March 2022

Analysis of Plant and Animal Proteins Using Raman Spectroscopy

Noopur Bapardekar

University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/masters_theses_2

Part of the Food Chemistry Commons

Recommended Citation

https://doi.org/10.7275/26881422 https://scholarworks.umass.edu/masters_theses_2/1149

This Open Access Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
ANALYSIS OF PLANT AND ANIMAL PROTEINS USING RAMAN SPECTROSCOPY

A Dissertation Presented

by

NOOPUR BAPARDEKAR

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

February 2022

Department of Food Science
ANALYSIS OF PLANT AND ANIMAL PROTEINS USING RAMAN SPECTROSCOPY

A Dissertation Presented

by

NOOPUR BAPARDEKAR

Approved as to style and content by:

_________________________________________

Lili He, Chair

_________________________________________

David Julian McClements, Member

_________________________________________

Amanda Kinchla, Member

_________________________________________

Lynne McLandsborough, Department Head

Food Science
ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Lili He, for her relentless support and helping me throughout my journey at UMass Amherst. She has been a constant source of motivation for me and I really appreciate the structured guidance provided to me in the course of my Master’s degree. Another notable mention would be Prof Matthew Steffens who helped me pursue my quest for animal and plant proteins in the real world and gave product development insights. Without his support this thesis and my research experience would have been incomplete.

I would also like to thank my committee members Dr David Julian McClements and Prof Amanda Kinchla for their time and research recommendations. Many thanks to all my lab mates for their feedback and help, especially Janam and Andy. Their mentorship and advice over the years are gratefully acknowledged.

And last but not the least I wish to thank my parents for always believing in me and the friends I made here that made living through COVID-19 not as difficult. Without their love and encouragement, this study would never be fulfilled.
ABSTRACT

ANALYSIS OF PLANT AND ANIMAL PROTEINS USING RAMAN SPECTROSCOPY

FEBRUARY 2022

NOOPUR BAPARDEKAR, B.E., MUMBAI UNIVERSITY
M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Lili He

There has been a notable rise in the alternative protein market in the recent years which promotes an interest in the research of both animal and plant proteins to establish better structure-function relationships. Over the years many analytical tools have been used to study proteins and compare them, however Raman Spectroscopy and Surface Enhanced Raman Spectroscopy (SERS) have not been as much used for this application. SERS consolidates Raman Spectroscopy that primarily measures molecular vibrations and nanostructures that enhance the weak Raman signals. The objective of this study is to explore the capability of the Raman instrumentation in combination with different substrates for spectroscopic analysis of 3 animal proteins viz. whey, k-casein and albumin from chicken egg white and 4 plant proteins namely mung bean, soy, pea and faba bean. Herein, we firstly established a method that could be applied to all proteins to detect characteristic peaks that are related to their structure. Of all the methods, SERS with silver dendrites was the most promising method that detected protein characteristic peaks, particularly the shifts around 700-900 cm\(^{-1}\) attributed to the CN stretch and tryptophan groups. Although different proteins exhibit similar spectral characteristics, they were discriminated using principal component analysis. Then we explored the optimal method to study the effect of different environmental conditions including pH and salt concentration on the protein spectroscopic analysis. The limitations of the substrates were better understood during this process as Ag dendrites failed to provide a spectrum in the high pH range but was compatible with different salt concentrations. The peaks in the Amide-I region were
used as a marker to study the effect of change in pH and salt. Most proteins showing a shift in the band suggesting a transition from α-sheet to a random coil conformation. The acquired spectra and subsequent PCA results depicted that pea protein was the most susceptible to change in pH amongst other proteins whereas faba bean was susceptible to a change in salt concentration. Finally, these learnings were applied to analyze a real-world food product to compare its spectroscopic characteristic with the standards we have. In conclusion, we demonstrated that Raman Spectroscopy and SERS was able to provide distinct spectroscopic characteristics of plant and animal proteins that may be used to facilitate the quality control or product development of novel plant-based food products. Future work will investigate the relationship between the spectroscopic characteristics and the structural function of proteins.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS .............................................................................................................. iv
ABSTRACT ............................................................................................................................... v
LIST OF TABLES ....................................................................................................................... ix
LIST OF FIGURES ..................................................................................................................... x
LIST OF ABBREVIATIONS ....................................................................................................... xii
CHAPTER 1 ................................................................................................................................. 1
  1. Introduction ......................................................................................................................... 1
  1.1 Animal proteins and their functional properties ................................................................. 1
  1.2 Need for Animal Protein Replacement and its probable alternatives ................................. 2
  1.3 Plant Protein as a Potential Alternative ............................................................................. 5
  1.4 Molecular Spectroscopy Methods for Protein Analysis ....................................................... 6
  1.6 Goals and Objectives ......................................................................................................... 10
CHAPTER 2 ............................................................................................................................... 11
  DEVELOP AND OPTIMIZE METHOD FOR PLANT AND ANIMAL PROTEIN
  SPECTROSCOPIC ANALYSIS USING NORMAL RAMAN AND SERS .......................... 11
    2.1 Introduction ...................................................................................................................... 11
    2.1.1 Characteristic spectrum of proteins ........................................................................... 11
    2.1.2 Noble metal substrates ............................................................................................. 12
    2.1.3 Objectives of this study .............................................................................................. 12
    2.2 Materials and methods .................................................................................................... 13
    2.2.1 Materials .................................................................................................................. 13
    2.2.2 Preparation of Ag dendrites ..................................................................................... 13
    2.2.3 Sample Preparation ................................................................................................. 13
    2.2.4 Raman Instrument and Data Analysis ...................................................................... 14
    2.3 Result and Discussion ....................................................................................................... 15
    2.3.1 Characterization of Raman and SERS with Ag substrates spectra of whey and mung
         protein .............................................................................................................................. 15
    2.3.2 SERS Analysis of protein structure using silver dendrites .......................................... 17
    2.3.3 Principal Component Analysis of Proteins ................................................................. 20
    2.4 Conclusion ....................................................................................................................... 21
CHAPTER 3 ................................................................................................................................ 22
  STUDY EFFECT OF CHANGE IN PH ON SPECTROSCOPIC
  CHARACTERIZATION OF ANIMAL AND PLANT PROTEINS USING
  RAMAN SPECTROSCOPY ................................................................................................. 22
  3.1 Introduction ....................................................................................................................... 22
  3.1.1 Effect of change in pH on protein ................................................................................. 22
  3.1.2 Objectives of the study ............................................................................................... 22
  3.2 Sample Preparation .......................................................................................................... 23
  3.2.1 Raman Instrument and Data Analysis ....................................................................... 23
  3.3 Result and Discussion ....................................................................................................... 26
  3.3.1 Characterization of Raman and SERS with Ag substrates spectra of whey and mung
         protein at pH 9 ............................................................................................................... 26
  3.3.2 Normal Raman analysis to study the effect of change in pH on structure of proteins .... 28
  3.3.3 Principal Component analysis of proteins at pH 3, 6 and 9 ......................................... 31
  3.4 Secondary Structure Analysis .......................................................................................... 33
  3.5 Conclusion ....................................................................................................................... 35
CHAPTER 4

STUDY EFFECT OF CHANGE IN SALT CONCENTRATION ON SPECTROSCOPIC CHARACTERIZATION OF ANIMAL AND PLANT PROTEINS USING RAMAN SPECTROSCOPY

4.1 Introduction ........................................................................................................... 36
4.1.1 Effect of change in salt concentration on protein .............................................. 36
4.1.2 Objectives of the study ..................................................................................... 36
4.2 Materials and methods ......................................................................................... 37
4.2.1 Materials .......................................................................................................... 37
4.2.2 Sample Preparation ......................................................................................... 37
4.2.3 Raman Instrument and Data Analysis ............................................................... 38
4.3 Result and discussion ........................................................................................... 39
4.3.1 Characterization of Raman and SERS with Ag substrates spectra of whey and mung protein at different salt concentrations ...................................................... 39
4.3.2 SERS analysis to study the effect of change in salt concentration on structure of proteins .............................................................................................................. 42
4.3.3 Principal Component analysis of proteins at ionic strength 0.2, 0.5 and 0.9 .......... 44
4.4 Conclusion ............................................................................................................. 46

CHAPTER 5

5.1.1 Proteins in the Food Matrix ............................................................................. 48
5.1.2 Objectives of this study ................................................................................... 49
5.2 Materials and methods ......................................................................................... 49
5.2.1 Materials .......................................................................................................... 49
5.2.2 Sample Preparation ......................................................................................... 49
5.2.4 Raman Instrument and Data Analysis ............................................................... 51
5.3 Result and Discussion ........................................................................................... 52
5.3.1 Comparison of Just Egg MPI and egg albumin ............................................... 52
5.3.2 Principal Component Analysis of Proteins ....................................................... 53
5.4 Conclusion ............................................................................................................. 54

CHAPTER 6

CONCLUSION .................................................................................................................. 55
REFERENCES .................................................................................................................. 56
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1. Typical Wavenumbers of Raman Bands &amp; General Assignments in Raman Protein Spectra</td>
<td>8</td>
</tr>
<tr>
<td>Table 2. Summary of Raman bands characterized by SERS using Ag dendrites</td>
<td>19</td>
</tr>
<tr>
<td>Table 3. Principal Amide I Frequencies Characteristic of Protein Secondary Structures</td>
<td>33</td>
</tr>
<tr>
<td>Table 4. Percentages of protein secondary structure of proteins at pH 3, pH 6 and pH 9</td>
<td>34</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Schematic illustration of the sample preparation method</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>Representative spectra of WPI with Normal Raman, AgNPs background, WPI with AgNPs, Ag dendrites background and WPI with Ag dendrites</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>Representative spectra of MPI with Normal Raman, AgNPs background, MPI with AgNPs, Ag dendrites background and MPI with Ag dendrites</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>Schematic illustration of the sample preparation method</td>
<td>24</td>
</tr>
<tr>
<td>7</td>
<td>Representative spectra of WPI with Normal Raman, Ag dendrites background and WPI with Ag dendrites at pH 9</td>
<td>26</td>
</tr>
<tr>
<td>8</td>
<td>Representative spectra of MPI with Normal Raman, Ag dendrites background and MPI with Ag dendrites at pH 9</td>
<td>27</td>
</tr>
<tr>
<td>9</td>
<td>Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 3</td>
<td>29</td>
</tr>
<tr>
<td>10</td>
<td>Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 6</td>
<td>29</td>
</tr>
<tr>
<td>11</td>
<td>Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 9</td>
<td>30</td>
</tr>
<tr>
<td>12</td>
<td>Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI and faba bean protein isolate at pH 3</td>
<td>31</td>
</tr>
<tr>
<td>13</td>
<td>Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 6</td>
<td>32</td>
</tr>
<tr>
<td>14</td>
<td>Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 9</td>
<td>32</td>
</tr>
<tr>
<td>15</td>
<td>Schematic illustration of the sample preparation method</td>
<td>38</td>
</tr>
<tr>
<td>16</td>
<td>Representative spectra of WPI with Normal Raman, Ag dendrites background and WPI with Ag dendrites at ionic strength 0.9</td>
<td>40</td>
</tr>
</tbody>
</table>

x
Figure 17. Representative spectra of MPI with Normal Raman, Ag dendrites background and MPI with Ag dendrites at ionic strength 0.9 .................................................................41

