final thesis submission.

This project did not require approval from the Texas A&M University Research Compliance & Biosafety office.

TABLE OF CONTENTS

Page

ABSTRACT

The Composition of Quantitative Nutritional Analysis Using Raman Spectroscopy

Axell Rodriguez
Department of Biochemistry & Biophysics
Texas A&M University

Faculty Research Advisor: Dr. Dmitry Kurouski
Department of Biochemistry & Biophysics
Texas A&M University

With the substantial global population growth, it is estimated that the population will be nearing around 9.7 billion by 2050. With this in hand, the demand for food production is an imminent matter in question that must be fulfilled to avert a food crisis of global malnutrition. This is where we introduce our research to enhance the upcoming, cutting-edge technology that is Raman spectroscopy (RS), known for being non-invasive, non-destructive, and chemical free. Through this method, we aim to strengthen and test the instrument's potential in its ability to accurately identify and classify different vibrational frequencies to enhance our nutritional analysis. Through advancements in nutritional analysis, we can identify important macromolecules such as carbohydrates, proteins, and fats to reach a precise quantification of one's nutrient intake. By analyzing popular international foods such as ramen noodles, bread, and similar carbohydrate foods, we can promote the use of RS to create a non-destructive and chemical-free method in the production of these foods. In this study, we hypothesized that through RS we could collect substantial data and process this data to produce accurate results to enhance our nutritional analysis. Our results indicate that RS is successful in its ability to

differentiate between foods and, more specifically, food brands. The data collected is then analyzed using a multi-paradigm programming language to further enhance this precise nutritional analysis. In this study we use RS to analyze and distinguish the samples tested, however, we are also studying the potential of RS towards our agricultural sciences to impact the impending global issue.

ACKNOWLEDGEMENTS

Contributors

I would like to thank my faculty advisor, Dr. Dmitry Kurouski, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my father for their encouragement, to my mother for her patience and love, and to my friends that have been by my side supporting me throughout my career.

1. INTRODUCTION

In 2015, a total of 107.7 million children and 603.7 million adults were obese, doubling the prevalence of obesity since 1980 in more than 70 countries. High body-mass index (BMI) has accounted for 4.0 million deaths where two-thirds of this statistic corresponded to cardiovascular disease [1]. Most causes of cardiovascular diseases can be traced back to unhealthy lifestyles such as poor diet and lack of physical activity. A noticeable factor that may lead to cardiovascular disease is high blood pressure (HBP), this is when the force of blood flowing in one's blood vessels is higher than an average persons. HBP is also known as hypertension, a systolic value of 140 mm Hg or higher, and a diastolic value of 90 mm Hg or higher. Ultimately, the leading cause of HBP is a heavy salt diet with little to no physical activity. Although some may suggest eating a healthier diet to be a simple task, it may not be as easy as one might think. Fortunately, we live in a highly advanced society where we can produce food at an extremely low cost with a high rate and efficiency. The production rate of foods has increased by tenfold over the years with such advancements in technology. However, with these advancements we face many shortcuts in the production and distribution of our food products.

With advancements in production such as the assembly line, later altered from automobiles, the upkeep for these processed foods had to be improved. Food preservation has been around since the early 12,000 B.C. [2]. These early methods included the use of the sun and wind to dry out the food of its moisture. A well-known method included the utilization of salt to cure, or dehydrate, the foods which allowed for a relatively long shelf time. It wasn't until the acknowledgement of microbes in foods were understood, that food preservation took a higher toll on public health. Important hurdles in food preservation include temperature, water activity, acidity, redox potential, preservatives (nitrite, sorbate, sulfite, etc.), and competitive microorganisms [3].

Let us apply the concept of food preservation in a notable popular international food that is stereotypically known to be the iconic food for college students, instant ramen noodles. Before that we must understand the origin of how ramen noodles first came to be. It is stated in the journal article *Noodle Odyssey: East Asia and Beyond,* that Japanese ramen, named *sina soba,* originated from Chinese noodles after their war with China in the years 1894 to 1895 [4]. Due to postwar repercussions, Japan faced severe food shortages and were reluctant to resolve their malnutrition crisis. This is where ramen noodles, formally known as *chuka soba*, began to spark an interest amongst the Japanese population. From this food shortage, the popularity of ramen spiked and had become a cultural and proper dish, gaining popularity internationally. This is not to be confused with *ramyeon*, Korean-style noodles, created by a businessman named Jeon Jung-Yung who frequently traveled to Japan for business trips. Here he adopted the rising Japanese ramen noodles and began the production of *ramyeon*.

Ramyeon was first advertised to be containing and be produced from beef, where red meat was considered a rich commodity during this era. The high poverty-stricken population of Korea in the mid-1900s were amazed at how such an inexpensive item could be produced and sold to the public. This is how the popularity of Korea's *ramyeon* began to compete against Japan's *ramen* noodles. However, both Japanese and Korean food industries rejected the popular inexpensive commodity mainly for falsely stating its nutritional value.

In 1958, Momofuku Ando's Sanshī Shokuhin, a now well-renowned industry named Nissin Foods Corporation, first released their instant ramen noodles with the brand name Chikin Ramen [5]. The introduction of this postwar instant food marketed as containing high nutritional value and benefits, was rejected by Japan's food industry who stated that these instant ramen noodles would not be successful. According to an article dating back to the year 2000, BBC News reported that the Fuji Research Institute stated that from Japanese surveys, they believe that ramen noodles are seen as their best export [6]. The Fuji Research Institute division focuses on the development of ingredients and technologies to contribute to the endorsement of new foods released to society.

In 2010, the global population consumed around 95.3 billion units of instant ramen noodles, and half of this statistic was in Eastern Asian countries specifically China, Japan, and Indonesia [5]. Taking this into account, the total global population in 2010 was roughly 6.9 billion people [7]. This value is 13.8 times higher than the population of 2010, the amount of consumption in that year alone is remarkable. This standalone statistic is incredible, but if we relate this statistic to the previous cardiovascular disease statistic, how many people in this population were diagnosed with a cardiovascular disease while having this instant food play a factor in their diet? Although compared to many instant foods, instant ramen noodles do not contain heavy artificial food preservatives. They do, however, contain high amounts of sodium, as many other instant and frozen products. We previously stated that high contents of sodium in a person's diet played a big factor in their health, developing conditions and diseases affecting their health. With multiple studies conducted on the processing and nutritional analysis of these foods containing artificial ingredients and preservatives, how can we properly detect their true nutritional value hidden away in its chemical composition?

Modern techniques in nutritional analysis, specifically food composition analysis through food composition databases (FCD), provide quantitative chemical analysis for most foods and beverages [8]. This practice derives data collected from food samples containing macronutrients, or macromolecules, such as carbohydrates, lipoids, and proteins. With various types of macromolecules, there are multiple different methods to identify them. A common studied characteristic of many foods is their moisture content, and by utilizing infrared spectroscopy (IR) we can determine the amount of moisture. By studying the absorption factor of food, the atomic absorption spectrometer (AAS) utilizes the physical properties of free atoms that contribute to the amount of absorption. With many other surpluses of methods which additionally characterize absorption, moisture, extraction, and more, these methods tend to always utilize the sample entirely which results in the possible destruction of the sample. This destruction or use of chemicals does not always pertain to the process of analytics in foods, but also pertains to the agricultural studies in plants. With the introduction of Raman spectroscopy, not to be confused with ramen noodles, this idealizes the protection of the sample from physical and chemical means through non-destructive and chemical-free analysis.

