

INVESTIGATION OF STRUCTURAL AND FUNCTIONAL PROPERTIES OF
OMPG MUTANTS BY FTIR SPECTROSCOPY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

İREM YILMAZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
PHYSICS

AUGUST 2017

Approval of the Thesis:

**INVESTIGATION OF STRUCTURAL AND FUNCTIONAL PROPERTIES
OF OMPG MUTANTS BY FTIR SPECTROSCOPY**

submitted by **İREM YILMAZ** in partial fulfillment of the requirements for the degree
of **Master of Science in Physics Department, Middle East Technical University**
by,

Prof. Dr. Gülbin Dural Ünver
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Altuğ Özpineci
Head of Department, **Physics** _____

Assoc. Prof. Dr. Alpan Bek
Supervisor, **Dept. of Physics, METU** _____

Assoc. Prof. Dr. Filiz Korkmaz Özkan
Co-supervisor, **Dept. of Electrical and Electronics Eng.,
Atılım University** _____

Examining Committee Members:

Assoc. Prof. Dr. Kemal Efe Eseller
Dept. of Electrical and Electronics Eng., Atılım University _____

Assoc. Prof. Dr. Alpan Bek
Dept. of Physics, METU _____

Assoc. Prof. Dr. Serhat Çakır
Dept. of Physics, METU _____

Assoc. Prof. Dr. Hande Toffoli
Dept. of Physics, METU _____

Assist. Prof. Dr. Emre Yüce
Dept. of Physics, METU _____

Date: 18.08.2017

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: İREM YILMAZ

Signature :

ABSTRACT

INVESTIGATION OF STRUCTURAL AND FUNCTIONAL PROPERTIES OF OMPG MUTANTS BY FTIR SPECTROSCOPY

Yılmaz, İrem
M.Sc., Department of Physics
Supervisor: Assoc. Prof. Dr. Alpan Bek
Co-Supervisor: Assoc. Prof. Dr. Filiz Korkmaz Özkan

August 2017, 93 pages

The subject of the current study is a bacterial protein OmpG from *Escherichia coli*. OmpG-16S and OmpG-16SL are the mutants of OmpG that are genetically engineered for potential biotechnological applications. In this study, structural and functional similarities and/or differences of these mutants with respect to the wild type OmpG were revealed using IR spectroscopy. Also, a new spectrometer attachment enabling in situ $^1\text{H}/^2\text{H}$ exchange was developed. For the collection of mass amount of spectral data from each experiment, an automated data collection computer code was developed as well.

Keywords: IR spectroscopy, ATR-FTIR spectroscopy, Outer Membrane Protein (OMP), Porin, Genetically engineered protein.

ÖZ

OMPG MUTANTLARININ YAPISAL VE FONKSİYONEL ÖZELLİKLERİNİN FTIR SPEKTROSKOPİSİ İLE İNCELENMESİ

Yılmaz, İrem
Yüksek Lisans, Fizik Bölümü
Tez Yöneticisi: Doç. Dr. Alpan Bek
Ortak Tez Yöneticisi: Doç. Dr. Filiz Korkmaz Özkan

Ağustos 2017, 93 sayfa

Bu çalışmanın konusu *Escherichia coli* bakterisine ait OmpG proteinidir. Bu proteinin potansiyel biyoteknolojik uygulamaları olabilecek OmpG-16S ve OmpG-16SL mutantları ise genetik mühendisliği ile tasarlanmıştır. Bu çalışmada mutantların doğal yapıda olan OmpG ile olan yapı ve fonksiyon benzerlik/farklılıkları IR spektroskopisi kullanılarak araştırılmıştır. Ayrıca yerinde $^1\text{H}/^2\text{H}$ değişimi yapabilen bir spektrometre ataçmanı da geliştirilmiştir. Bunun yanında her deneyde ortaya çıkan büyük miktarda spektral bilgiyi otomatik olarak toplayabilmeye yarayan bir bilgisayar kodu da geliştirilmiştir.

Anahtar Kelimeler: IR spektroskopisi, ATR-FTIR spektroskopisi, Dış Zar Proteini, Porin, Gen mühendisliği ürünü protein.

*Dedicated to the memory of
my beloved grandfather,
Bayram HANÇER*

*I wonder every day
how my life would be with you...*

ACKNOWLEDGMENTS

First of all, I would like to thank my supervisor Assoc. Prof. Dr. Alpan Bek for giving me this opportunity. His enthusiasm, patience and tolerance made me continue at hard times throughout this thesis. He always welcomed me whenever I knocked on his door and I have always felt so grateful for the small but fruitful talks he shared spontaneously.