Figure 18. Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at I=0.2 ........................................................................................................42

Figure 19. Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at I=0.9 ........................................................................................................43

Figure 20. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate ionic strength 0.2 ........................................45

Figure 21. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate ionic strength 0.5 ..............................................45

Figure 22. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate ionic strength 0.9 ..............................................46

Figure 24. Schematic illustration of the sample preparation method ........................................50

Figure 25. Representative spectra of Egg Albumin, MPI and extracted MPI from Just Egg ..........52

Figure 26. Principal component scores 3D Display model of Egg Albumin, MPI and extracted MPI from Just Egg .................................................................53
WHC - Water holding capacity
FAO - Food and Agricultural Organization
MetS - Metabolic Syndrome
WHO - World Health Organization
CD - Circular Dichroism
FS - Fluorescence Spectroscopy
NMR - Nuclear Magnetic Resonance
IR - Infrared Spectroscopy
NIR - Near Infrared Spectroscopy
FTIR - Fourier Transform Infrared Spectroscopy
PCA - principal component analysis
SERS - Surface-enhanced Raman spectroscopy
localized surface plasmon resonance (LSPR)
Trp - Tryptophan
Tyr - Tyrosine
AgNPs - silver nanoparticles
Ag dends - silver dendrites
WPI - whey protein isolate
SPI - soy protein isolate
MPI - mung protein isolate
PPI - pea protein isolate
SNV - standard normal variate
CHAPTER 1

Introduction

1. Justification

Proteins play an important role in human diet. Since all vital amino acids cannot be synthesized from endogenous precursors, dietary protein acts as primary source of amino acids. (Berryman, Lieberman, Fulgoni, and Pasiakos, 2018; Boye et al., 2012; Chongtham et al., 2011; Almeida, Moreno and Carciofi, 2020). Animal protein is considered of highest quality due to its ability to provide all essential amino acids and has been an inseparable part of the human diet since a long time. However, in recent times animal protein alternatives are seen to be on the rise, especially in developed countries owing to health and environmental concerns as well as animal welfare. Food companies are investing more resources than ever to mirror the palette experience of animal-based protein for the consumers. According to the current market trend, the largest source of alternative protein viz. plant-based food remains the top choice as numbers suggest a 17% increase in sales in 2018 with a growing usage of alternative protein as a food ingredient in consumer products. The market base for alternative protein is approximately $2.2 billion when compared to the global meat market of approximately $1.7 trillion, making the growth rate of the alternative proteins marginal to the overall meat market (Bashi, McCullough, Ong, & Ramirez, 2019).

1.1. Animal proteins and their functional properties

Animal proteins like meat, seafood and dairy are considered the highest quality protein source as they consist of all the essential amino acids with a composition profile that meet the adult essential amino acid requirements as per the World Health Organization. They have a high nutritional value and account for many remarkable functional properties like emulsification, gelation, thickening and water-holding capacity (WHC) (Rickey, 2017). Animal food proteins can be classified into three categories 1) Sarcoplasmic proteins constituting mainly of enzymes 2)
Myofibrillar proteins like actin, myosin, tropomiosin, troponin, etc. and 3) Connective tissue like collagen with each playing a significant role in determining functionalism (Herrero, 2008). Myofibrillar proteins determine most of the functional properties and texture of the muscle food, fish and meat. Collagen contributes significantly to the textural properties of muscle in mammals but not most fish muscles (Shenouda, 1980; Haard, 1990; Sikorski and Kolakowska, 1994; Venugopal and Shahidi, 1996). Molecular characteristics like conformation, charge and polarity as also the characteristics of the surrounding matrix viz. pH, ionic strength and temperature play an important role in determining the functionality of these proteins (Pardo, Joye and McClements, 2015). When subjected to suitable conditions that interfere with either of these characteristics, animal proteins tend to undergo structural changes and interactions to enable the functional characteristics. For example, gelation occurs as a result of matrix formation by extracted myofibrillar proteins and collagen protein-protein interactions in meat that immobilize water, fat and other constituents. This can be attributed to the high length-to-diameter ratio and bipolar structural arrangement of myofibrillar proteins that provide them with an excellent emulsification capability. This arrangement enables the hydrophobic site of the protein to interact with fat and their hydrophilic site to interact with water further reducing the surface tension of fat particles and form a rigid protein membrane in fat emulsion (Lotte, 2018). These functional properties contribute to the overall characteristics of animal protein products, including texture, appearance, mouthfeel, juiciness, and physical stability during storage (Rickey, 2017).

1.2 Need for Animal Protein Replacement and its probable alternatives

As beneficial animal protein may be for humans, its consumption comes with an environmental cost since meat production is one of the primary sources of greenhouse gas emissions. Approximately 70 billion animals are reared as domestic animals annually and thereby are a big contributor to global warming. As per a report by the Food and Agriculture Organization (FAO) of the United Nations “Tackling Climate Change through Livestock”, it is estimated that
14.5% of human induced global greenhouse gas emissions are caused by the livestock industry with approximately 7.1 gigatons carbon dioxide (CO₂) released to the atmosphere per year. Consequently, the global meat production is predicted to double from 229 million tons in 1999 to 465 million tons by 2050 indicating a rise in the market. Milk production is also expected to increase from 580 to 1043 million tons (Gerber, Steinfeld, Henderson, Mottet, Opio, Dijkman, Falcucci and Tempio, 2013).

With the ever-increasing population and the need to suffice their nutritional needs it is important to devise animal protein alternatives while addressing the concerns about the major effects of emissions on the environment as well as on human health and the economics of the food system. (Godfray, 2018). These concerns have formulated consumer groups like the vegetarians, vegans and consolidate the emergence of new groups. The flexitarians are one new group, for example, who consciously reduce animal protein consumption without removing it entirely from their diet (Alves and Tavares, 2019). Taking everything into consideration it is important to have an increased market availability and subsequently ensure improvements in research. Currently, the challenges faced by most food manufacturers while developing animal protein alternatives products is to acquire same or similar attributes such as texture, taste, flavor and mouth feel as the animal-derived versions. In terms of functional properties many alternative protein options tend to have poor solubility, foaming and emulsification properties compared to animal proteins that provide an additional challenge for the food industry to incorporate them in the food matrix and commercialize them (Brown, 2017). It is important to bridge this gap by finessing these functionalities through physical, chemical, enzymatic, and genetic modifications. Physical modification can either be partial denaturation or unfolding of the protein under controlled heating and shear conditions. Chemical modification on the other hand is based on the induction of structural changes in protein at the secondary, tertiary, and quaternary structural levels, or alteration of the hydrophobicity-hydrophilicity balance. Subsequently, hydrolysis and polymerization reactions can be considered...
as enzymatic modifications. The structure-function relationship of enzymes and mutative technologies are widely used in genetic modification while studying proteins. Application of these techniques are used to improve the functional properties of food proteins (Damodaran, 2005).

In recent times, there has been a development to meet this need of protein by introducing newer formats of meat alternatives like cultured meat, plant proteins or blends of animal and plant protein with the former being used in limited quantities than the latter. Cultured meat is a technology that involves expanding stem cells followed by differentiation into muscle cells and then processing into meat products. Chemical/biological cues in the cell culture media and mechanical stimulation are some of the techniques employed (Stephens, Di Silvio, Dunsford, Ellis, Glencross and Sexton, 2018). This in vitro production of meat using tissue engineering techniques may potentially lead to a reduction in environmental pollution and land use associated with current meat production systems. It broadens the scope of sustainable production by reducing the animal suffering significantly, facilitate production of chemically safe and disease-free meat with desirable nutritional profile since the conditions in an in vitro meat production system can be controlled. According to Kearney consultancy, plant-based products will account for 10% of global meat consumption within five years, although Kearney analysts expect cultured meat to eventually account for a larger share of the market than plant-based. While in vitro meat production or cultured meat may seem like the solution, it demands more research and development with respect to the cost effectiveness of the technology, and ethical and societal issues before achieving an effective large-scale production (Bhat, Kumar and Fayaz, 2015).

On the other hand, blended animal and plant protein still hold the issue of incorporation of animal proteins and may be treated as a temporary solution to the underlying long-term challenges.
1.3 Plant Protein as a Potential Alternative

There is an increased interest in substituting animal proteins with plant proteins due to its high protein delivery efficiencies in terms of energy use as also the reduced greenhouse gas emission (Gonzales, 2011). Apart from the environmental benefits, a decreased meat consumption is associated with lower consumption of potentially harmful fats. Subsequently an increased consumption of vegetables and plant-based foods, such as legumes, seeds, and whole grains, promotes better health and is associated with fewer metabolic syndrome (MetS) noncommunicable diseases such as obesity, type II diabetes, hypertension, cardiovascular disease, and nonalcoholic fatty liver disease (American Heart Association 2016; Dietary Guidelines 2017; Pérez-Martínez and others 2017; WHO 2017). This information has transcended into the commercial market and the consumers well, as statistics regarding plant-based protein have depicted an ever-increasing rise since 2010. Non-animal meats are no longer niche products. Recent numbers and survey show that sales of frozen and refrigerated meat substitutes in major outlets (grocery, drug, mass market, military, and select club and dollar store retailers) accounted to 19.1% which is equivalent to approximately $680.5 million for the year 2019, according to data from market research firm IRI (Kuhn, 2020). A 2019 survey by the Hartman Group, a research and consulting firm, found that only 12% of plant-based product purchasers described themselves as vegetarians, while 41% said they were omnivores. Further strengthening Hartman’s analysis, Gallup has also reported that 70% have cited health as a major reason for never eating meat or cutting it back. In addition, 49% of these consumers characterize environmental issues as a major factor, and 41% cite concern about animal welfare (McCarthy and Dekoster 2020).

Like animal proteins, each of the plant proteins also have their distinct functional properties such as emulsifying, foaming, solubility, water holding capacity, fat absorption, and gelling that influence the overall quality and sensory perception of foods. (Chan and Lacroix, 2017). However, a direct usage of plant proteins in the food matrix is accompanied with issues such as antinutrients,
poor WHC, off-flavors, and non-meat like textural properties (Joshi, 2016; Day, 2013, Silva et al., 2019b, Wouters et al., 2017). The favorable nutritional characteristics and functional properties of meat proteins have been very difficult to reproduce by any other food proteins or non-protein functional ingredients. Also, during operations like extrusion or extraction, processing conditions (i.e. temperature, pH, and ionic strength) highly influence protein functionality. For example, many heat treatment or temperature change operations alter the ultimate protein functionality as proteins unfold, exposing enfolded hydrophobic groups and promoting formation of covalent bonds between proteins. This results in new three-dimensional structures or aggregates of the proteins (Ainsworth, 2012). However, plant proteins also display some advantages that include a lower cost of production, low allergenicity (with some exceptions, e.g. gluten) and unique techno-functional properties.

The focus of this thesis would be to study these conditions and observe the structural changes in proteins from animal and plant sources using molecular spectroscopy methods.