Raman spectroscopy (RS) is a surfacing analytical technique with the potential to overcome impending issues by providing near accuracy in its molecular vibrational analysis. This non-destructive and chemical-free instrument gathers the vibrational frequencies to produce what we call the mode of frequency, the average taken from multiple frequencies at a given wavelength. This emerging analytical technique was first discovered by Dr. CV Raman dating back to the year 1928 but wasn't practical until the invention of the laser in the 1960s [9]. The early theory of Raman spectroscopy included the use of a laser that is dispersed by the instrument where the light is then refracted from an internal mirror. This refraction then proceeds through a filter and then released out of the microscope and into the sample. This resulting scattered light is then gone back into the microscope and then filtered for any stray scattered light known as Rayleigh scattering. The filtered light then proceeds into a spectroscope for spectral analysis which is then detected by a detector which we now use charged coupled devices (CCDs) to complete the process.

One of its best-known features is the ability to perform at a faster pace when compared to other spectroscopic methods, and additionally it requires minimal sample handling and preparation [10]. Which brings us to the question of how exactly RS differs when compared to other forms of spectroscopy. The key factor in distinguishing between spectroscopies is that RS utilizes the laser emitted from the spectrometer to both excite the molecules in the sample and acquire the quantitative data shortly after. The quantified data is acquired through the means of scattered photons, generated by the excited vibrating molecules, where the spectrometer quantifies these Raman photons by relating their energies using a coupled form of spectroscopy. This coupling device is known as a charged coupled device (CCD), where the linear conformation of this device was used as the first application for spectroscopic methods [11]. After acquiring data from the scattered light, the spectrometer can then process and produce a statistical figure to portray the quantitative analysis. Some spectrometers may have a built-in system to automatically provide a preprocessing method, such as a linear baseline function. To enhance the spectroscopic data obtained, data preprocessing is conducted to improve the model figure produced by the instrument.

Through the utilization of data preprocessing and chemometrics, a method for statistical chemical analysis, we can produce concrete predicted results without the need to reproduce the experiment. As stated before, there are spectrometers capable of performing data preprocessing through built-in functions. The instrument that was used for the experiment contains this built-in function to automatically apply a baselining correction to yield a smoother model. The model figures, or averaged spectra, created by the instrument can then be transferred to a programming language for further analysis using chemometrics. There are thousands of citations across different fields which utilize chemometric-like data processing methods including computer science, informatics, chemoinformatic, bioinformatics, and many engineering disciplines [12]. The advantages to using this enhanced form of data analysis is that it allows us to validate the data acquired by running the data through algorithms provided by the application. Through this, the algorithm provides a prediction of accuracy by classification as if the experiment were to be conducted repeatedly. After optimizing the data to create a filtered and cleaner model, we can then produce further supportive statistical data by creating analysis of variance (ANOVA) tests. Consuming all this information on bioinformatics and spectroscopy, how does this all apply to ramen noodles?

The reason for selecting instant ramen noodles as a choice of study is to investigate a popular international food by identifying its nutritional contents. Common ingredients that make up these ramen noodles include a flour variant, oil, water, salt, and what is known as kansui. It is a tradition to use kansui, or Mongolian mineral-rich water, which is rich in the minerals sodium and potassium carbonate, this gives the ramen noodles its yellow color while also giving it a firm texture [13]. While we can be informed by the manufacturer on the composition of these noodles, we cannot determine its true nutritional value through basic methods. This is where we can utilize the use of RS coupled with chemometrics for a detailed analysis on ramen noodles. How exactly can we determine the nutritional composition of these ramen noodles without having a basis on what each ingredient will be translated through the spectrometer?

We can then utilize the data acquired from previously conducted experiments to provide a template as a form of translation for these macromolecules. By assigning peaks at certain wavenumbers, we refer to as the Raman shift, with units of reciprocal centimeters, we can establish a classification system for the different macromolecules. For example, we understand that we should obtain some form of a lipid in the ramen noodles, and this can then be identified and assigned at a peak around 1445 cm^{-1} [14]. This is how we assigned the significance of each band present in the ramen noodles, resulting in the quantification of the nutritional composition. This was the first step to the study, comparing each ramen noodle brand by its brand manufacturer and testing the potential of RS in its capability to distinguish and classify each brand. The second part of the study was testing RS by focusing solely on the peak that was assigned to be gluten, so two classes were created to participate in the study. The first being all the ramen noodle brands that were considered regular, or gluten-rich, and the second classified as organic and gluten-free ramen. Lastly, after completing these first two studies, we hypothesized that RS coupled with chemometrics could predict and classify the ramen noodles based on their origin of nationality. By doing so, we can set an example of how RS can provide an accurate analysis which can be implemented in future food production as a method for quality control.

The goal in investigating foods containing gluten using RS is to support the gluten-intolerant population in their dietary needs and knowing which could be the best to include in one's daily diet. The comparison of gluten-content and researching how certain brands and types differ in the production process, such as potentially identifying oils, fats, and sugars through RS. Through this chemical-free and non-destructive technique, we are aiming to push this technological advancement forward to help, not only on a national level, but on an international scale. This RS-focused study is to ultimately improve the quality control of foods in the overall food industry, market, public stores, and restaurants for a precise nutritional analysis.

2. METHODS

2.1 Ramen Noodles

The ramen noodles were purchased from a local international food store and other local grocery stores located in College Station, TX. Ten different ramen noodle packaging were purchased for the study, consisting of six gluten-free ramen and four gluten-containing ramen. Each ramen brand was scanned 50 times to keep the data consistent with enough scans for accuracy. After collecting using the handheld Resolve Agilent spectrometer, the data was extracted and then processed using the MATLAB program. This data is baselined by the Resolve Agilent spectrometer, then extracted to be tested with different combinations of preprocessing to achieve the best results using PLS Toolbox 9.0 (Eigenvector Research, Inc., Manson, WA). Before preprocessing the data, a few outliers consisting of one to two spectra were removed to prevent any alterations to the averaged spectra. The chosen preprocessing for the dataset included taking the 1st derivative (Sav Gol), a filtering method, and then mean centered, a scaling and centering method.

2.2 Gluten-Rich vs Gluten-Free

The same ten ramen noodle brands were utilized in this portion of the experiment with six brands classified as gluten-free, and the resulting four being the regular gluten-rich ramen noodle brands. The brands were then split into their respective groups to average the first group of gluten-rich together, and the second group as gluten-free. The gluten-rich group consisted of four of the ramen noodle brands: Buldak, Gomtang, Nissin, and Nongshim. The gluten-free group consisted of six of the ramen noodle brands: Annie Chun's, Dr. McDougall's, HakuBaku, Koyo, Lotus Foods, and Mike's. The same number of spectra was used in the data analysis, 50 spectra for each brand, while removing the few outliers in the data set. Like the first part, the same programming language was used along with the same preprocessing. With the exception of normalizing a certain peak to reduce the prediction bias from the prediction algorithm.

2.3 Origin of Nationality

The same data set from the first part of the experiment was used to maintain the consistency of the data. In this study, the ten ramen brands were split into their respective groups based on where they were manufactured. The ramen brands could be classified respectively as: Annie Chun's (Korea), Buldak (Korea), Dr. McDougall's (USA), Gomtang (Korea), HakuBaku (Australia), Koyo (USA), Lotus Foods (USA), Mike's Mighty Good (USA), Nissin (Japan), & Nongshim (Korea). The same programming language was used to preprocess the data, with the same preprocessing methods as before. An image of all ramen brands can be found in the appendix section of the article.