Second of all, I would like to thank my co-supervisor Assoc. Prof. Dr. Filiz Korkmaz Özkan for providing me numerous opportunities, and confiding in me. I feel blessed that I had the opportunity of being guided by the person who is such hard-working, passionate, broad-minded and professional. She has not only been an instructor, but also a mentor and a friend to me and will always be a role model for me. I am deeply thankful for the efforts she made.

I would like to thank my parents who have always supported me and given me their never-ending, unconditional love. I am also thankful to my sister and her husband for their confidence in me, and for computer technical support.

I would like to thank Halil İbrahim Aydemir for his companionship at Atılım University.

I thank my dear friends Merve Yağmur Yardımcı, Simay Canlar and Onur Utku Uçak for constant support and soothing words in the time of need. I also express my gratitude to Yeliz Günhan, my cousin Serkan Yılmaz, and Ateş Aliefendioğlu for just being there.

I would like to thank Dr. Özkan Yıldız and Katharina van Pee at the Department of Structural Biology, Max Planck Institute of Biophysics for biological samples.

Finally, I kindly acknowledge financial support from TÜBİTAK under project no: 214Z085.

TABLE OF CONTENTS

ABSTRACT.....	v
ÖZ.....	vi
ACKNOWLEDGMENTS.....	viii
TABLE OF CONTENTS.....	ix
LIST OF TABLES.....	xi
LIST OF FIGURES.....	xii
LIST OF SYMBOLS AND ABBREVIATIONS.....	xv

CHAPTERS

1. INTRODUCTION.....	1
2. BACKGROUND.....	3
2.1 Protein Structure.....	3
2.1.1 Components of Protein.....	3
2.1.2 Chemical Bonding that Creates Protein Stability.....	7
2.1.3 Structural Characteristics of Proteins.....	9
2.2 Protein Function.....	13
2.3 Porin and Outer Membrane Protein G (OmpG).....	13
2.3.1 Porin.....	13
2.3.2 OmpG.....	15
2.4 Fourier Transform Infrared Spectroscopy (FTIR).....	18
2.4.1 Interaction of Electromagnetic Radiation with Matter.....	18
2.4.2 Infrared (IR) Spectroscopy.....	22
2.4.2.1 The Spectrometer.....	25
2.4.2.2 Fourier Transform.....	29
2.4.3 FTIR Vibrational Frequencies of a Protein.....	34
3. MATERIALS AND METHODS.....	37
3.1 Materials.....	37

3.1.1	Chemicals.....	37
3.1.2	Buffers.....	37
3.1.3	FTIR Spectra of H ₂ O and ² H ₂ O.....	39
3.1.4	OmpG-WT and Mutants.....	40
3.2	Methods.....	41
3.2.1	FTIR Spectroscopy.....	41
3.2.2	Data Process of the Spectra in order to Analyze Protein Structure.....	42
3.2.2.1	Buffer Subtraction.....	43
3.2.2.2	Second Derivative.....	43
3.2.2.3	Side Chain Subtraction.....	44
3.2.2.4	Curve-fitting.....	45
3.2.3	ATR-FTIR Spectroscopy.....	48
3.2.3.1	¹ H/ ² H Exchange.....	51
3.2.3.2	Analysis of Spectra in the ¹ H/ ² H Exchange Experiment.....	52
4.	RESULTS AND DISCUSSION.....	55
4.1	Band Assignment on FTIR Absorbance Spectra.....	55
4.2	pH-Dependent Conformational Changes and Secondary Structures of OmpG-WT, OmpG-16S and OmpG-16SL.....	60
4.3	Thermal Stability of OmpG-WT, OmpG-16S and OmpG-16SL.....	71
4.4	¹ H/ ² H Exchange Properties of OmpG-WT, OmpG-16S and OmpG-16SL.....	77
5.	CONCLUSION.....	83
	REFERENCES.....	87
	APPENDIX.....	93