1.4 Molecular Spectroscopy Methods for Protein Analysis

Spectroscopy has been one of the favored modes of analysis by researchers to understand the properties, interaction and reactions of proteins since the early 1980s. Circular dichroism (CD) is a robust analytical method for the study of protein structure and folding in solution under various conditions (Compton and Johnson, 1986; Pelton and McLean, 2000). In the food industry, CD finds its application in monitoring the structural transitions of food and its bioproduct proteins due to the changes in varied processing conditions (Martin and Schilstra, 2008). However, this technique has its own limitations. It is not suitable for direct spectroscopic measurements of samples such as muscle food systems which are in the solid state (Herrero, 2008). Fluorescence spectroscopy (FS) is another biophysical method to study proteins, majorly based on tryptophan fluorescence which is only excited by certain wavelengths light. FS successfully provides information for both qualitative and quantitative analysis. But its inability to determine the peptide backbone structure
and inefficient functioning in environmental conditions apart from tyrosine and tryptophan side chains serves as a drawback (Sz‘oll‘osi et al., 1987; Herrero, 2008). Nuclear magnetic resonance (NMR) has become one of the preferred analytical techniques for metabolomics studies due to its inherent nondestructive nature and rapid analysis (Hatzakis, 2019). Recently, it is being applied in the domain of protein analysis as well, as it provides a detailed and specific information about peptides and proteins in solution. But NMR tends to be limited to low molecular weight proteins (McKelvy et al., 1996; Scotter, 1997; Pelton and McLean, 2000; Herrero, 2008). Another technique employed for analysis of the conformational structure of proteins and polypeptides is infrared spectroscopy (IR). However, near infrared (NIR) spectroscopy provides a low spectral resolution for proteins while, mid-IR spectroscopy has very strong infrared absorption of water which interfere with the spectra of proteins (Herrero, 2008). Fourier transform infrared spectroscopy (FTIR) is another widely utilized analytical tool that monitors changes in protein and polypeptide secondary structure during processing. The need for sophisticated optical materials proves to be a major challenge for this method (Arrondo et al., 1993; McKelvy et al., 1996; Pelton and McLean, 2000; Herrero, 2008).

And lastly Raman spectroscopy - an analytical technique that provides information on the peptide backbone structure, the environment of some side chains like the hydrophobic groups and the local conformations of disulfide bonds and methionine residues. Some of the typical assignments of major Raman bands are represented in Table 1. Raman Spectroscopy addresses most of limitations faced by other spectroscopic methods and could be used to determine the protein structure but has a few drawbacks. (Frushour and Koening, 1975a; Tu, 1982; Pelton and McLean, 2000; Tuma, 2005; Herrero, 2008). Its relatively weak signals and fluorescence interference makes it a lesser used instrument than IR in food analysis. However, it was found that if the surface of the substrate is altered using noble metal nanostructures the weak Raman scattering signals are greatly
enhanced while the fluorescence would be suppressed. This phenomenon was later coined the term Surface Enhanced Raman Spectroscopy (Zheng and He, 2014).

Table 1. Typical Wavenumbers of Raman Bands & General Assignments in Raman Protein Spectra

<table>
<thead>
<tr>
<th>Origin</th>
<th>Wavenumber (cm(^{-1}))</th>
<th>Assignment</th>
<th>Structural Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>cystine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>510</td>
<td>S–S stretch</td>
<td>gauche–gauche-gauche conformation</td>
</tr>
<tr>
<td></td>
<td>525</td>
<td>S–S stretch</td>
<td>gauche–gauche-trans conformation</td>
</tr>
<tr>
<td></td>
<td>545</td>
<td>S–S stretch</td>
<td>trans-gauche–trans conformation</td>
</tr>
<tr>
<td>tyrosine</td>
<td>850/830</td>
<td>Fermi resonance between ring fundamental and overtone</td>
<td>state of phenolic OH group (exposed or buried, hydrogen-bond donor or acceptor)</td>
</tr>
<tr>
<td>tryptophan</td>
<td>760, 880, 1360</td>
<td>indole ring</td>
<td>sharp intense band indicates buried residues; sensitive to environment polarity</td>
</tr>
<tr>
<td>phenylalanine</td>
<td>1003</td>
<td>ring breathe</td>
<td>conformation insensitive; useful as an internal intensity standard</td>
</tr>
<tr>
<td>aspartic and glutamic acids</td>
<td>1400–1430</td>
<td>C=O stretch of COO-</td>
<td>ionized carboxyl groups</td>
</tr>
<tr>
<td>aliphatic residues</td>
<td>1450, 1465</td>
<td>C–H bending</td>
<td>microenvironment, polarity</td>
</tr>
<tr>
<td>amide I</td>
<td>1650–1660</td>
<td>amide C=O stretch, N–H wag</td>
<td>(\alpha)-helix, antiparallel (\beta)-sheet, random coil</td>
</tr>
<tr>
<td>amide III</td>
<td>1260–1300</td>
<td>N–H in-plane bend, C–N stretch</td>
<td>(\alpha)-helix, antiparallel (\beta)-sheet, random coil</td>
</tr>
<tr>
<td></td>
<td>1230–1240</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1245 ± 4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Zhao, Ma, Yuen & Phillip, 2004; Li-Chan, 1996)

Apart from procurement of sample spectra, it is also important to analyze data acquired from various molecular spectroscopic methods to have a better understanding. Chemometric analysis is an effective way for multivariate data processing and analysis in spectroscopy techniques and hyperspectral imaging. It builds calibration and prediction models that best describe the system under analysis by measurement, quantification, classification, identification, detection, and assessment of quality and safety of food. The spectral data can be analyzed directly after a few pretreatments that eliminate the nonlinearities caused by light scattering, fluorescence and baseline
shift. Principal Component Analysis (PCA) is the most widely used method for spectral analysis that performs an initial investigation, visualizes the spectral data and examines any possible grouping of samples according to spectral features of the sample (Cheng and Sun, 2015).

1.5 Surface Enhanced Raman Spectroscopy for Protein Analysis

Surface Enhanced Raman Spectroscopy (SERS) is a rapid, non-destructive and ultra-sensitive spectroscopic technique capable of femtomolar levels of detection (Shin, Oh, Kang and Choi, 2020). It utilizes electromagnetic enhancement and chemical enhancement to improve the strength of Raman signals. Electromagnetic enhancement is induced by localized surface plasmon resonance (LSPR) on the noble metal nanostructures, often referred to as ‘hot spots’. On the other hand, chemical enhancement relies mainly on the arrangement and morphology of the noble metal surface as molecules of the analyte are adsorbed onto the roughened substrate surfaces (Pham, Hoang and others, 2019). Most commonly used SERS-active substrates can be classified into two categories: solid surface-based substrates and colloid-based substrates (Xie, Pu and Sun, 2017). Solid surface-based substrates are nanoparticles stabilized on solid substrates that offer relatively consistent results and a simplicity of sample preparation, but maybe an expensive option. Colloid-based substrates on the other hand are formulated by stabilizing colloidal nanoparticles in the solution-based system. They are comparatively cheap, convenient to make and hence well suited for both large-scale commercial and laboratory production (Zheng and He, 2014).

It has been observed that SERS substrates on surface roughening or aggregation (such as flowers-like or dendrites) showed larger enhancement factors of SERS activity due to the creation of multiple hot-spots and uniform distribution of nanostructures on the surface. These dendritic nanostructures are formed from the tips and the sharp edges of the trunk and branches of the nano-dendrites, and the narrow gap between the branches (Yin, 2015; Pham, Hoang and others, 2019). When silver is employed to form these nanostructures viz. Ag dendrites, it highly improves the application of the substrate due to its plasmonic properties that allows breaking the diffraction limit
enabling strong field enhancements (Gather, 2012). They are fabricated by electrochemical deposition method to increase high density hot-spots for enhanced SERS activity (Rafailovic, 2015). This method seems promising for protein analysis and hence has been employed in this thesis. SERS has been widely used to detect adulteration in various foods but research on food protein analysis using this molecular spectroscopy technique has been comparatively limited (Craig, Franca and Irudayaraj, 2013). Individual research on proteins focusing on a particular parameter only has been reported by publications in the past. However, an elaborate comparison of proteins using Raman Spectroscopy or SERS has not been done before thus offering a promising future to the research.

1.6 Goals and Objectives

We aim to develop an approach based on Raman Spectroscopy and SERS to study animal and plant proteins. To achieve this goal, there are three objectives of this study.

Objective 1: Establish methods for animal and plant protein spectroscopic analysis using Normal Raman and SERS.

Objective 2: Study the effects of pH and salt on the spectroscopic characterization of proteins using Raman Spectroscopy.

Objective 3: Apply the developed method to study the spectroscopic characteristics of a plant protein in a food product of animal protein alternatives.
CHAPTER 2

DEVELOP AND OPTIMIZE METHOD FOR PLANT AND ANIMAL PROTEIN SPECTROSCOPIC ANALYSIS USING NORMAL RAMAN AND SERS

2.1 Introduction

As mentioned in chapter 1, attempts have been made to utilize various analytical tools to study the structure of proteins. However, comparatively Raman spectroscopy and SERS tend to be an under-used technique utilized for protein differentiation and analysis. Even if many methods exist to understand the structure of proteins, very few can be utilized to study proteins under conditions that are applicable to their eventual use in the food matrix (Yada, Jackman and Smith, 1994). Our lab has successfully used various colloid based substrates for different analytical applications and specific protein detection (Zheng, Zhao, Tian and He, 2017), it would be interesting to know if that work can be further extended to study food proteins and potentially create a generalized method that can be used for other proteins. Here, we investigate different protein isolates using Raman and SERS with different colloid-based substrates and discuss if either is advantageous over the others for protein differentiation and peak identification.

2.1.1 Characteristic spectrum of proteins

From literature and previous work done in our lab we know that Raman and SERS is capable of generating a distinct spectrum for proteins. It provides information on the secondary structure, conformational changes and microenvironment of protein side chains. Among many distinct vibrational modes of the amide (-CO-NH-), the most readily available regions of Raman spectra for determining protein secondary structures are the amide I (1645–1685 cm\(^{-1}\)) and III (1200– 1350 cm\(^{-1}\)) bands which comprise of \(\alpha\)-helices, \(\beta\)-sheets, turns and random coil structures. Other characteristic bands like the phenylalanine residue (1004 cm\(^{-1}\)) also tend to be important. It is typically reported to be insensitive to conformation or microenvironment changes and hence used as an internal standard for most Raman Spectroscopic studies (Ismail et al., 2013). Some
information regarding the tertiary structure can also be acquired by the tryptophan (Trp) bands, tyrosine (Tyr) bands, inter-chain disulfide bands and aliphatic hydrophobic residues in the Raman spectra. The frequency position of these bands depends strongly on the protein state, environment and the intermolecular interactions thus providing peculiar information of the protein structure. (Blanpain-Avet et al., 2012).

2.1.2 Noble metal substrates

As mentioned in chapter 1, Raman spectroscopy is a powerful tool for protein analysis but has low sensitivity, which limits its applications to powders or highly concentrated samples. SERS on the other hand can generate a spectrum with trace amount of sample to the limit detection order of parts per billion with the help of noble metal substrates. It has been seen that the interaction of Ag-substrates is the most efficient for protein study using Raman and hence only Ag based substrates viz. silver nanoparticles (AgNPs) and silver dendrites (Ag dendrites) are considered for this research to evaluate the feasibility of adopting substrates for the structural study of proteins. Each of the substrates is manufactured differently to create a unique interface for silver to interact with the sample. Factors such as the shape, size and surface morphology of the substrate determines the ability of signals enhancement and creation of hotspots on the nanostructure (Boisselier and Astruc 2009). Because the interaction between these substrates and proteins is unclear, we first investigated the spectra of one animal and one plant protein i.e. whey and mung respectively using the above mentioned Ag-substrates before applying it to other proteins.