2.4 Acquisition

The instrument used to collect the data discussed is the portable, hand-held Agilent Resolve Raman spectrometer equipped with an 830 nm laser. The experimental parameter for use includes an acquisition time of 1 second, a 495-mW power, and a surface scan mode-setting. The instrument has a baseline spectral subtraction through its device software, yielding processed results in a quick ~30 to 60 second scan. The data is exported as comma separated value (CSV) files onto an excel sheet where outliers are removed, data can be classified, and then imported into the MATLAB program to perform a chemometric analysis on the data.

2.5 Statistical Analysis

The spectra are imported into the MATLAB program to be analyzed through partial least squares discriminant analysis (PLS-DA). Specifically, through the PLS Toolbox 9.0

(Eigenvector Research, Inc., Manson, WA), the imported data set is analyzed, edited, plotted, to clean up and omit any residing outliers or poor data that could alter our results. Along with being run through PLS-DA, the model optimizer function is run to compare the different combination of preprocessing methods using the latent variables (LV) obtained. The preprocessing method is chosen by comparing the cross-validation (CV) results for each LV, usually ranging from 1-7 for the best accuracy. This step consists of trial and error to yield the best results possible. Cross-validation is the program's bias, or ability to classify, its predicted data results.

3. RESULTS & DISCUSSION

3.1 Differentiation of Ramen Brands

Raman spectra were collected from 10 different ramen brands, these noodles were scanned dry and fresh out the package. The 10 different brands can be seen in Figure 1, where they vary in intensities at certain bands. The different brands were split into two categories, gluten-free and gluten-containing. There were 6 gluten-free ramen in this study including Annie Chun’s, Mr. McDougall’s, HakuBaku, Koyo, Lotus Foods, and Mike’s ramen. The other 4 packages were not gluten-free, Buldak, Gomtang, Nissin, and Nongshim. Keep in mind that the experiment was on the ramen noodles dry, meaning without the seasoning or any toppings the packaging may include.

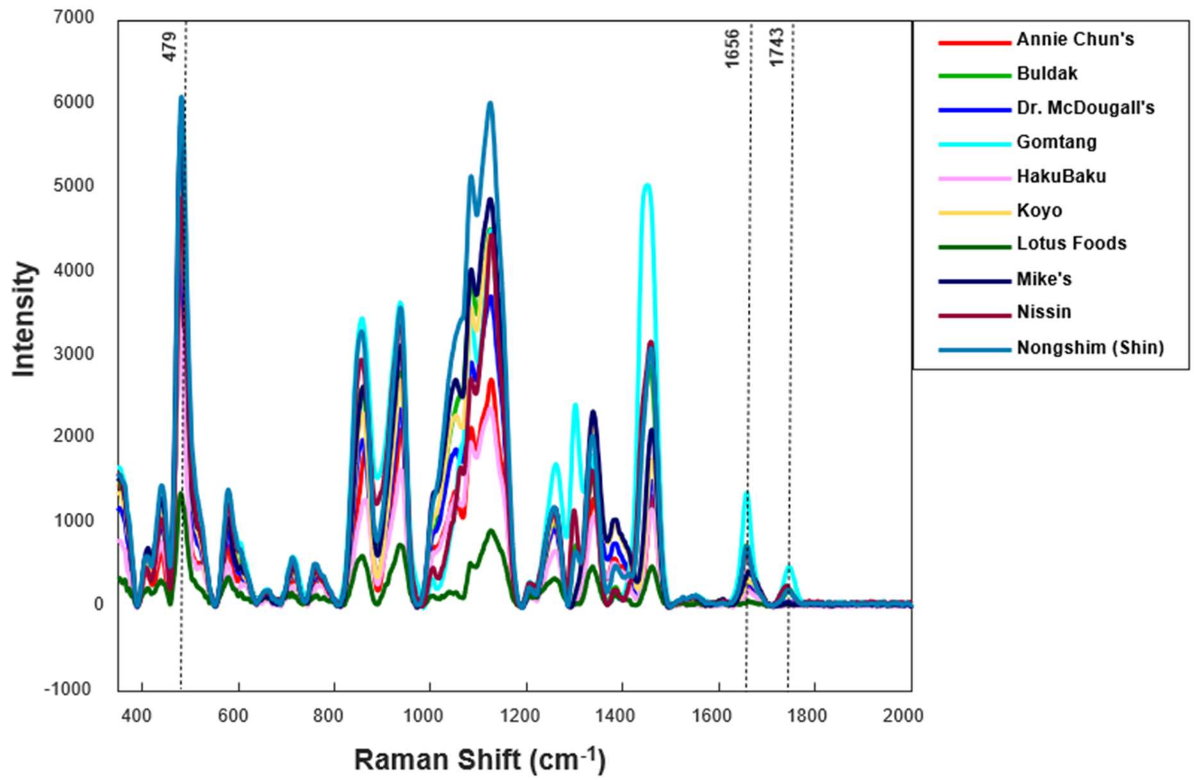

Figure 1: Averaged Raman spectra collected from the 10 different ramen noodle brands; no preprocessing.

Starting by comparing all 10 brands of ramen, we can identify the different levels of intensity between the samples. Returning to Figure 1, many of the unclassified vibrational peaks consist of a carbohydrate classification (e.g., peaks around 441, 481, 580, 714, 857, 936, 1005, 1085, 1124, 1147, 1260, 1338, and 1384 cm^{-1}). Table 1 identifies the assignment of each of these corresponding intensities, ranging from 441 to 1743 cm^{-1}. The peaks focused on this experiment correspond to the vibrational frequencies of gluten at 1005 and 1656 cm^{-1}. The averaged spectra for each of the ten brands were then preprocessed using the programming language, MATLAB. Alongside the programming language, a toolbox equipped with various analysis methods was utilized to perform a partial least squares discriminant analysis (PLS-DA).

Table 1: Vibrational band assignment for Raman spectra

Band	Vibrational mode	Assignment
1743	C=O stretching	Lipids[38]
1656	C=O stretching (amide I)	Proteins[27,39]
1604	Aromatic ring	Proteins[40]
1458	δ(CH)+δ(CH_2)	Aliphatic[41]
1445	δ(CH)+δ(CH_2)	Aliphatic[41]
1384	δ(C-O-H) - coupling of the CCH and COH deformation modes	Carbohydrates[41]
1338	δCH_2 bending vibration	Aliphatic[41]
1298	CH_2 twisting	Lipids[38]
1260	δ(C-C-H)+δ(O-C-H)+δ(C-O-H)	Carbohydrates[41,42]
1147	ν(C-O-C), ν(C-C) in glycosidic linkage, asymmetric ring breathing	Carbohydrates[43]
1124	ν (C-O)+ν(C-C)+δ(C-O-H)	Carbohydrates[41]
1085	ν(C-O)+ν(C-C)+δ(C-O-H)	Carbohydrates[41]
1062	ν (C-O-C)	Lipids
1050	ν(C-O)+ν(C-C)+δ(C-O-H)	Carbohydrates[41]
1005	Phenylalanine ring stretching mode	Proteins[39]
936	δ(C-O-C)+δ(C-O-H)+ν(C-O) α-1,4 glycosidic linkages	Carbohydrates[41]
892	CH_2 wagging	Lipids[38]
857	δ(C-C-H)+δ(C-O-C) glycosidic bond; anomeric region	Carbohydrates[41]
714	δ(C-C-O) related to glycosidic ring skeletal deformations	Carbohydrates[41]
580	δ(C-C-O)+τ(C-O)	Carbohydrates[41]
481	CCO and CCC deformations; Related to glycosidic ring skeletal deformations δ(C-C-C)+τ(C-O) Scissoring of C-C-C and out-of-plane bending of C-O	Carbohydrates[41]
441	Skeletal modes of pyranose ring	Carbohydrates[41]

The spectral data that was first analyzed can be looked back at Figure 1, containing no preprocessing. Taking these averaged spectra of all ten of the ramen noodle brands, the data was preprocessed by taking the first derivative of the spectra with a polynomial order of 2, and a derivative order of 1 through a Savitzky-Golay smoothing filter (1st derivative Sav-Gol). The second preprocessing done was mean centering, an autoscaling and variable centering calibration. Lastly for the ANOVA and prediction, an area normalization at the 1338 cm^{-1} peak was performed to improve the initial resulting data to improve the predicted classification results.