LIST OF TABLES

TABLES

Table 2.1 The chemical characteristics of amino acids along with the three-letter and one-letter codes.....	6
Table 2.2 Types of interactions among molecules or atoms and their corresponding force dependences on the separation within the power ranges.....	9
Table 3.1 Chemicals used in the study.....	37
Table 3.2 Calculated masses of Na_2HPO_4 and NaH_2PO_4 in double distilled water and deuterated water for two different pH and $p^2\text{H}$ values.....	39
Table 3.3 Stock solution properties of OmpG-WT and mutants.....	41
Table 4.1 Protein secondary structure elements and their position ranges in wavenumbers compiled from the literature.....	57
Table 4.2 Amino acids side chain positions in wavenumbers.....	59
Table 4.3 The percentages of the secondary structure elements of OmpG-WT and OmpG-16S in $p^2\text{H}$ 5.5 and $p^2\text{H}$ 7.5 obtained from curve-fitting results are taken from literature.....	67
Table 4.4 The curve-fitting results of OmpG-16SL in the amide I region.....	67
Table 4.5 Transition temperature values of the three proteins at both $p^2\text{H}$ values...	75
Table 4.6 The percentages of $^1\text{H}/^2\text{H}$ exchange of the three proteins for pH 5.5 and 7.5 after 18 hours	78

LIST OF FIGURES

FIGURES

Figure 2.1 A general amino acid structure (non-zwitterionic state).....	4
Figure 2.2 Peptide bond formation (dehydration synthesis). Residues are enumerated in sequence beginning from the left (amino group).....	5
Figure 2.3 Coplanar units and torsion angles between the peptide units in the protein backbone.	8
Figure 2.4 An example to a simple chain of amino acids, which as a whole is called the primary structure of a protein.....	10
Figure 2.5 Secondary structure elements: alpha helix (A), and antiparallel beta sheet (B).....	11
Figure 2.6 An example demonstrating the interactions within a tertiary structure...	12
Figure 2.7 Schematic representation of outer and inner membrane structures (double membrane) of a Gram-negative bacteria (Gram-positive ones do not own an outer membrane).....	14
Figure 2.8 Examples of integral membrane proteins shown in secondary structure forms.....	15
Figure 2.9 Structure of OmpG in the case of open (a,b) and closed (c,d) conformation.....	16
Figure 2.10 Two strands and their connecting loops are shown in acidic pH (purple) and in neutral pH (green).....	18
Figure 2.11 Electromagnetic wave propagating in the x-direction with speed of light.....	19
Figure 2.12 Energy-level diagram demonstrating electronic energy levels and both vibrational and rotational sub-levels.....	20
Figure 2.13 Three modes of vibration of H ₂ O (in gas phase) where ν_1 (3650 cm ⁻¹) is the symmetric stretch, ν_2 (1600 cm ⁻¹) is the bending mode and ν_3 (3750 cm ⁻¹) is the asymmetric stretch.....	25
Figure 2.14 Michelson Interferometer.....	26

Figure 2.15 A diagram of monochromatic light passing through the Michelson interferometer and reaching the detector.....	27
Figure 2.16 Spectral radiance distribution of incident radiation at three different superposition cases and their corresponding interference patterns (interferograms).....	28
Figure 2.17 Basic components inside a commercial FTIR spectrometer.....	29
Figure 2.18 The graphs on the left column are the Fourier transformed spectra of the interferograms on the right column.....	33
Figure 2.19 An interferogram of a sample over data points and its FTIR spectrum (absorbance vs. wavenumbers)	34
Figure 2.20 A representative spectrum showing amide bands and their corresponding molecular groups shown in boxes on the protein backbone in accordance with color coding.....	35
Figure 3.1 The FTIR absorbance spectra of H ₂ O (red) and ² H ₂ O (green).....	40
Figure 3.2 The original absorbance spectrum (upper graph) of OmpG-16S in pH 5.5 and its second derivative spectrum (bottom graph) in the region 1700-1500 cm ⁻¹	44
Figure 3.3 The absorbance spectrum (black) and the amino acid subtracted spectrum (purple) of OmpG-16SL in p ² H 7.5.....	45
Figure 3.4 Lorentzian and Gaussian functions.....	46
Figure 3. 5 An ATR sample cell and the path of the IR beam	49
Figure 3.6 The figure demonstrates the total internal reflection of an incoming IR wave penetrating into the IRE and the evanescent wave established by that reflection.....	49
Figure 3.7 ATR-IR microdialysis cell which is mounted into the FTIR spectrometer.....	50
Figure 3.8 A representative loop written for the second hour of the exchange in Macro and used for automated scanning.....	52
Figure 3.9 The absorbance spectrum of OmpG-16S in pH 7.5 (black line) and in p ² H 7.5 (red line).....	53
Figure 4.1 A characteristic absorbance spectrum of OmpG-16S mutant in deuterated buffer.....	58
Figure 4.2 Absorbance spectra of OmpG-16S at pH 5.5 and pH 7.5 at ~ 22 °C.....	60
Figure 4.3 Absorbance spectra of OmpG-WT at p ² H 5.5 and p ² H 7.5 at ~ 22 °C.....	61