2.1.3 Objectives of this study

The objectives of this study were to (1) determine a method for secondary structure analysis and peak detection for whey and mung protein using Raman and SERS (2) apply method to 5 other proteins (casein, egg albumin from chicken egg white, soy, pea and faba bean protein) and (3) compare and differentiate protein spectra using PCA.
2.2 Materials and methods

2.2.1 Materials

Protein isolates with protein content above 90% were acquired from following commercial sources. Whey protein isolate (WPI) with 93.5% protein was purchased from Agropur (La Crosse, WI), k-casein and Albumin from chicken egg white with 99.99% protein content from Sigma Aldrich (St Louis, MO), Soy Protein Isolate (SPI) (SUPRO EX 45) with 92% protein from Essex Food Ingredients (Frazer, PA), Mung bean protein isolate (MPI) with 93% protein from Scouler (Omaha, NE), Pea protein isolate (PPI) and Faba bean protein isolate with 92.1% protein from AGT Foods (Saskatoon, SK, Canada).

0.02 mg/ml 40nm bare (citrate) AgNPs solution and 0.2 mg/ml 60nm citrate coated AgNPs solution was purchased from NanoComposix (San Diego, CA). For silver dendrites Zn foil (99.99%) and silver nitrate (99%) was acquired from Fischer Scientific (Rochester, NY).

2.2.2 Preparation of Ag dendrites

Silver dendrites were prepared using a simple replacement reaction using zinc and silver nitrate (AgNO₃). Briefly, zinc plate was immersed in 100 mM HCl followed by immersion in 200 mM AgNO₃ to enable the deposition of the dendrites on the zinc foil. The dendrites are then scraped off the zinc foil and washed using DI water and stored at room temperature.

2.2.3 Sample Preparation

The limit of detection is different for Raman and SERS and hence two different methods of sample preparation have been adopted with different concentration of protein. Since the isoelectric point (pl) of all proteins is around 4.5, the phosphate buffer used was at pH 3 to ensure no isoelectric precipitation. For Normal Raman, 1% solution of protein was mixed with pH 3 phosphate buffer in equal parts. After incubating for 15 mins an aliquot (10uL) was dropped on the gold slide and allowed to air dry. The sample preparation of SERS used 100ppm protein solution mixed with pH 3 phosphate buffer in 2:1 proportion for 5 mins. Then 1 ml of the solution was
transferred to an Eppendorf tube and 20uL of Ag substrate was added and pipette mixed for 1 min. The substrate were then allowed to settle to the bottom of the Eppendorf tube taking about 15-20 mins assuring enough time for the protein interaction with the nanostructures. Then 10uL of precipitate under the bottom was deposited on the gold slide and air dried for further analysis. The preparation method for each is demonstrated in Figure 1, Raman spectra were collected individually for each of the methods.

2.2.4 Raman Instrument and Data Analysis

The aliquots were analyzed using a DXR Raman microscope (Thermo Fischer Scientific, Madison, WI). The instrument is equipped with a 780 nm excitation laser and a 20X objective.

confocal microscope, 50um slit width, 10 s integration time and a spectral range of 400-2000cm⁻¹. Due to different sensitivities of each of the methods, the laser power for Normal Raman and SERS was set differently. A 24mW laser power was used to collect spectra for Normal Raman while SERS spectra was collected using a laser power of 4mW. The Raman instrument was controlled by OMNIC™ software (version 9.1). Ten spectra were selected from each of the aliquots.

Figure 1. Schematic illustration of the sample preparation method
and then averaged by OMNIC™ software. Later the grouped spectra were analyzed by Thermo Scientific TQ Analyst (version 8.0). All Raman spectra were calculated from at least three replicates. Standard normal variant, second derivative transformation, and smoothing were applied when necessary to reduce spectral noise, normalize the spectra (against phenylalanine band at 1004 cm$^{-1}$), separate overlapping bands, and remove baseline shifts. An averaged spectrum was generated by the software for each sample (representing each treatment or control).

Principal component analysis (PCA), a feature of TQ Analyst to discriminate data was used to analyze the variances of the spectra. For PCA analysis, the spectra were pre-processed using secondary derivative transformation with Norris derivative filter (Segment length 7–11 cm$^{-1}$, gap between segment 7–11 cm$^{-1}$), and then used Standard normal variant based on the region (400-2000 cm$^{-1}$), which include the phenylalanine peak (Ismail et al., 2013). It is possible to substantially reduce a multidimensional dataset to its most dominant features, eliminate random variation and better understand the variation between spectra through the retained principal components (PCs). This information can potentially give an estimate of the variance within a class and between different classes. If two clusters have no overlap, it means that the two samples are significantly different from each other (Zheng, Zhao, Tian and He, 2017).

2.3 Result and Discussion

2.3.1 Characterization of Raman and SERS with Ag substrates spectra of whey and mung protein

To study the effect of SERS, spectrum of WPI was acquired using various silver-based substrates as also without any substrate as represented in Fig. 2. In general, all methods except AgNPs could successfully identify the signature phenylalanine peak (1004 cm$^{-1}$) of proteins. WPI with Normal Raman was able to detect most of the characteristic peaks mentioned in section 2.1.1. and hence could be a potential method for peak detection. However, that was not the case for AgNPs which failed to show the comparatively inert phenylalanine peak and suggest that the
protein may not have interacted with the nanoparticles. Apart from the phenylalanine peak other characteristically significant peaks like the Amide I and Amide III were also absent. Based on literature, we know that these bands are a crucial part of the protein backbone structure and hence AgNPs was deemed incapable of enhancing protein signals for WPI. Ag dendrites gave significantly good results as it was able to successfully detect the peaks that could provide information about the protein backbone structure. The configuration of the protein based on its peaks would be discussed further in detail in section 2.3.3.

![Figure 2. Representative spectra of WPI with Normal Raman, AgNPs background, WPI with AgNPs, Ag dendrites background and WPI with Ag dendrites.](image)

Having studied an animal protein, it is also important to understand if this method could be applied to plant proteins as well. To this end, spectrum of MPI was acquired using various silver-based substrates as also without any substrate as represented in Fig. 3. Similar to the results of WPI, all methods except AgNPs could perfectly depict a high intensity phenylalanine peak (1004 cm\(^{-1}\)) of proteins. MPI with Normal Raman was able to detect most of the characteristic peaks mentioned in section 2.1.1. and hence could be a possible method for peak detection. However, that was not
the case for AgNPs which failed to show phenylalanine, Amide I and Amide III peak suggesting the possibility that the protein may not have interacted with the nanoparticles. Just like WPI, AgNPs failed to show the protein spectrum and was deemed incapable of enhancing protein signals for MPI. Ag dendrites gave significantly good results as it was able to successfully detect the peaks that could provide information about the protein backbone structure. It was beneficial to use Ag dendrites since the background does not have any overlapping peaks with the protein structure that would interfere with the protein spectrum. It is also seen that the spectrum generated using dendrites not just identified the phenylalanine, Amide I and Amide III but also tryptophan and CN stretch in the range of 600-900 cm$^{-1}$ under similar conditions as the Normal Raman. The configuration of the protein based on its peaks would be discussed further in detail in section 2.3.3.

![Figure 3. Representative spectra of MPI with Normal Raman, AgNPs background, MPI with AgNPs, Ag dendrites background and MPI with Ag dendrites.](image)

2.3.2 SERS Analysis of protein structure using silver dendrites

Based on the results acquired in section 2.3.1, it is seen that SERS with Ag dendrites can be termed as a better method for peak detection for both WPI and MPI. Hence, the method was
further applied to other proteins using the same protocol to generate spectra (Fig. 4). All proteins depicted a strong phenylalanine peak across which the spectra is normalized and baseline corrected. The SERS spectrum of all proteins also has a wide amide-I band at 1674 cm\(^{-1}\) with a peak width of approximately 50 cm\(^{-1}\) for all of the proteins. It is important to note that the asymmetry of the amide I band along with the peak width showcase a distribution of secondary structures. The band at 1665 cm\(^{-1}\) is a marker band for alpha-sheet and the shoulder at 1653 cm\(^{-1}\) is a marker band associated with R-helical peptide bonds (Maiti et al, 2003).

![SERS spectra of different proteins](image)

**Figure 4.** Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate.

All proteins except casein also projected information on the secondary structures through the amide III region of the spectrum. It is seen to attain maximum height at about 1252 cm\(^{-1}\), attributed to the undefined secondary structures in the protein solution. However, there were some peaks that were only identified in plant protein spectra. The tryptophan peaks (700- 800 cm\(^{-1}\)) in the plant proteins that tended to be absent in the animal proteins was a significant difference. These
peaks show that the buried tryptophan residues from a hydrophobic microenvironment became exposed which was not observed in animal proteins. The S-S stretches originating from cystine also had a slightly different frequency for the animal-based (510 cm\(^{-1}\)) and plant-based (525 cm\(^{-1}\)) proteins. The S-S stretch in animal based proteins depicted a gauche-gauche-gauche conformation while the plant proteins had a gauche-gauche-trans conformation. The band at 1400-1430 cm\(^{-1}\) which is attributed to the dissociated or ionized carboxyl (COO\(-\)) groups was only found in faba bean, PPI & SPI.

Table 2. Summary of Raman bands characterized by SERS using Ag dendrites

<table>
<thead>
<tr>
<th>Origin</th>
<th>Wavenumber (cm(^{-1}))</th>
<th>Assignment</th>
<th>Proteins comprising it</th>
</tr>
</thead>
<tbody>
<tr>
<td>cystine</td>
<td>510</td>
<td>S–S stretch</td>
<td>Whey, casein and egg albumin</td>
</tr>
<tr>
<td></td>
<td>525</td>
<td>S–S stretch</td>
<td>MPI, SPI, PPI and faba bean</td>
</tr>
<tr>
<td>tryptophan</td>
<td>740</td>
<td>indole ring</td>
<td>MPI, SPI, PPI and faba bean</td>
</tr>
<tr>
<td>phenylalanine</td>
<td>1003</td>
<td>ring breathe</td>
<td>All proteins</td>
</tr>
<tr>
<td>aspartic and glutamic acids</td>
<td>1400</td>
<td>C=O stretch of COO-</td>
<td>SPI, PPI and faba bean</td>
</tr>
<tr>
<td>aliphatic residues</td>
<td>1450</td>
<td>C–H bending</td>
<td>All proteins</td>
</tr>
<tr>
<td>amide I</td>
<td>1650–1660</td>
<td>amide C=O stretch, N–H wag</td>
<td>All proteins</td>
</tr>
<tr>
<td></td>
<td>1667–1673</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1665 ± 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>amide III</td>
<td>1260–1300</td>
<td>N–H in-plane bend, C–N stretch</td>
<td>All proteins</td>
</tr>
<tr>
<td></td>
<td>1230–1240</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1245 ± 4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It will be interesting to observe the intensity differences of these peaks when the protein environments are changed and quantify the change. For now, peaks observed in both classes of proteins viz animal and plant have been presented (Table 2)
2.3.3 Principal Component Analysis of Proteins

As discussed in the previous section, it has been seen that the spectrum of animal and plant proteins has certain similarities and dissimilarities. However, it can be difficult to summarize the difference based on visible SERS spectra pattern and peaks. To enable a better relevance between experimental data and the structure of proteins, an innovative non-linear model is necessary to be established. Therefore, we analyzed the SERS data with principal component analysis (PCA). Generally, the PCA plot is able to show significant statistical discrimination of spectra from different data groups and based on this principle we shall attempt to see differences between animal and plant-based proteins (Fig 5.)