A cross-validation table is created from the resulting data that was preprocessed at a LV of 11 which allows us to depict the accuracy for the prediction model. In table 2, we can first identify the model's ability to classify each ramen brand with an accuracy of exactly 90.2%. In other words, RS is capable of accurately predicting the different ramen noodle brands from their nutritional composition. The model suggests that RS was able to successfully identify 5 of the ramen noodle brands with 100% accuracy, including the brands Annie Chun's, HakuBaku, Lotus Foods, Mike's, and Nissin. Taking a closer look, the model suggests that RS successfully classified these five ramen noodle brands where each 50 spectra were correctly identified as their respective brand. The next three, Gomtang, Koyo, and Nongshim, showed high accuracy but did not represent a 100% accuracy. Unfortunately, the model predicts that Buldak and Mr. McDougall's to be the least accurate amongst the data set. The percentage accuracy in the model depicts how well the classification performed was by RS. A high percentage would mean that RS is successful in classifying which spectra from the data set corresponds to its respective ramen noodle brand. A low accuracy suggests that RS misclassified the spectra to be other ramen noodle brands, most likely due to their similarity in their nutritional composition.

Table 2: Classification table of ramen brands from the cross-validation model created by PLS-DA

	Accuracy (%)	Annie Chun's	Buldak	Dr. McDougall's	Gomtang	HakuBaku	Koyo	Lotus Foods	Mike's	Nissin	Nongshim
Annie Chun's	100	50	0	1	0	0	1	0	0	0	0
Buldak	68	0	32	1	4	0	0	0	0	0	1
Dr. McDougall's	64	0	0	32	0	0	6	0	0	0	1
Gomtang	92	0	8	0	47	0	0	0	0	0	2
HakuBaku	100	0	0	0	0	50	0	0	0	0	0
Koyo	86	0	1	12	0	0	43	0	0	0	0
Lotus Foods	100	0	1	2	0	0	0	43	0	0	0
Mike's	100	0	0	1	0	0	0	0	50	0	0
Nissin	100	0	0	0	0	0	0	0	0	50	0
Nongshim	92	0	5	1	0	0	0	0	0	0	46

To support this cross-validation model created from the PLS-DA, we created ANOVA tables using the Kruskal-Wallis testing method. Looking at Figure 2, this first figure depicts the difference in intensity for one of the carbohydrate peaks at 479 cm^{-1}. The carbohydrate intensity varied between the different ramen noodle bands, all containing different values. The figure suggests that Lotus Foods contains the lowest amount of carbohydrates at this certain brand, and interesting enough the second highest carbohydrate-containing brand is the gluten-free Mike's brand. This data result suggests that the ramen noodle brand's carbohydrate content has no relation to whether the noodle itself is considered gluten-rich or gluten-free. The second figure, Figure 2, identifies the vast difference in gluten content between the ten ramen noodle brands. The figure suggests that the ramen brand Gomtang contains the highest amount of gluten while the Buldak brand is the second highest. It also appears that Dr. McDougall's contained the lowest amount of gluten content. Analyzing the ANOVA figures, the horizontal line represents the confidence interval for the spectra with a meaning of all the spectra which can be represented as a circle. The last ANOVA figure, Figure 2, represents the amount of lipid intensity at 1730 cm^{-1}, here we identify that again Gomtang contains the highest lipid content with Dr. McDougall's contains the lowest lipid content. Overall, the results from this first portion of the experiment shows that RS is successful in identifying the ramen noodle brands by correctly classifying them with their respective spectra. It is also capable of distinguishing the brands at their different peak intensities, ranking them based on their level of content.