Figure 4.4 Absorbance spectra of OmpG-16S at p ² H 5.5 and p ² H 7.5 at ~ 22 °C.....	62
Figure 4.5 Absorbance spectra of OmpG-16SL at p ² H 5.5 and p ² H 7.5 at ~ 22 °C.....	62
Figure 4.6 Second derivative of OmpG-WT absorbance in p ² H 5.5 and p ² H 7.5 at ~ 22 °C shown in the region 1720-1600 cm ⁻¹	64
Figure 4.7 Second derivative of OmpG-16S absorbance in p ² H 5.5 and p ² H 7.5 at ~ 22 °C shown in the region 1720-1600 cm ⁻¹	64
Figure 4.8 Second derivative of OmpG-16S absorbance in p ² H 5.5 and p ² H 7.5 at ~ 22 °C shown in the region 1720-1600 cm ⁻¹	65
Figure 4.9 Curve-fitting results of OmpG-16SL (1720-1500 cm ⁻¹ region).....	68
Figure 4.10 The second derivative comparisons of the original absorption (main) spectrum and the co-fitted spectrum (1720-1500 cm ⁻¹ region) of OmpG-16SL.....	69
Figure 4.11 IR spectra of OmpG-WT in p ² H 5.5 in the region 1720-1500 cm ⁻¹ during a temperature ramp.....	73
Figure 4.12 The second derivatives of the absorbance IR spectra of OmpG-WT in p ² H 5.5 (IR spectra in Figure 4. 11).....	73
Figure 4.13 T-ramp 3-D plot of protein OmpG-WT in p ² H 5.5 in the region 1720-1500 cm ⁻¹	74
Figure 4.14: Comparative thermal stability graphs of OmpG-WT and mutants in both p ² H values.....	76
Figure 4.15 ¹ H/ ² H exchange percentage profiles of the three proteins.....	79

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

C_{α}	alpha carbon
^2H or D	deuterium
T_t	transition temperature

ABBREVIATIONS

ATR	Attenuated Total Reflection
CD	Circular Dichroism
DTGS	Deuterated Triglycine Sulfate
FWHH	Full width half height
IR	Infrared
IRE	Internal Reflection Element
MCT	Mercury Cadmium Telluride
NMR	Nuclear Magnetic Resonance
OG	n-octyl β -D-glucoside
OmpG	Outer membrane protein G
OPD	Optical path difference
t-ramp	temperature ramp
WT	Wild type
ZPD	Zero path difference

CHAPTER 1

INTRODUCTION

Infrared (IR) spectroscopy has been used for decades by Physicists and Chemists for identifying the constituents of unknown substances and evaluating semiconductor materials. It is utilized in Biophysics for probing the polypeptide backbone and amino acid side chains in proteins among many other applications. In this study, the structure and dynamics of three proteins, OmpG-WT, OmpG-16S and OmpG-16SL were investigated by IR spectroscopy.

OmpG is a bacterial protein having a unique nature. It is a stable monomer creating a pore in the outer membrane of the bacteria *Escherichia coli*. It allows ions and molecules outside the membrane to enter the cell at neutral pH (open state) and blocks the passage of substances at acidic pH (closed state) using one of its loops as the gatekeeper.

Engineering a monomer that can adapt new functions is much more favorable than engineering a protein having sub-units (i.e., trimers) (Conlan & Bayley, 2003, p. 9454). Since OmpG is a stable monomer and is able to attain pH-dependent open and closed states, it is considered to be a good candidate in biotechnological applications (i.e., biosensor applications and in targeted drug delivery systems). Using a protein-based agent rather than a manmade one like carbon nano-tubes will have better chances for being candidates in such applications. Thus, in principle, the first structural alteration was to enlarge the pore diameter while keeping the stability unaltered. With this motivation, two mutants (OmpG-16S and OmpG-16SL) were designed at Max Planck Institute of Biophysics, Frankfurt am Main, Germany.