Figure 5. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate.

It was seen that the datapoints of all proteins were all clustered and could be very well distinguished from the animal proteins that lie in a different plane. There was significant overlap between the datapoints of PPI and faba bean protein. The spectrum of mung bean protein isolate data seemed to be the most distinct as it lied in a different plane from all other proteins. This indicates that there is significant statistical difference in the two groups of proteins. Hence, both
the SERS spectra and the PCA model had the capability to discriminate between animal and plant-based proteins.

2.4 Conclusion

This chapter discusses the various methods employed to study and predict the secondary structure of the protein using Normal Raman and SERS. Silver based substrates namely AgNPs and Ag dendrites were used for the optimization of this approach. A large amount of spectral data were classified and analyzed for identifying the characteristic Normal Raman and SERS signals of whey protein isolate and mung protein isolate. Based on the results acquired from these two proteins, it was concluded that SERS with Ag dendrites could be the better method for peak detection and classification and was hence applied to 5 other proteins. The similarities and differences in the spectrum of both classes of proteins were then studied by applying PCA for better clarification that distinguished the proteins. It was observed that some peaks were exclusive to each class of protein and hence, it would be interesting to understand the effect of change in environmental conditions on these peaks. Future work includes the study of effect of change in pH and ionic strength on the secondary structure of the protein and potentially develop a method for the same.
CHAPTER 3
STUDY EFFECT OF CHANGE IN PH ON SPECTROSCOPIC CHARACTERIZATION OF ANIMAL AND PLANT PROTEINS USING RAMAN SPECTROSCOPY.

3.1 Introduction

In Chapter 2, we established a method to detect peaks and allocate functional groups that would provide more information on the structure of the proteins. We illustrated the merits and demerits of using Raman Spectroscopy for structural analysis with and without substrates and concluded that Ag dendrites deem to be the most successful method to identify peaks. Herein, we would forward this research to establish a potential method to study the effect of change in pH and quantify it. The focus would majorly be on studying the differences in the spectrum of protein subject to change in pH and how compatible would the substrates be to detect protein signals in varied environments.

3.1.1 Effect of change in pH on protein

The solubility of proteins in aqueous solutions is largely governed by the pH of the solution. At the isoelectric pH, protein exhibit minimum solubility as the hydrophobic interaction between proteins reaches maximum and hinders unfolding of the protein molecules. There is no surface net charge at the isoelectric point of the protein while there exists a charge on the surface above and below it. These charges result in the folding and unfolding of proteins that would eventually expose or hide certain groups that may result in a different spectrum across the pH range.

3.1.2 Objectives of the study

The objectives of this study were to (1) determine a method to study structural changes in the protein due to change in pH for whey and mung protein using Raman and SERS (2) apply method to 5 other proteins (casein, egg albumin from chicken egg white, soy, pea and faba bean protein) and (3) compare and differentiate protein spectra using PCA.
3.2 Materials and methods

3.2.1 Materials

Protein isolates with protein content above 90% were acquired from following commercial sources. Whey protein isolate (WPI) with 93.5% protein was purchased from Agropur (La Crosse, WI), k-casein and Albumin from chicken egg white with 99.99% protein content from Sigma Aldrich (St Louis, MO), Soy Protein Isolate (SPI) (SUPRO EX 45) with 92% protein from Essex Food Ingredients (Frazer, PA), Mung bean protein isolate (MPI) with 93% protein from Scouler (Omaha, NE), Pea protein isolate (PPI) and Faba bean protein isolate with 92.1% protein from AGT Foods (Saskatoon, SK, Canada). Here, all proteins taken have an isoelectric point (pI) of around 4.5

For silver dendrites Zn foil (99.99%) and silver nitrate (99%) was acquired from Fischer Scientific (Rochester, NY).

3.2.2 Sample Preparation

Similar to section 2.2.2, two different methods of sample preparation have been adopted with different concentration of protein since the limit of detection is different for Raman and SERS. For Normal Raman, 1% solution of protein was mixed with pH 7.4 phosphate buffer in equal parts. Since the isoelectric points of all proteins were around 4.5, the pHs of the solution were modulated to either 3,6 or 9 using 0.5M HCl or 0.5M NaOH for pH analysis to avoid isoelectric precipitation. The solution was mixed for 15 mins. After incubating for another 15 mins the pH of the solution was checked and an aliquot (10uL) was dropped on the gold slide and allowed to air dry. Then, another 10uL solution was dropped on the air-dried aliquot to ensure uniformity and produce stronger signals under the Raman instrument. The sample preparation of SERS used 100ppm protein solution mixed with pH 7.4 phosphate buffer in 2:1 proportion for 5 mins. The pH of the solution was modulated to either 3,6 or 9 using 0.5M HCl or 0.5M NaOH. The solution was mixed
for 15 mins. Then 1 ml of the solution was transferred to an Eppendorf tube and 20uL of Ag substrate was added and pipette mixed for 1 min. The substrate was then allowed to settle to the bottom of the Eppendorf tube taking about 15-20 mins assuring enough time for the protein interaction with the nanostructures. Then 10uL of precipitate under the bottom was deposited on the gold slide and air dried for further analysis. The preparation method for each is demonstrated in Figure 5, Raman spectra were collected individually for each of the methods.

Figure 6. Schematic illustration of the sample preparation method
3.2.3 Raman Instrument and Data Analysis

The aliquots were analyzed using a DXR Raman microscope (Thermo Fischer Scientific, Madison, WI). The instrument is equipped with a 780 nm excitation laser and a 20X objective confocal microscope, 50μm slit width, 10 s integration time and a spectral range of 400-2000cm⁻¹. Due to different sensitivities of each of the methods, the laser power for Normal Raman and SERS was set differently. A 24mW laser power was used to collect spectra for Normal Raman while SERS spectra was collected using a laser power of 4mW. The Raman instrument was controlled by OMNIC™ software (version 9.1). Ten spectra were selected from each of the aliquots and then averaged by OMNIC™ software. Later the grouped spectra were analyzed by Thermo Scientific TQ Analyst (version 8.0). All Raman spectra were calculated from at least three replicates. Standard normal variant, second derivative transformation, and smoothing were applied when necessary to reduce spectral noise, normalize the spectra (against phenylalanine band at 1004 cm⁻¹), separate overlapping bands, and remove baseline shifts. An averaged spectrum was generated by the software for each sample (representing each treatment or control).

Principal component analysis (PCA), a feature of TQ Analyst to discriminate data was used to analyze the variances of the spectra. For PCA analysis, the spectra were pre-processed using secondary derivative transformation with Norris derivative filter (Segment length 7–11 cm⁻¹, gap between segment 7–11 cm⁻¹), and then used Standard normal variant based on the region (400-2000 cm⁻¹), which include the phenylalanine peak (Ismail et al., 2013). It is possible to substantially reduce a multidimensional dataset to its most dominant features, eliminate random variation and better understand the variation between spectra through the retained principal components (PCs). This information can potentially give an estimate of the variance within a class and between different classes. If two clusters have no overlap, it means that the two samples are significantly different from each other (Zheng, Zhao, Tian and He, 2017).
3.3 Result and discussion

3.3.1 Characterization of Raman and SERS with Ag substrates spectra of whey and mung protein at pH 9

To study the effect of SERS, spectrum of WPI was acquired using various silver dendrites as also without any substrate as represented in Fig. 7. Here, AgNPs were not taken into consideration due to its established incapability to produce signals at pH 3 which is one of the values we would operate with eventually. As observed in section 2.3.1 AgNPs could not identify the signature phenylalanine peak (1004 cm\(^{-1}\)) of proteins and hence we conclude that use of this substrate may not be functional for this objective.

![Figure 7. Representative spectra of WPI with Normal Raman, Ag dendrites background and WPI with Ag dendrites at pH 9.](image)

WPI with Normal Raman was able to detect most of the characteristic peaks even at pH 9 as mentioned in section 2.1.1. and hence could be a potential method for structure analysis. However, that was not the case for Ag dendrites which failed to show the comparatively inert phenylalanine peak and suggest that the protein may not have interacted with the nanoparticles. Apart from the phenylalanine peak other characteristically significant peaks like the Amide I and
Amide III were also absent. Based on literature, we know that these bands are a crucial part of the protein backbone structure and hence Ag dendrites were incapable of enhancing protein signals for WPI. Normal Raman gave significantly good results as it was able to successfully detect the peaks that could provide information about the protein backbone structure. The configuration of the protein based on its peaks would be discussed further in detail in section 3.3.2. Having understood the capabilities of Normal Raman and inability of Ag-substrates to study an animal protein at high pH, it is also important to observe if this phenomenon is seen in plant proteins as well. To this end, spectrum of MPI was acquired using Ag dendrites substrates as also without any substrate as represented in Fig. 8.

![Figure 8. Representative spectra of MPI with Normal Raman, Ag dendrites background and MPI with Ag dendrites at pH 9.](image)

Normal Raman could perfectly depict a high intensity phenylalanine peak (1004 cm\(^{-1}\)) for mung protein. MPI with Normal Raman was able to detect most of the characteristic peaks mentioned in section 2.1.1. and hence could be a possible method for structure analysis. However, the same cannot be said about Ag dendrites which failed to show phenylalanine, Amide I and Amide III peak suggesting the possibility that the protein may not have interacted with the substrates. Just
like WPI, Ag dendrites was unsuccessful in showing the protein spectrum and could not enhance protein signals for MPI. Normal Raman gave significantly good results for both the proteins as it was able to successfully detect the peaks that could provide information about the protein backbone structure. As seen in section 2.3.1 Normal Raman was able to provide a distinct Raman spectrum at pH 3 as well signifying that the application of the method is not influenced by changes in pH unlike Ag dendrites whose sensitivity was compromised. Herein, Normal Raman is the method that would be utilized for studying the effect of change in pH on the structure of proteins.

3.3.2 Normal Raman analysis to study the effect of change in pH on structure of proteins

Based on the results acquired in section 2.3.1 and 3.3.1, it can be said that Normal Raman is a better method for structural analysis for both WPI and MPI at pH 3 & pH 7. Hence, the method was further applied to other proteins using the same protocol to generate spectra at pH3, 6 and 9 (Fig. 9, 10 & 11). All proteins depicted a strong phenylalanine peak across which the spectra is normalized and baseline corrected. The spectrum of all proteins showed a wide amide-I band at 1674 cm$^{-1}$ with a peak width of approximately 50 cm$^{-1}$ for all of the proteins except pea protein at pH 3. It is important to note that the asymmetry of the amide I band along with the peak height and width showcase a distribution of secondary structures. The band at 1665 cm$^{-1}$ is a marker band for alpha-sheet and the shoulder at 1653 cm$^{-1}$ is a marker band associated with R-helical peptide bonds (Maiti et al, 2003).
Figure 9. Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 3.