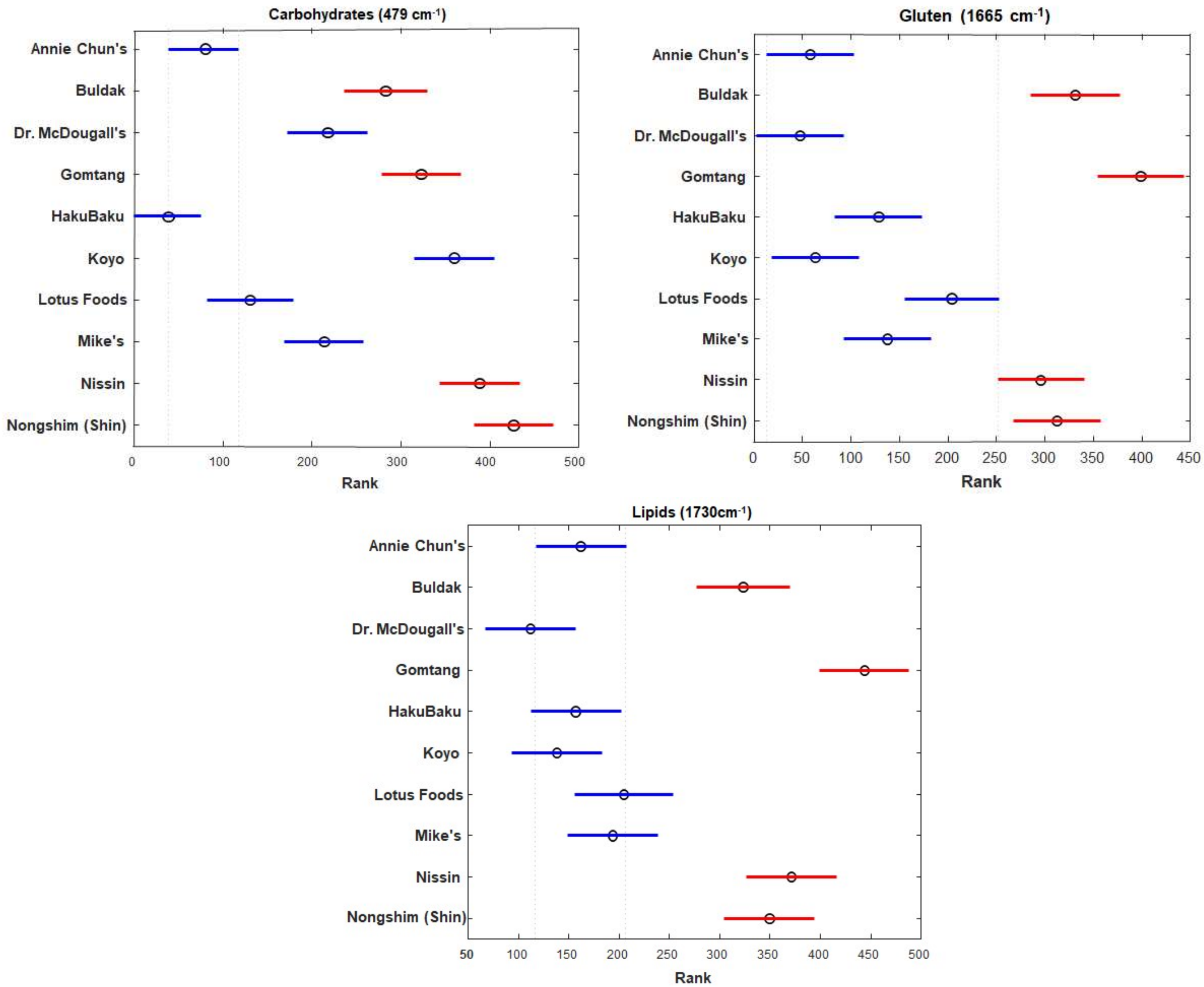

Figure 2: ANOVA for the ramen brands at the three peaks of interest

3.2 Gluten-rich vs Gluten-free

By focusing on these peaks, we can visually classify the different ramen noodle brands where the gluten-free noodles had either low intensity or no intensity at the corresponding Raman shift (cm^{-1}). The surprising results were seeing these intensities at these gluten peaks for the gluten-free ramen. We expected to find no intensity at these peaks considering the ramen noodle packaging label. However, through Raman spectroscopy we can determine the amount of gluten these noodles contain and acknowledge the instrument's capability to compare this quantitative data with similar food types or even different food types. To further support this collected data, we performed a quantitative analysis on the dataset. Using MATLAB, with an incorporated Toolbox, we can perform various chemometric functions to accurately depict our results.

The first set of data consisted of categorizing the ten different ramen noodle brands. In doing so, we compared the nutritional composition of each individual ramen noodle brand amongst each other to obtain the needed values to identify and assign. After performing the data analysis on all ten brands to obtain the peaks of interest, we identified two peaks of interest at 1005 & 1656 cm^{-1} which correspond with gluten. However, the peak at 1656 cm^{-1} will be used to focus on the ramen's gluten content since it can be easily identified. Knowing this, the ten ramen noodle brands were then split up into two groups, a gluten-rich ramen noodle group and a gluten-free group. The results can be portrayed by Figure 3 where the green line suggests the gluten-free averaged spectra from the six gluten-free ramen noodle brands while the red line represents the four gluten-rich ramen noodles. The figure depicts that the peak at around 1743 cm^{-1}, representing the lipid content in the noodles, nearly shows no intensity for the gluten-free noodles. The peak of interest at 1656 cm^{-1} shows the vast difference in intensity between the gluten-rich noodles and the gluten-free noodles. ANOVA were also created for each of these peaks but were not as impactful since the vast differences between the gluten-rich noodles and gluten-free noodles are noticeable.

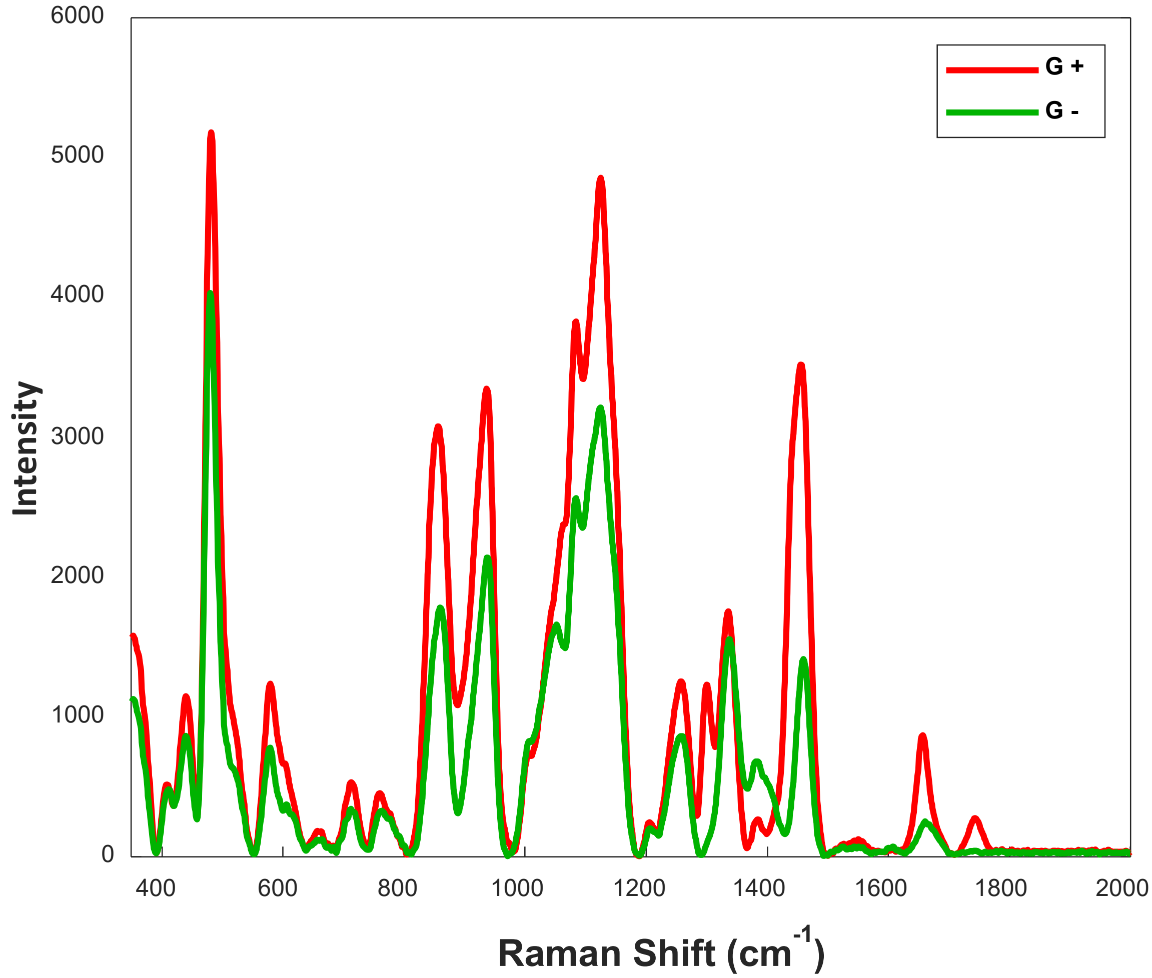

Figure 3: Averaged Raman spectra of gluten-free ramen compared to gluten-rich ramen.

To end this second part of the experiment, a classification table was created for these two groups to depict the accuracy of RS. Looking at Table 3, RS was nearly 100% accurate in determining which spectra corresponded to their respective gluten-rich or gluten-free brand. It shows that one spectrum was misclassified as gluten-rich when it should be classified as gluten-free. This could be due to an outlier that was not removed due to human error, resulting in a near 100% accuracy by the prediction model. To summarize these results, RS was successful in determining which ramen noodles were either gluten-free or gluten-rich.

Table 3: Classification table of Gluten-content in ramen from the cross-validation model created by PLS-DA.