Protein structures can be determined via different techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), circular dichroism (CD) and FTIR spectroscopy. Every technique provides relatively different kind of information

on protein structure and function when special conditions, depending on the technique, is met for the sample protein. For instance, information provided by X-ray crystallography is at atomic level and it presents much more accurate and detailed structural data if the sample protein can be crystallized. However, NMR determines dynamic (functional) information but with low sensitivity and only for small sized proteins. Furthermore, CD shows conformational changes in the protein structure but the protein sample must be optically active and it is mostly preferred in rapid structure determination since it does not give detailed information. On the other hand, FTIR spectroscopy is advantageous in many aspects. First of all, it is a non-invasive method with very few light scattering problems and with high signal-to-noise ratio. By using few amounts of protein, the environment (the buffer surrounding the protein) can be altered and the conformational changes in the structure can be detected. Also, protein dynamics can be monitored. Besides these upsides, there are downsides of FTIR spectroscopy but they can be handled. For example, complex proteins have overlapping spectral bands but further analysis can be performed on them or water bands overlap the spectra of the proteins but different treatments can be performed on buffers.

Throughout the upcoming pages, the chapters will be divided into sub-sections. The background chapter will contain detailed information on protein structure and function, OmpG structure, light-matter interaction, principles of FTIR spectroscopy, Fourier transform between time-frequency domains, and FTIR vibrational frequencies of proteins. Chapter 3 introduces materials used in experiments, experimental procedures when applying transmission FTIR spectroscopy, and ATR-FTIR spectroscopy along with the methods analyzing the outcomes. In the following chapters, these outcomes will be demonstrated, discussed and brought to a conclusion.

CHAPTER 2

BACKGROUND

2.1 Protein Structure

Every form of life from single-celled organisms to complex forms function and survive using proteins and enzymes as workers in the cellular level. Understanding the conditions and key elements of normal functioning is the first step in diagnosis and cure for many diseases. Examining the structure of proteins is a vital necessity in this regard. Proteins perform many functions that rely heavily on their structure.

Proteins constitute half of the dry weight of a cell. All proteins are flexible and can bind to other molecules due to specific recognition. Some external conditions experienced by a protein may alter its conformation. For example, due to ligand binding or pH change, the altered conformation of the protein may give clues on its functional properties, or the conformational changes due to temperature change may provide information about structural properties of a given protein.

2.1.1 Components of Protein

Proteins are made up of amino acids. 21 different amino acids take part in making of proteins in all livings. In the structure of an amino acid (Figure 2. 1), if the α -Carbon (C_α) is assumed to be the central component, there are four other components conjoint to it.

The four components conjoint to C_α are namely a separate hydrogen atom, an amino group (NH_3^+), a side chain (R group) and a carboxyl group (COO^-). The side chain in the amino acid structure is the only variant component. 21 amino acids found in all organisms are different from each other due to differing side chains.

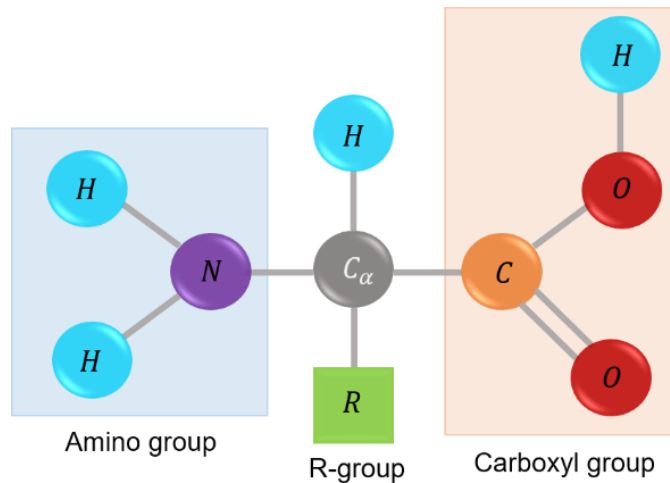


Figure 2. 1: A general amino acid structure (non-zwitterionic state).

Amino acids, linked to each other with peptide bonds by elimination of one water molecule, are called residues (Figure 2. 2). During protein synthesis, in the formation of peptide bonds, (NH)-(C_αH)-(C=O) is the repeating main chain section, called the *protein backbone*.