Figure 10. Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 6.
All proteins except pea depicted stability across all pHs with the mentioned protein showing low intensity amide bands. All protein also projected information on the secondary structures through the amide III region of the spectrum. The peak at 940 depicts the CN stretch in proteins which was significantly noticeable at pH 3 & 6 but was absent at pH 9. This could be attributed to its interaction and unfolding that exposed these groups under acidic pH and which were missing at basic. Visibly the spectrum of casein was the most distinct at all pHs as it identified the tryptophan bands in all pHs which were missing in other animal proteins. However, there were some peaks that were only identified in plant protein spectra. Most changes happened to be in the range of 800-1000cm⁻¹. These peaks show that the buried tryptophan residues from a hydrophobic microenvironment became exposed which was not observed in animal proteins. The band at 1400-1430 cm⁻¹ which is attributed to the dissociated or ionized carboxyl (COO⁻) groups was found in all proteins. Extreme pH values also led to decreases in the tyrosine doublet band intensity which indicates an increased participation of the tyrosine phenolic groups as hydrogen bond donors (Ma,
These were the changes observed by studying the spectrum, however it is important to analyze these peaks with a more definitive method and hence we employ PCA to compare proteins.

### 3.3.3 Principal Component analysis of proteins at pH 3, 6 and 9

As discussed in the previous section, it has been seen that the spectrum of animal and plant proteins has certain similarities and dissimilarities. However, it can be difficult to summarize the difference based on visible Raman spectra pattern and peaks. To enable a better relevance between experimental data and the structure of proteins, an innovative non-linear model is necessary to be established. Therefore, we analyzed the Raman data with principal component analysis (PCA). Generally, the PCA plot is able to show significant statistical discrimination of spectra from different data groups and based on this principle we shall attempt to see differences between animal and plant-based proteins at different pHs (Fig 12, 13 & 14).

![Image of PCA plot showing protein distribution at pH 3](image-url)

**Figure 12.** Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI and faba bean protein isolate at pH 3.

At pH 3 the animal proteins were well distinguishable with whey, casein and egg albumin having significantly different structures. The structures of SPI, PPI and faba bean at overlaps in their spectrum hinting that the proteins were not structurally that different from each other.
However, MPI was the protein which was not a part of the other plant protein cluster and was close to WPI and lied on the same plane as casein. This is important note and see if this behavior is observed for other proteins at higher pHs as well.

Changes were observed with a change in pH as all plant proteins attained a spectrum similar to that of casein. The animal proteins could still be distinguished properly as they were placed far

Figure 13. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 6.

Figure 14. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at pH 9.
apart in separate planes altogether. There was significant overlap between the datapoints of SPI and PPI. However, the faba bean protein isolate data seemed to be scattered but still could be distinguished from the animal proteins with no overlapping datapoints.

However, as pH changed to a basic pH of 9 it was seen that the spectrum of egg albumin protein became relevant to that of SPI and PPI while faba bean formed its separate cluster showing a notable change. A constant overlap between SPI and PPI has been observed throughout the entire pH range while significant differences were observed for all other proteins especially plant and animal proteins. This indicates that there is significant statistical difference in the two groups of proteins. Hence, both the Raman spectra and the PCA model had the capability to provide information on the effect of pH on the structure of proteins.

3.4 Secondary Structure Analysis

The Amide-I peak showed Raman bands at 1650, which depicts C=O stretching and N-H in plane bending of peptide groups. This vibrational mode depicts a characteristic spectral profile that can provide an approximate quantitative information about the secondary structure of the proteins (Herrero et.al, 2008). The amide-I band can be further interpreted into conformational information on the motifs viz. α-helix, antiparallel β-sheet, parallel β-sheet turn and unordered random coil structures. The frequencies allocated to the individual motifs is discussed in Table 3 (Alix, Pedanou and Berjot, 1988; Pelton and McLean, 2000; Kumasinski and Farrell, 1993).

<table>
<thead>
<tr>
<th>Conformation</th>
<th>Frequency (cm⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-helix</td>
<td>1650-1657</td>
</tr>
<tr>
<td>Antiparallel β-sheet</td>
<td>1612-1640</td>
</tr>
<tr>
<td>Parallel β-sheet turn</td>
<td>1655-1696</td>
</tr>
<tr>
<td>Unordered random coil</td>
<td>1640-1651</td>
</tr>
</tbody>
</table>
Based on this information secondary structure of acquired spectra was studied by initially isolating the Amide-I band and then determining % secondary structure by dividing the area under the curve for allocated conformation by the entire area of the Amide-I band. This acts as an approximation to quantify the amount of individual motifs and the changes the protein undergoes with a change in pH. In general all proteins have a secondary structure composition of $35 \pm 2\%$

Table 4. Percentages of protein secondary structure of proteins at pH 3, pH 6 and pH 9

<table>
<thead>
<tr>
<th>Protein</th>
<th>Conformation</th>
<th>pH 3</th>
<th>pH 6</th>
<th>pH 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>WPI</td>
<td>α-helix</td>
<td>8.93</td>
<td>8.7</td>
<td>8.52</td>
</tr>
<tr>
<td></td>
<td>β-sheet</td>
<td>36.44</td>
<td>36.28</td>
<td>36.96</td>
</tr>
<tr>
<td></td>
<td>β-turn</td>
<td>43.6</td>
<td>43.23</td>
<td>42.61</td>
</tr>
<tr>
<td></td>
<td>Random coil</td>
<td>11.03</td>
<td>11.79</td>
<td>12.01</td>
</tr>
<tr>
<td>Casein</td>
<td>α-helix</td>
<td>8.37</td>
<td>8.34</td>
<td>8.91</td>
</tr>
<tr>
<td></td>
<td>β-sheet</td>
<td>36.37</td>
<td>37.36</td>
<td>30.5</td>
</tr>
<tr>
<td></td>
<td>β-turn</td>
<td>43.42</td>
<td>42.58</td>
<td>48.41</td>
</tr>
<tr>
<td></td>
<td>Random coil</td>
<td>11.84</td>
<td>11.73</td>
<td>12.18</td>
</tr>
<tr>
<td>Egg Albumin</td>
<td>α-helix</td>
<td>8.67</td>
<td>8.37</td>
<td>8.42</td>
</tr>
<tr>
<td></td>
<td>β-sheet</td>
<td>35.76</td>
<td>35.29</td>
<td>36.06</td>
</tr>
<tr>
<td></td>
<td>β-turn</td>
<td>43.91</td>
<td>44.67</td>
<td>43.71</td>
</tr>
<tr>
<td></td>
<td>Random coil</td>
<td>11.66</td>
<td>11.68</td>
<td>11.81</td>
</tr>
<tr>
<td>MPI</td>
<td>α-helix</td>
<td>8.61</td>
<td>10.32</td>
<td>10.22</td>
</tr>
<tr>
<td></td>
<td>β-sheet</td>
<td>37.5</td>
<td>31.57</td>
<td>34.21</td>
</tr>
<tr>
<td></td>
<td>β-turn</td>
<td>43.3</td>
<td>46.14</td>
<td>43.03</td>
</tr>
<tr>
<td></td>
<td>Random coil</td>
<td>10.59</td>
<td>11.96</td>
<td>12.83</td>
</tr>
<tr>
<td>SPI</td>
<td>α-helix</td>
<td>8.14</td>
<td>9.09</td>
<td>18.15</td>
</tr>
<tr>
<td></td>
<td>β-sheet</td>
<td>36.79</td>
<td>32.34</td>
<td>36.03</td>
</tr>
<tr>
<td></td>
<td>β-turn</td>
<td>45.03</td>
<td>47.08</td>
<td>43.46</td>
</tr>
<tr>
<td></td>
<td>Random coil</td>
<td>10.05</td>
<td>11.49</td>
<td>12.47</td>
</tr>
<tr>
<td>PPI</td>
<td>α-helix</td>
<td>11.16</td>
<td>10.1</td>
<td>8.36</td>
</tr>
<tr>
<td></td>
<td>β-sheet</td>
<td>33.91</td>
<td>35.58</td>
<td>35.52</td>
</tr>
<tr>
<td></td>
<td>β-turn</td>
<td>46.21</td>
<td>44.75</td>
<td>43.48</td>
</tr>
<tr>
<td></td>
<td>Random coil</td>
<td>9.59</td>
<td>10.51</td>
<td>12.64</td>
</tr>
<tr>
<td>Faba Bean</td>
<td>α-helix</td>
<td>8.96</td>
<td>9.1</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>β-sheet</td>
<td>35.75</td>
<td>31.73</td>
<td>34.47</td>
</tr>
<tr>
<td></td>
<td>β-turn</td>
<td>45.2</td>
<td>47.36</td>
<td>44.12</td>
</tr>
<tr>
<td></td>
<td>Random coil</td>
<td>10.09</td>
<td>11.81</td>
<td>12.21</td>
</tr>
</tbody>
</table>
β-sheet and 9 ± 2% α-helix structures (Wong et al., 2009) which was also see in Table 4. With an increase in pH there was an increase in the random coil structures of the proteins thus indicating unfolding of the protein structure. As seen most predominantly in the spectrum of PPI, the Amide-I band underwent a shift in the peak region which suggests a transition from β-sheet to random coil also seen in other proteins understood from the Table above. Normally, most proteins tend to be stable near their isoelectric point, where repulsive forces are low and the proteins remain in their native state. Large net charges are induced at pHs far from the isoelectric point and hence proteins tend to partially unfold due to intramolecular side-chain charge repulsion leading to rupture of hydrogen bonds and a breakup of hydrophobic interactions (Ma, Rout, Chan, and Phillips, 2000).

3.5 Conclusion

This chapter discusses the various methods employed to study and predict the secondary structure of the protein using Normal Raman and SERS. Silver based substrates AgNPs were not taken into consideration due to its failure to provide spectrum in previous chapter at pH 3. Spectral data for whey protein isolate and mung protein isolate were classified and analyzed for identifying the characteristic Normal Raman and SERS signals to identify a potential method. Based on the results acquired from these two proteins, it was concluded that Normal Raman was the better method for studying protein structure with changing pH and was hence applied to 5 other proteins. The similarities and differences in the spectrum of both classes of proteins were then understood by applying PCA for better clarification that distinguished the proteins. The effect of change in pH was the most apparent for pea protein compared to others while PCA showed the effect of pH on the protein spectra in more detail. The secondary structure analysis proved to be a quantitative way to understand the change in the secondary structure of the protein by observing the increase in amount of random coil structures with an increase in pH. Future work includes the study of effect of change in ionic strength on the secondary structure of the protein and identify a method for the same.
CHAPTER 4

STUDY EFFECT OF CHANGE IN SALT CONCENTRATION ON SPECTROSCOPIC CHARACTERIZATION OF ANIMAL AND PLANT PROTEINS USING RAMAN SPECTROSCOPY

4.1 Introduction

In Chapter 3, we established a method to allocate functional groups and studied the effect of change in pH on the structure of proteins. We illustrated the merits and demerits of using Raman Spectroscopy for structural analysis with and without substrates and concluded that Normal Raman was the most successful method to identify peaks and understand the change in pH. Herein, we would forward this research to establish a potential method to study the effect of change in salt concentration and quantify it. The focus would majorly be on studying the differences in the spectrum of protein subject to change in salt concentration and how compatible would the substrates be to detect protein signals in varied environments.