	Accuracy (%)	Gluten-rich	Gluten-free
Gluten-rich	100	198	1
Gluten-free	99	0	292

3.3 Origin of Nationality

The third part of the overall experiment was out of curiosity whether RS coupled with chemometrics could identify the relationship between the ramen noodle brands and their origin of nationality. This one was interesting because each country has their own way of manufacturing ramen noodles, the process may be similar, but the ingredients vary. If RS could successfully identify the ramen noodles based on their origin of nationality, we wanted to see which country contained the highest and lowest content from the peaks of interest. This is where the same data set was taken but split into four different groups based on their origin. The results can be seen in Figure 4 where Japan and Korea both contained higher levels of carbohydrate, lipids, and gluten content. Although the results are interesting, these may not be accurately depicted on a wider scale of ramen brands. In this study, two of our ramen brands originated from Australia and the USA. Meaning that as the figure depicts, these two nations produce the healthier ramen noodle brands.

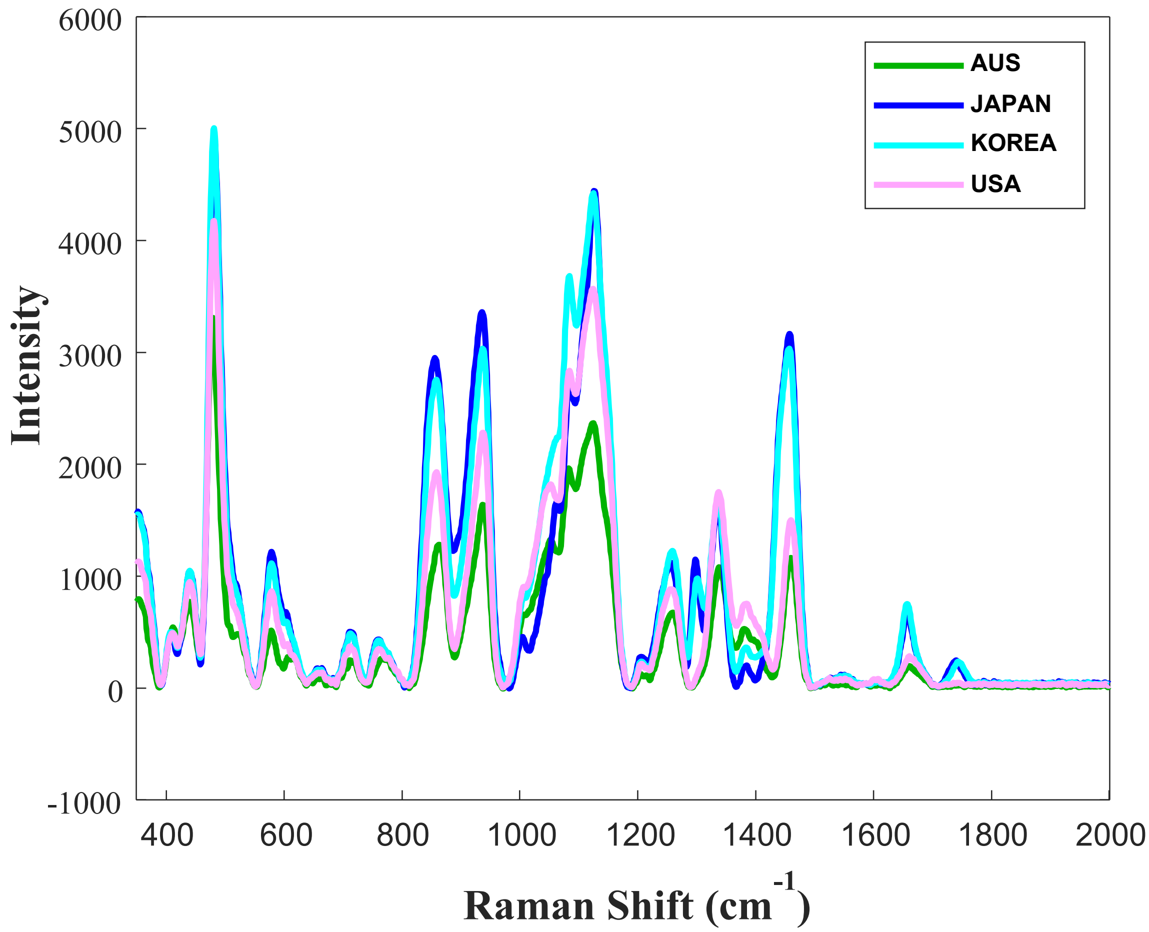

Figure 4: Averaged Raman spectra comparing the ramen brands' nationality.

The classification table for the processed data resulted in RS successfully classifying each ramen brand to its respective origin of nationality, as seen in Table 4. The resulting percentage for Australia and Japan were an exact representation of 100% classification. However, for Korea and the USA, we obtained values of 98% and 99% respectively. This misclassification could be interpreted as the Korean gluten-free ramen brand being classified as a gluten-free ramen brand from the USA. Although the results aren't a perfect representation, it's still impressive at how accurate RS is in classifying these ramen brands based on their nationality.

Table 4: Classification table of nationality in ramen from the cross-validation model created by PLS-DA.

	Accuracy (%)	Australia	Japan	Korea	USA
Australia	100	50	0	0	0
Japan	100	0	50	0	0
Korea	98	0	0	195	1
USA	99	0	0	3	192

According to the ANOVA test conducted for the different ramen noodle nationalities, in Figure 5, Japan is considered to contain higher levels of gluten compared to its high content partner, Korea. However, from the previous studies, Korean ramen brands Gomtang and Nongshim contained higher levels of gluten, lipid, and carbohydrate content. This could be interpreted as a change in the results compared to the previous data analyzed. After taking into consideration the third Korean ramen noodle brand, Annie Chun's, we can conclude that this gluten-free organic ramen is the sole reason why the averaged spectrum for Korea is ranked lower than the Japanese ramen data.