In order to read an amino acid chain, the series initiate from the unlinked amino group called the N-terminal and end with the carboxyl group at the other end of the sequence, called the C-terminal. The name of one amino acid can be abbreviated to both three-letter and one-letter codes. All amino acids belong to a group because they retain different chemical characters among each other. Hydrophobic amino acids tend to avoid contact with water. Alanine, isoleucine, leucine and valine have hydrophobic-aliphatic side chains; while phenylalanine, tryptophan and tyrosine have hydrophobic-aromatic side chains.

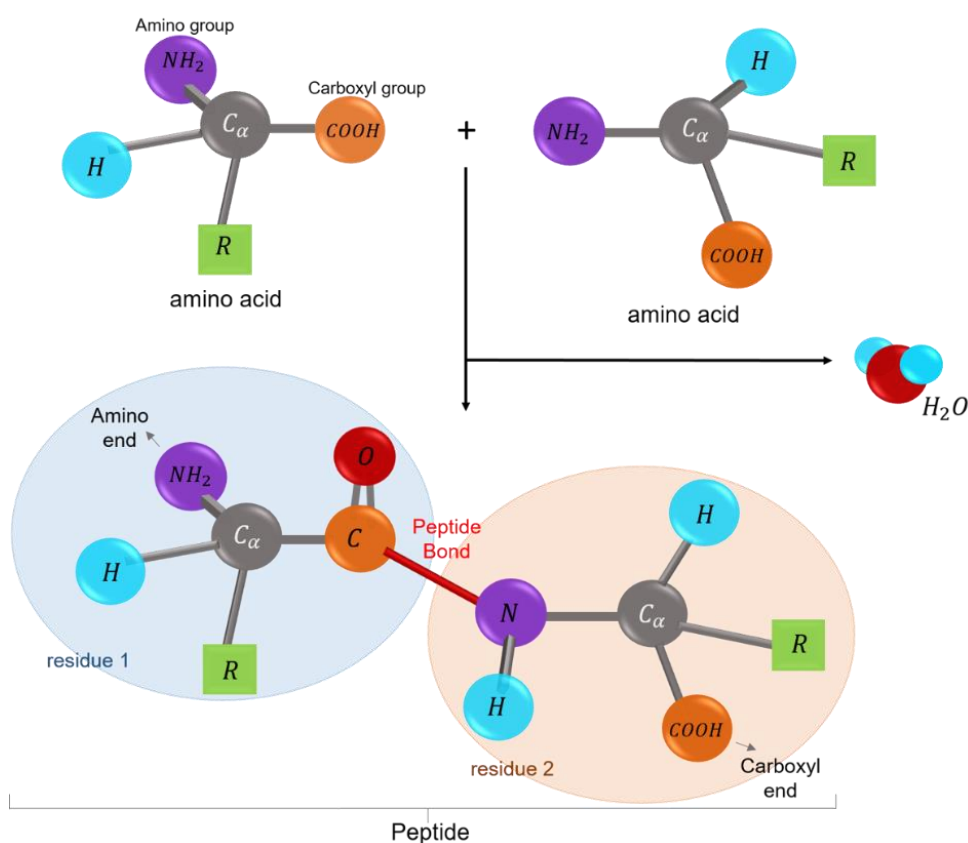





Figure 2. 2: Peptide bond formation (dehydration synthesis). Residues are enumerated in sequence beginning from the left (amino group).

Polarity relates with hydrophilicity. Some very soluble side chains in water tend to make contact with water. Asparagine, cysteine, glutamine, methionine, serine and threonine have neutral-polar side chains. Aspartic acid (aspartate) and glutamic acid (glutamate) have positively charged; and arginine, histidine and lysine have negatively charged side chains. Moreover, there are three amino acids that do not belong to any kind of groups due to their side chains' unique nature; i.e., glycine, proline and selenocysteine. While glycine has only one hydrogen atom as its R-group, proline has no amide group and cannot make hydrogen bond. The structure of selenocysteine can be compared to the structure of cysteine because only sulfur atom in cysteine is replaced with the selenium atom. It is less reducible than cysteine. The structures and codes of these 21 amino acids can be seen in Table 2. 1.

Table 2. 1: The chemical characteristics of amino acids along with the three-letter and one-letter codes.