4.1.1 Effect of change in salt concentration on protein

Salt concentration is an important factor to consider when studying the environmental conditions of proteins. Electrostatic shielding is an important effect imparted by changing the ionic strength of the medium and influences the protein functionality. It depends majorly on the ionic strength and not the nature of the ion. Salts are also seen to induce ion-specific effects in hydrophobic interaction and affect the stability of higher concentration (Qin, L., 1989). The ionic strength of a solution also determines the overall charge of the protein molecule. Ionic strength affects protein solubility (salting-in or salting-out) dependent on the hydrophilicity-hydrophobicity characteristics of the protein surface (Ismail et al., 2013).

4.1.2 Objectives of the study

The objectives of this study were to (1) determine a method to study structural changes in the protein due to change in salt concentration for whey and mung protein using Raman and SERS
(2) apply method to 5 other proteins (casein, egg albumin from chicken egg white, soy, pea and faba bean protein) and (3) compare and differentiate protein spectra using PCA.

4.2 Materials and methods

4.2.1 Materials

Protein isolates with protein content above 90% were acquired from following commercial sources. Whey protein isolate (WPI) with 93.5% protein was purchased from Agropur (La Crosse, WI), k-casein and Albumin from chicken egg white with 99.99% protein content from Sigma Aldrich (St Louis, MO), Soy Protein Isolate (SPI) (SUPRO EX 45) with 92% protein from Essex Food Ingredients (Frazer, PA), Mung bean protein isolate (MPI) with 93% protein from Scouler (Omaha, NE), Pea protein isolate (PPI) and Faba bean protein isolate with 92.1% protein from AGT Foods (Saskatoon, SK, Canada).

0.02 mg/ml 40nm bare (citrate) AgNPs solution and 0.2 mg/ml 60nm citrate coated AgNPs solution was purchased from NanoComposix (San Diego, CA). For silver dendrites Zn foil (99.99%) and silver nitrate (99%) was acquired from Fischer Scientific (Rochester, NY).

4.2.2 Sample Preparation

The limit of detection is different for Raman and SERS and hence two different methods of sample preparation have been adopted with different concentration of protein. For Normal Raman, 1% solution of protein was mixed with phosphate buffer with designated ionic strength made using NaCl (0.1M, 0.5M & 0.9M) in equal parts. After incubating for 15 mins an aliquot (10uL) was dropped on the gold slide and allowed to air dry. The sample preparation of SERS used 100ppm protein solution mixed with designated ionic strength phosphate buffer in 2:1 proportion for 5 mins. Then 1 ml of the solution was transferred to an Eppendorf tube and 20uL of Ag substrate was added and pipette mixed for 1 min. The substrate was then allowed to settle to the bottom of the Eppendorf tube taking about 15-20 mins assuring enough time for the protein interaction with the nanostructures. Then 10uL of precipitate under the bottom was deposited on the gold slide and
air dried for further analysis. The preparation method for each is demonstrated in Figure 14, Raman spectra were collected individually for each of the methods.

4.2.3 Raman Instrument and Data Analysis

The aliquots were analyzed using a DXR Raman microscope (Thermo Fischer Scientific, Madison, WI). The instrument is equipped with a 780 nm excitation laser and a 20X objective.

![Sample preparation method diagram](image)

Figure 15. Schematic illustration of the sample preparation method

Due to different sensitivities of each of the methods, the laser power for Normal Raman and SERS was set differently. A 24mW laser power was used to collect spectra for Normal Raman while SERS spectra was collected using a laser power of 4mW. The Raman instrument was controlled by OMNIC™ software (version 9.1). Ten spectra were selected from each of the aliquots and then averaged by OMNIC™ software. Later the grouped spectra were analyzed by Thermo Scientific TQ Analyst (version 8.0). All Raman spectra were calculated from at least three replicates. Standard normal variant, second derivative transformation, and smoothing were applied when necessary to
reduce spectral noise, normalize the spectra (against phenylalanine band at 1004 cm\(^{-1}\)), separate overlapping bands, and remove baseline shifts. An averaged spectrum was generated by the software for each sample (representing each treatment or control).

Principal component analysis (PCA), a feature of TQ Analyst to discriminate data was used to analyze the variances of the spectra. For PCA analysis, the spectra were pre-processed using secondary derivative transformation with Norris derivative filter (Segment length 7–11 cm\(^{-1}\), gap between segment 7–11 cm\(^{-1}\), and then used Standard normal variant based on the region (400-2000 cm\(^{-1}\)), which include the phenylalanine peak (Ismail et al., 2013). It is possible to substantially reduce a multidimensional dataset to its most dominant features, eliminate random variation and better understand the variation between spectra through the retained principal components (PCs). This information can potentially give an estimate of the variance within a class and between different classes. If two clusters have no overlap, it means that the two samples are significantly different from each other (Zheng, Zhao, Tian and He, 2017).

4.3 Result and discussion

4.3.1 Characterization of Raman and SERS with Ag substrates spectra of whey and mung protein at different salt concentrations

To study the effect of SERS, spectrum of WPI was acquired using various silver dendrites as also without any substrate as represented in Fig. 15. Like section 2.3.1 AgNPs were not taken into consideration due to its established incapability to produce signals and identify the signature phenylalanine peak with phosphate buffers. Data was acquired at I=0.9 to see if any of the methods hold potential to study the structure of proteins in different salt concentrations.
WPI with Normal Raman was able to detect most of the characteristic peaks even at ionic strength of 0.9 as mentioned in section 2.1.1. and hence could be a potential method for structure analysis. However, Ag dendrites gave significantly good results as it was able to successfully detect the peaks that could provide information about the protein backbone structure. Since SERS has more sensitivity, a proper interaction of the protein with the substrate can yield better and clear peaks. It successfully identified the Amide-I and Amide-III peaks along with the CN stretch and tryptophan peaks that were mentioned in Table 2.

Having studied an animal protein, it is also important to understand if this method could be applied to plant proteins as well. To this end, spectrum of MPI was acquired using silver dendrites as also without any substrate as represented in Fig. 16. MPI with Normal Raman was able to detect most of the characteristic peaks mentioned in section 2.1.1. and hence could be a possible method for peak detection. Ag dendrites gave significantly good results as it was able to successfully detect the peaks that could provide information about the protein backbone structure.
It was beneficial to use Ag dendrites since the background does not have any overlapping peaks with the protein structure that would interfere with the protein spectrum. It is also seen that the spectrum generated using dendrites not just identified the phenylalanine, Amide I and Amide III but also tryptophan and CN stretch in the range of 600-900 cm\(^{-1}\) under similar conditions as the Normal Raman.

![Graph showing representative spectra of MPI with Normal Raman, Ag dendrites background and MPI with Ag dendrites at ionic strength 0.9.](image)

Figure 17. Representative spectra of MPI with Normal Raman, Ag dendrites background and MPI with Ag dendrites at ionic strength 0.9

Based on this data we can conclude that SERS using Ag dendrites could be a potential method to study the effect of change in salt concentration and will now be applied to other proteins.
4.3.2 SERS analysis to study the effect of change in salt concentration on structure of proteins

Based on the results acquired in section 4.3.1, it can be said that SERS with Ag dendrites is a better method for structural analysis for both WPI and MPI at ionic strength 0.9. Hence, the method was further applied to other proteins using the same protocol to generate spectra at ionic strength 0.2, 0.5 and 0.9 (Fig. 17, 18 & 19). All proteins depicted a strong phenylalanine peak across which the spectra is normalized and baseline corrected. The spectrum of all proteins except MPI showed a wide amide-I band at 1674 cm\(^{-1}\) with a peak width of approximately 50 cm\(^{-1}\) for all of the proteins. It is important to note that the asymmetry of the amide I band along with the peak height and width showcase a distribution of secondary structures. The band at 1665 cm\(^{-1}\) is a marker band for alpha-sheet and the shoulder at 1653 cm\(^{-1}\) is a marker band associated with R-helical peptide bonds (Maiti et al, 2003).

![Figure 18. Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at I=0.2](image.png)
Figure 18. Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at I=0.5

Figure 19. Representative SERS spectra of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate at I=0.9
All proteins except casein also projected information on the secondary structures through the amide III region of the spectrum. It is seen to attain maximum height at about 1252 cm\(^{-1}\), attributed to the undefined secondary structures in the protein solution. However, there were some peaks that were only identified in plant protein spectra. The tryptophan peaks (700-800 cm\(^{-1}\)) in the plant proteins that tended to be absent in the animal proteins was a significant difference. These peaks show that the buried tryptophan residues from a hydrophobic microenvironment became exposed which was not observed in animal proteins. The S-S stretches originating from cystine for the animal-based (510 cm\(^{-1}\)) varied in all proteins across different ionic strengths. All proteins were affected by a change in salt concentration and the peaks represented at ionic strength 0.9 were the most distinct. At lower salt concentrations the electrostatic shielding effect comes into picture that results in less functional groups being exposed.

4.3.3 Principal Component analysis of proteins at ionic strength 0.2, 0.5 and 0.9

As discussed in the previous section, it has been seen that the spectrum of animal and plant proteins has certain similarities and dissimilarities. However, it can be difficult to summarize the difference based on visible Raman spectra pattern and peaks. To enable a better relevance between experimental data and the structure of proteins, an innovative non-linear model is necessary to be established. Therefore, we analyzed the Raman data with principal component analysis (PCA). Generally, the PCA plot is able to show significant statistical discrimination of spectra from different data groups and based on this principle we shall attempt to see differences between animal and plant-based proteins at different salt concentrations. (Fig 20, 21 & 22).

At ionic strength 0.2 all animal proteins and plant proteins except SPI and PPI could be distinguished from each other. The structures of SPI and PPI at overlaps in their spectrum hinting that the proteins were not structurally that different from each other which were in the same plane as egg albumin but could be distinguished. The overlapping in these 3 proteins became more prominent at ionic strength 0.9. The datapoints of other proteins always remained clustered except
casein. This is important note and see if this behavior is observed for other proteins at higher ionic strengths as well.

Figure 20. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate ionic strength 0.2

Figure 21. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate ionic strength 0.5
Figure 22. Principal component scores 3D Display model of WPI, casein, egg albumin, MPI, SPI, PPI, and faba bean protein isolate ionic strength 0.9

The proteins seem to vary largely from each other at ionic strength 0.5 where clusters became more scattered for the plant proteins. This signifies that the said ionic strength could be a transition phase for proteins dominating over the electrostatic shielding effect of the salt.