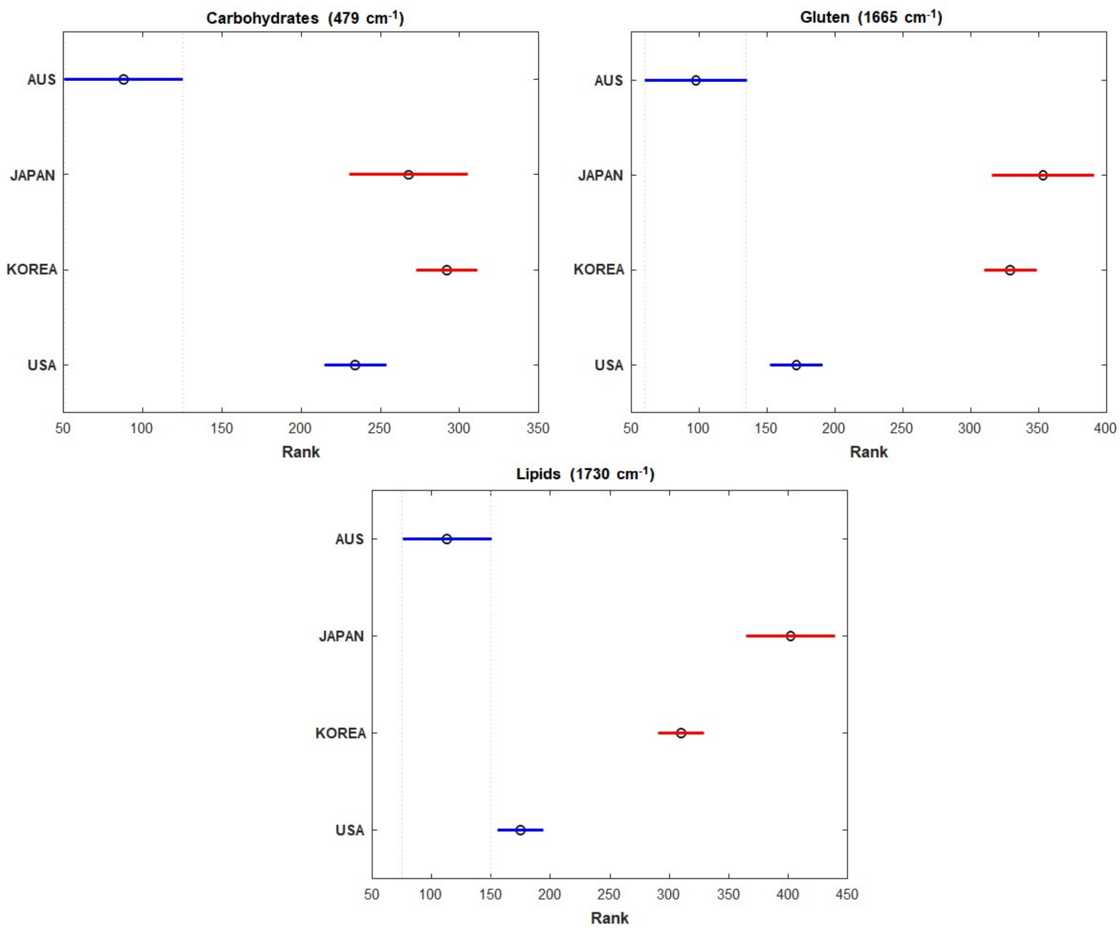

Figure 5: ANOVA for the three peaks of interest between the different nationalities.

To resolve this bias and prevent any future bias in upcoming experiments, obtaining equal amounts of ramen noodle brands from each country and equal amounts of brands both gluten-free and gluten-rich would fix this issue. However, overall, this part of the experiment was successful in testing the potential of RS in identifying the ramen noodle brands' origin of nationality.

4. CONCLUSION

Raman spectroscopy being an emerging analytical technique, the initial concept of this study was to test the potential of Raman spectroscopy and its ability to characterize macromolecules in the ramen noodles. The first round of data analysis that was conducted resulted in the understanding of the instant ramen noodles' composition. Utilizing the data, we were able to characterize the different types of macromolecules present in the ramen noodles, and even quantify the amount in each of the ramen noodles. The three key macromolecules we were able to identify consisted of starch, fats, and gluten. The compelling peaks that can be relayed back to in Figure A.2 portray an example of the three important regions of interest.

The first peak that caught our interest was what we portrayed as a fats peak. This corresponds with a variety of different fatty acids, but triacylglycerols was the main peak of interest for this group. The starch peak consists of many carbohydrate structures which can be identified in multiple parts of the spectra. Ramen noodles being a carbohydrate heavy food confirms that the spectra should portray the high amounts of these macromolecules. However, the ramen noodle brands being labeled as either gluten-rich or gluten-free did not play a factor when it came down to the carbohydrate intensity. Both gluten-free and gluten-rich ramen noodles varied in intensities. Lastly the peak we corresponded to gluten was initially a mystery peak which we predicted to be a protein or some form of lipid or fat. This was later confirmed by conducting an analytical test on the raw form of gluten to obtain a spectrum that we could associate this unidentified peak with.

The gluten spectra collected matched the wavenumber, around 1650 cm^{-1}, and was successfully identified as gluten. After this result, we determined that since gluten was able to be identified using RS, that we could conduct an additional study on its gluten-free content. This is where we started the studies on gluten-free ramen to then compare the resulting spectra to gluten-rich ramen. The data suggested that some gluten-free ramen could be identified as gluten-rich ramen. This is the result we were hoping to find, as the ramen noodles that were chosen for the study were marketed to be organic and gluten-free. However, the study suggests that some of the gluten-free ramen does in fact contain some form of gluten, or at least a relatively low level of this protein.

Taking the data collected into consideration, we could determine that RS can yield accurate and substantial results in such a short period of time. However, with its capability of identifying these macromolecules, we decided to compare these values to find similarities and differences between the different brands of ramen noodles. After conducting the study on whether RS could distinguish between ramen noodle brands, we can infer that RS can successfully identify these different brands with a near accuracy of 100%. Through this we gathered the gluten-rich ramen and labeled this as the first group to be compared between each other. The second group consisted of the gluten-free ramen we obtained and analyzed it to also be compared amongst each other. The last group was a combination of these two groups to ultimately determine the similarities and differences of all ramen brands based on their gluten content. Resulting in RS identifying which ramen brand was gluten-rich and which was gluten-free with again a near accuracy of 100%.

Lastly, we hypothesized that after all this analysis, could RS be able to differentiate between the brand’s national origin? Although the ramen we had at hand was not of equal value, meaning there was more ramen of one nationality than the others, we decided to still put this to the test and work with the data obtained from the study. Again, RS proved to us that it's capable of accurately determining the difference in nationality between the ramen noodle brands. To clarify this, each food manufacturer tends to have similarities and differences when it comes down to the ingredients used to produce the products or the process used. From the data collected, we determined that Korean-based ramen noodle brands would yield the highest fat-containing products. Seeing that each country has their own methods to produce and obtain their ingredients, they all share similar ingredients but could be produced in different amounts and variations.

We aim to further support this data through additional research conducted on other carbohydrate-heavy foods which can also be partnered with their gluten-free variant. Also, in hope for the data collected and analyzed to be utilized as a supportive analysis to any future research conducted using RS in related experiments. The study of different types of bread ranging from store bought bread and comparing between white, wheat, and whole-grain bread to list a few examples. Additionally, a study on international types of bread ranging from French baguettes, to pita bread, to sourdough bread. Another possibility would be to conduct a study on crackers and cookies, which can also enhance the study on store bought food. Perhaps in the future this study will be reproduced but differ by adding onto the amount of ramen noodle brands initially studied.

Ultimately, the overall results of the experiment depict that by using RS, we can successfully determine the nutritional value, its respective brand, and origin of nationality of a popular international food. A small steppingstone that can potentially result in the use of RS in other production of foods. Where these foods can be labeled with near 100% accuracy in their nutritional composition using an analytical spectroscopic technique known as RS. These more accurate labels can then benefit those with underlying cardiovascular diseases or hope to follow a healthier lifestyle. With this study we hoped to achieve the promotion of using this emerging analytical technique and present its capabilities through non-invasive, non-destructive, and chemical-free methods.

REFERENCES

[1] The GBD 2015 Collaborators, “Health Effects of Overweight and Obesity in 195 Countries over 25 Years”. *New England Journal of Medicine*, vol. 377, no. 1, New England Journal of Medicine, 2017, pp. 13–27, doi:10.1056/nejmoa1614362.

[2] Nummer, Brian A., and A. Brian. "Historical Origins of Food Preservation." *National Center for Home Food Preservation. University of Illinois Extension*, 2002.

[3] Leistner, L. “Basic Aspects of Food Preservation by Hurdle Technology.” *International Journal of Food Microbiology,* vol. 55,1-3, 2000, pp.181-186, doi:10.1016/s0168-1605(00)00161-6.