Chemical Characters of Amino Acids and Amino Acid Codes					
Hydrophobic -Aliphatic Amino Acids	$\begin{array}{c} \text{NH}_2 \\ \\ \text{CH}_3\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Alanine, Ala, A</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{CH}_3\text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Isoleucine, Ile, I</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{CH}_3\text{CHCH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Leucine, Leu, L</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{CH}_3\text{CHCH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Valine, Val, V</i>	
Hydrophobic -Aromatic Amino Acids	$\begin{array}{c} \text{NH}_2 \\ \\ \text{C}_6\text{H}_5-\text{CH}_2-\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Phenylalanine, Phe, F</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$  <i>Tryptophan, Trp, W</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{HO}-\text{C}_6\text{H}_4-\text{CH}_2-\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Tyrosine, Tyr, Y</i>		
Polar Amino Acids	$\begin{array}{c} \text{O} \\ \\ \text{H}_2\text{N}-\text{C}-\text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Asparagine, Asn, N</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{HSCH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Cysteine, Cys, C</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{H}_2\text{N}-\text{C}-\text{CH}_2\text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Glutamine, Gln, Q</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{CH}_3\text{SCH}_2\text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Methionine, Met, M</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{CH}_3\text{CHCH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Threonine, Thr, T</i>
Acidic- Charged Amino Acids	$\begin{array}{c} \text{O} \\ \\ \text{HO}-\text{C}-\text{CH}_2\text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Glutamic Acid, Glu, E</i>	$\begin{array}{c} \text{O} \\ \\ \text{HO}-\text{C}-\text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Aspartic Acid, Asp, D</i>			
Basic- Charged Amino Acids	$\begin{array}{c} \text{NH} \\ \\ \text{H}_2\text{N}-\text{C}-\text{NHCH}_2\text{CH}_2\text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Arginine, Arg, R</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$  <i>Histidine, His, H</i>		$\begin{array}{c} \text{NH}_2 \\ \\ \text{H}_2\text{NCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Lysine, Lys, K</i>	
Particular Amino Acids	$\begin{array}{c} \text{O} \\ \\ \text{NH}_2\text{CH}_2-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Glycine, Gly, G</i>	$\begin{array}{c} \text{O} \\ \\ \text{C}-\text{OH} \\ \\ \text{O} \end{array}$  <i>Proline, Pro, P</i>	$\begin{array}{c} \text{NH}_2 \\ \\ \text{HSeCH}_2\text{CH}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$ <i>Selenocysteine, Sec, U</i>		

2.1.2 Chemical Bonding that Creates Protein Stability

Bonding is related with the electron arrangement of atoms. The stability character of a compound (here, of a protein) depends on the chemical forces between the atoms or to be more precise, between the electrons in these atoms (see Appendix). Protein stability can only be understood by studying chemical bonding.

The covalent bonding (peptide bonds in the protein backbone) holding the residues together in a polypeptide chain possesses partial double bonding (resonance) because the electrons of the C=O and the C-N bond are delocalized. This peptide bonding leads to steadiness since it is a base for explicit properties such as bond angles or bond lengths. C_α, C, O, N and H constitute a coplanar unit and this leads to steadiness. Hence, only two bonds in the protein backbone are rotating. These bonds are formed between C-C_α and N-C_α so that the rotation is restricted and each unit rotates in full. The angles between C-C_α and N-C_α are called psi (Ψ) and phi (Φ), respectively (Figure 2. 3). The amount of possible conformations a polypeptide chain can attain is, thus, considerably limited because of this phenomenon. Another covalent bonded structure in proteins is built up between two R groups of cysteine residues resulting in disulfide bridges.

A noncovalent bond kind, ionic bonding, may be generated when positive and negative charges interact. For example, oppositely charged side chains can hold ionic bonds. The forces arisen between the charges in a biological macromolecule are denoted by Coulomb's Law:

$$F = \frac{1}{4\pi\epsilon} \frac{q_1 \cdot q_2}{r^2}$$

where ϵ is the permittivity of the medium, r is the distance between the two charges (q_1 and q_2).

Van der Waals interactions are in act when atoms get close enough to one another. The existence of random fluctuations in the electron distribution of one atom induces a dipole moment on the nearby atom. Any permanent dipole-dipole interaction ends

up being attractive. However, when two atoms get closer to each other so that at a particular distance their occupied orbitals start to overlap, their interactions become repulsive. For hydrophobic side chains, the effect of these interactions is immense. However, when interacting atoms are detached from each other, the van der Waals interactions lessen. That is why they are considered as weakly noncovalent interactions, yet they contribute noticeably to the stabilization energy of the proteins. Van der Waals interactions between nonpolar (both aliphatic and aromatic) side chains are specially known as hydrophobic effect. These hydrophobic side chains when they encounter with water escape from water molecules and clump together so that they are positioned away from water.