4.4 Conclusion

This chapter discusses the various methods employed to study and predict the secondary structure of the protein using Normal Raman and SERS with different ionic strengths. Silver based substrates AgNPs were not taken into consideration due to its failure to provide spectrum in Chapter 2. Spectral data for whey protein isolate and mung protein isolate were classified and analyzed for identifying the characteristic Normal Raman and SERS signals to identify a potential method. Based on the results acquired from these two proteins, it was concluded that SERS with Ag dendrites was the better method for studying protein structure with changing salt concentration and was hence applied to 5 other proteins. The similarities and differences in the spectrum of both classes of proteins were then understood by applying PCA for better clarification that distinguished
the proteins. The effect of change in ionic strength was the most apparent for all proteins with the clusters being the most scattered at an ionic strength of 0.5 signifying the transition of the proteins overcoming the electrostatic shielding effect. Future work includes the study of proteins in commercial food products and analyze their secondary structure and compare it to the spectrum of proteins acquired so far.
CHAPTER 5
APPLY DEVELOPED METHOD TO STUDY THE SPECTROSCOPIC
CHARACTERISTICS OF PLANT PROTEIN IN A FOOD PRODUCT OF ANIMAL
PROTEIN ALTERNATIVES

5.1 Introduction

In chapter 2, we established a method to study the structure of proteins using SERS with Ag dendrites being the optimal substrate and supported it with advanced data analysis. The results acquired were promising as we could get an insight on the functional groups that govern the structure of the protein that enabled us better basis for prediction. It would be interesting to see if this research can be applied to proteins in the food matrix in real-world products so information can be acquired about the backbone structure of proteins. Here, we would investigate protein isolated from alt-protein food products and compare it with its corresponding animal protein using SERS and discuss if a correlation could be drawn between their structures.

5.1.1 Proteins in the Food Matrix

It can be said that the main contributors to the textural attributes and functional properties to foods are proteins and these parameters help to determine the quality of this food. The importance of food proteins in functional properties (solubility, extractability, viscosity, water holding capacity, etc.) and texture is to a great extent in its native state. This holds true for both animal and plant proteins. In animal-based foods the protein decides the functionality of the product whereas in plant-based foods it is a combination of the protein and the gelling agent. One also needs to note that the structure of proteins is stabilized by different types of interactions: covalent and hydrogen bonds, hydrophobic interactions and plays an important role in protein functionality.
5.1.2 Objectives of this study

The objectives of this study were to (1) apply SERS with Ag dendrites to acquire spectrum and study the secondary structure of isolated proteins (2) compare and differentiate protein spectra using PCA.

5.2 Materials and methods

5.2.1 Materials

Protein isolate and food product was acquired from following commercial source. Albumin from chicken egg white with 99.99% protein was purchased from Sigma Aldrich (St Louis, MO), Mung bean protein isolate (MPI) with 93% protein from Scouler (Omaha, NE) and Just Egg formula from Whole Foods.

For silver dendrites Zn foil (99.99%) and silver nitrate (99%) was acquired from Fischer Scientific (Rochester, NY).

5.2.2 Sample Preparation

There is an additional step compared to other objectives of protein isolation. Initially 10mL of the Just Egg formula was taken to the pI of MPI which is a pH of 4.5 using 2M HCl and kept for 12hrs to ensure complete precipitation of the protein. The formula was centrifuged for 15mins at 4C and the precipitate was washed using DI water. The precipitate was then neutralized to pH 7 and freeze-dried at -25C for 18hrs to obtain the protein isolate. This protocol yields isolate with 90% protein content (Hadnadev et, al., 2017). The protein isolation method is presented in Figure 23. The sample preparation further remains the same after the protein is isolated as 100ppm protein solution is mixed with pH 3 phosphate buffer in 2:1 proportion for 5 mins. Then 1 ml of the solution was transferred to an Eppendorf tube and 20uL of Ag substrate was added and pipette mixed for 1 min. The substrate were then allowed to settle to the bottom of the Eppendorf tube taking about 20 mins assuring enough time for the protein interaction with the dendrites. Then 10uL of precipitate under the bottom was deposited on the gold slide and air dried for further analysis. The
Figure 23. Schematic illustration of protein isolation method

1. Take 10 ml of Just Egg formula
2. Adjust pH to 4.5 with 2M HCl
3. Centrifuge for 15 mins at 4C
4. Wash ppt with DI water for 24hr. Neutralize
5. Freeze dry for 18 hrs
6. MPI
7. 100ppm protein + buffer (2:1)
8. 20uL Ag substrates
9. Drop on Gold Slide

Figure 24. Schematic illustration of the sample preparation method
preparation method for each is demonstrated in Figure 24. Raman spectra were collected individually for each of the proteins.

5.2.4 Raman Instrument and Data Analysis

The aliquots were analyzed using a DXR Raman microscope (Thermo Fischer Scientific, Madison, WI). The instrument is equipped with a 780 nm excitation laser and a 20X objective confocal microscope, 50um slit width, 10 s integration time and a spectral range of 400-2000 cm\(^{-1}\). Due to high sensitivity of SERS, the laser power used was 4mW. The Raman instrument was controlled by OMNIC™ software (version 9.1). The Raman instrument was controlled by OMNIC™ software (version 9.1). Ten spectra were selected from each of the aliquots and then averaged by OMNIC™ software. Later the grouped spectra were analyzed by Thermo Scientific TQ Analyst (version 8.0). All Raman spectra were calculated from at least three replicates. Standard normal variant, second derivative transformation, and smoothing were applied when necessary to reduce spectral noise, normalize the spectra (against phenylalanine band at 1004 cm\(^{-1}\)), separate overlapping bands, and remove baseline shifts. An averaged spectrum was generated by the software for each sample (representing each treatment or control).

Principal component analysis (PCA), a feature of TQ Analyst to discriminate data was used to analyze the variances of the spectra. For PCA analysis, the spectra were pre-processed using secondary derivative transformation with Norris derivative filter (Segment length 7–11 cm\(^{-1}\), gap between segment 7–11 cm\(^{-1}\)), and then used Standard normal variant based on the region (400-2000 cm\(^{-1}\)), which include the phenylalanine peak (Ismail et al., 2013). It is possible to substantially reduce a multidimensional dataset to its most dominant features, eliminate random variation and better understand the variation between spectra through the retained principal components (PCs). This information can potentially give an estimate of the variance within a class and between different classes. If two clusters have no overlap, it means that the two samples are significantly different from each other and vice versa (Zheng, Zhao, Tian and He, 2017).
5.3 Result and Discussion

5.3.1 Comparison of Just Egg MPI and egg albumin

To study the proteins in the food matrix, SERS spectrum of MPI was acquired using Ag dendrites and compared with the spectrum of egg albumin and MPI as seen in Fig. 25. We can see that a high intensity signature phenylalanine peak (1004 cm\(^{-1}\)) peak is present in the extracted protein. Apart from the phenylalanine peak other characteristically significant peaks like the Amide I and Amide III were also present. Based on literature, we know that these bands are a crucial part of the protein backbone structure and hence we can confirm protein interaction with the substrate. On comparing the spectra, it is seen that there is significant difference in the MPI and the extracted MPI. The extracted MPI did not show the CN and tryptophan peaks in the 600-900 cm\(^{-1}\) region which were also absent in the spectrum for egg albumin. The distribution of the amide-I region was also significantly different for the Just Egg MPI compared to the isolate itself. It can also be seen that the peak height of the 900 CN peak was also reduced in case of the extracted MPI thus hinting

![Figure 25. Representative spectra of Egg Albumin, MPI and extracted MPI from Just Egg](image)

52
that the processing conditions that the protein was subjected to while product preparation may have altered the structure of the protein that it now resembles the structure of egg albumin. Results from Principal Component Analysis will better help us understand the similarities and dissimilarities in the spectrum by observing the overlapping of datapoints.

5.3.2 Principal Component Analysis of Proteins

As discussed in the previous section, it has been seen that the spectrum of albumin and both the MPI has certain similarities and dissimilarities. However, it can be difficult to summarize the difference based on visible SERS spectra pattern and peaks. To enable a better relevance between experimental data and the structure of proteins, an innovative non-linear model is necessary to be established. Therefore, we analyzed the SERS data with principal component analysis (PCA). Generally, the PCA plot can show significant statistical discrimination of spectra from different data groups and based on this principle we shall attempt to see differences between the proteins (Fig 26.)

![Figure 26. Principal component scores 3D Display model of Egg Albumin, MPI and extracted MPI from Just Egg.](image-url)
It was seen that the datapoints of MPI were all clustered and could be very well distinguished from the extracted MPI that was in a different plane. This showed that the protein spectrum had vastly changed from the commercial MPI. As seen and expected from the spectrum the datapoints of the extracted MPI were close to the egg albumin, however there were no overlaps in the cluster of both the proteins. This indicates that there is significant statistical difference in the proteins, but the degree was understood with there being no overlaps between the commercial MPI and the one extracted from the Just Egg formula. Hence, both the SERS spectra and the PCA model gave us information on the structure of the proteins.

5.4 Conclusion

This chapter discusses the various methods employed to study and understand the secondary structure of the protein in the food matrix; here MPI in the Just Egg formula using SERS. Silver based substrate Ag dendrites was used for the optimization of this approach SERS signals of isolates from egg albumin, mung and mung extracted from Just Egg commercial product. A large amount of spectral data was classified and analyzed for identifying the characteristic. It was seen that the spectrum of MPI had more peaks than that of the extracted MPI signifying some alterations in the structure of protein during processing. The similarities and differences in the spectrum of the proteins were then understood by applying PCA for better clarification. It was observed that some peaks were exclusive to each class of protein and that they could be distinguished from each other especially both the MPIs from both commercial source and from food product. In further study we could assess other proteins from more commercial products as well as try other noble-metal substrates.
CHAPTER 6

CONCLUSION

This work has established potential methods to rapidly study the secondary structure of animal and plant proteins using Raman Spectroscopy or SERS. SERS using Ag dendrites seemed to be the more reliable method throughout this research for peak detection and discrimination of proteins due to its high sensitivity. However, that did not seem to be the case for environments with extreme pHs where the functionality of the substrate is decreased as it does not depict any protein peaks especially the relatively inert phenylalanine (1004 cm\(^{-1}\)) peak. Unlike Ag dendrites which was not compatible in selective conditions AgNPs were not compatible with the proteins in any of the conditions as it failed to generate a spectrum. It was observed that some peaks were exclusive to each class of protein viz found in plant but not in animals and vice versa. Applying PCA showed us the similarities and dissimilarities between the protein spectra, especially distinguishing plant from animal proteins. The spectra of plant proteins like faba bean and PPI tend to overlap in most cases signifying their structural similarity which was not seen in any of the animal proteins that were largely different from each other. This work was also forwarded to the proteins in extracted proteins from real-world commercial products. It was seen that the structure undergoes a significant change under processing conditions that result in the spectrum of extracted protein being different from the original and more like the protein it intends to mimic.

Further work in this research can involve utilization of other methods and substrates to enhance protein signals and study the secondary structure of proteins. The environmental conditions varied in this research were pH and salt, however parameter like temperature and enzyme effect can also be studied. Moreover, continued study for alternative protein using Raman Spectroscopy seems to have a promising future.
REFERENCES


22) Hatzakis, E. (2019). Nuclear magnetic resonance (NMR) spectroscopy in food science: A comprehensive review. *Comprehensive reviews in food science and food safety, 18*(1), 189-220.