[4] Kyungkoo H., "Noodle Odyssey: East Asia and Beyond." *Korea Journal*, vol. 50, no. 1,2010, pp. 60-84, doi: 10.25024/kj.2010.20.1.60.

[5] Solt G. "Shifting Perceptions of Instant Ramen in Japan During the High-growth Era, 1958-1973." *International Journal of Asia-Pacific Studies* vol. 8(2), pp. 13-31, 2012.

[6] Haub C. “2010 World Population Data Sheet.” *PRB*, 2010.

[7] “Asia-Pacific | Japan Votes Noodle the Tops.” *BBC News*, BBC, 2000.

[8] Md Noh, M. F., et al. “Recent Techniques in Nutrient Analysis for Food Composition Database”. *Molecules*, vol. 25, no. 19, Molecules, 2020, p. 4567, doi:10.3390/molecules25194567.

[9] Yohk. “History of Raman Spectroscopy.” *Nanophoton*, 15 Oct. 2020.

[10] Smith E., Dent G. "Modern Raman Spectroscopy: A Practical Approach." Wiley, 2, pp.15-20. 2004, doi:10.1002/0470011831.

[11] Sweedler, J.V., Jalkian R.D., and Denton M.B. "A Linear Charge-coupled Device Detector System for Spectroscopy." *Applied spectroscopy* 43.6, 1989, pp. 953-962, doi: 10.1080/05704920701551530.

[12] Lavine, B.and J. Workman. “Chemometrics”. *Analytical Chemistry*, vol. 80, no. 12,

Analytical Chemistry, 2008, pp. 4519–31, doi:10.1021/ac800728t.

[13] Teacher I. “Ramen Noodle Ingredients – Kansui, etc.” *Yamato Noodle*, 2019.

[14] Talari, A. C. S., et al. “Raman Spectroscopy of Biological Tissues”. *Applied Spectroscopy Reviews*, vol. 50, no. 1, Applied Spectroscopy Reviews, 2015, pp. 46–111, doi:10.1080/05704928.2014.923902.

[15] Stover, P. J., King, J. C., "More Nutrition Precision, Better Decisions for the Health of Our Nation". *The Journal of Nutrition*, 2020, vol. 150 (12), pp. 3058-3060, doi:10.1093/jn/nxaa280.

[16] Lancelot E., Fontaine J., et. al., "Study of Structural Changes of Gluten Proteins During Bread Dough Mixing by Raman Spectroscopy". *Food Chemistry*, vol. 358, 2021, doi: 10.1016/j.foodchem.2021.129916.

[17] Jamieson, L. E., et al., “Ratiometric Analysis Using Raman Spectroscopy as a Powerful Predictor of Structural Properties of Fatty Acids”. *Royal Society Open Science*, vol. 5, no. 12, Royal Society Open Science, 2018, doi:10.1098/rsos.181483.

[18] Food and Agriculture Organization of the United Nations. "How to Feed the World 2050." *Rome: Food and Agriculture Organization*, 2009.

[19] Awad T.S., Moharram H.A., et. al., "Applications of Ultrasound in Analysis, Processing, and Quality Control of Food: A Review". *Food Research International*, vol. 48, 2012, pp. 410-427, doi: 10.1016/j.foodres.2012.05.004.

[20] “World Instant Noodles Association.” *世界ラーメン協会*, 2022.

APPENDIX: SUPPORTING FIGURES

Figure A.1: Images of the ten ramen brands that were analyzed.

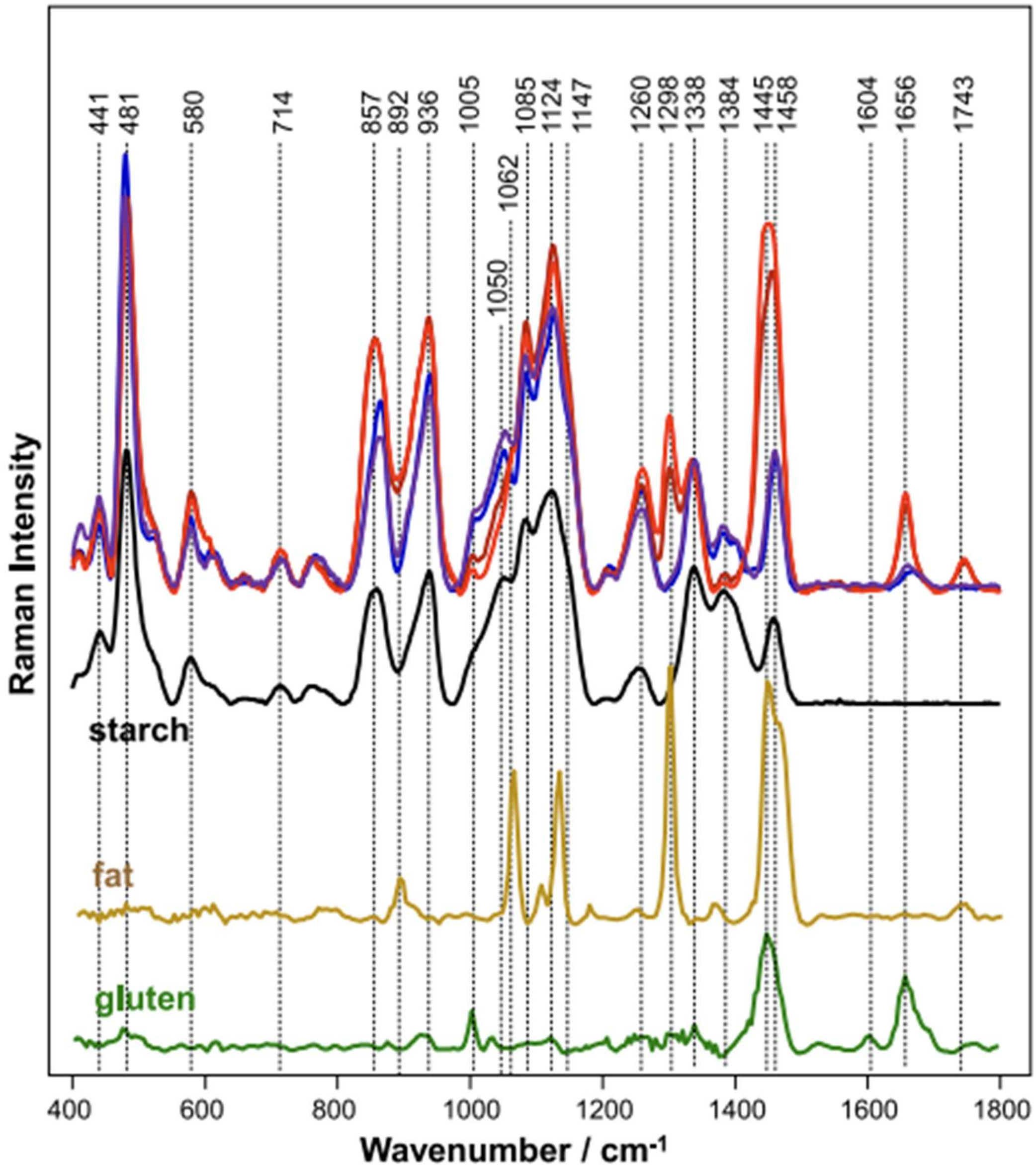

Figure A.2: Raman spectra of gluten-rich (Gomtang, red and Paldo, maroon) and gluten-free (Koyo, blue, and HakuBaku, purple). Raman spectra of starch (black), glyceral tristearate (gold), and gluten (green).

.