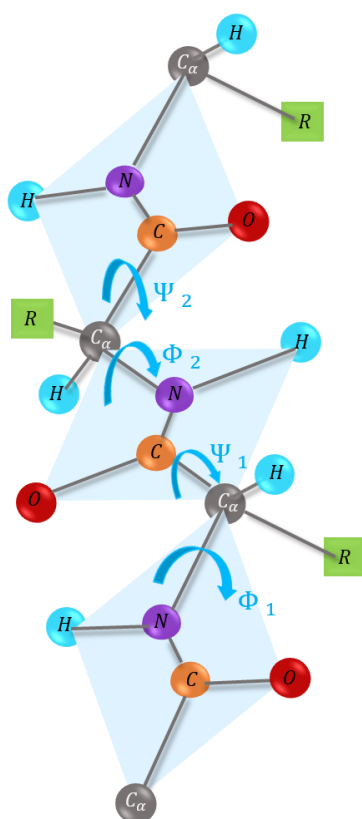


Figure 2. 3: Coplanar units and torsion angles between the peptide units in the protein backbone.

A Hydrogen bonding (another noncovalent bonding) is mainly set up between C=O and N-H groups in the protein backbone giving stability to the protein structure. Moreover, it is not only established between polar side chains and water molecules but also between backbone atoms and water molecules.

The magnitudes of the forces that make bonds are inversely proportional with the intermolecular distances between their molecules/atoms, which are shown in Table 2. 2.

Table 2. 2: Types of interactions among molecules or atoms and their corresponding force dependences on the separation within the power ranges.

Type of Interaction	Force Dependence on Intermolecular Distance
Ion-Ion	$1/r$
Ion-Dipole	$1/r^2$
Dipole-Dipole	$1/r^3$
Ion-Induced Dipole	$1/r^4$
Dipole-Induced Dipole	$1/r^6$
Repulsive Forces	$1/r^{12}$

2.1.3 Structural Characteristics of Proteins

The primary structure of proteins is the amino acid sequence (Figure 2. 4). Residues are covalently bonded in the sequence of amino acids and sequences are uniquely determined by nucleotide sequences (DNA) for all livings.



Figure 2. 4: An example to a simple chain of amino acids, which as a whole is called the primary structure of a protein.

A primary chain can fold into one of three types or may not fold at all, which then is called the unordered structure. The orderly patterns are alpha helix (α -helix), beta sheet (β -sheet) and turn. These variations pack in specific shapes so as to keep hydrophobic side chains away from water contact. Hence, hydrogen bonding is constituted between polar groups of side chains and water molecules at the outer surface of the protein, whereas in the interior region, hydrogen bonding is established between N-H and C=O parts of the protein backbone.

Alpha helices mainly with right-handed configuration, are shaped in spiral figure where C=O groups of residue n make hydrogen bonds with N-H groups of residue $n+4$ (Figure 2. 5). The first N-H and the last C=O parts of the alpha helix do not make hydrogen bonds with each other.

As another type of secondary structure, beta sheets also originate from hydrogen bonding between N-H and C=O groups between different residues of a single primary chain (Figure 2. 5). Even if the strands are distant from each other, they can be formed side by side. There are also outer groups standing on the outer strands without making any hydrogen bonds with each other. These outer N-H and C=O parts of the beta sheets may connect with water, or strands of other peptide chains. If these parts connect with each other like one boundary strand linked to the other boundary strand of the same sheet, the new shape is called beta barrel (resembles a closed cylinder). There are three types of beta sheets namely parallel, antiparallel and mixed. In parallel beta sheets, strands position in the same direction; however, in antiparallel, they stand in the opposite direction. Mixed sheets contain both parallel and antiparallel strands connected. The steadiest hydrogen bonding due to its linearity belongs to antiparallel sheets. The bond angles are generally -130° and $+125^\circ$ for phi and psi, respectively.

Lastly, beta turns are formed by C=O groups of residue n making hydrogen bonds with N-H groups of residue $n+3$. They change the direction of the chains. This direction reversing helps the protein to maintain a regular and tight shape. Moreover, there are uneven patterns called loops which are connectors of secondary structure elements. Long loops are called random coils while short loops are called turns.

When elements of secondary structure (alpha helices and beta strands) are connected by turns or loops, and side chains in the protein backbone interact with each other or they are diversely bonded with other molecules, the tertiary structure is formed.

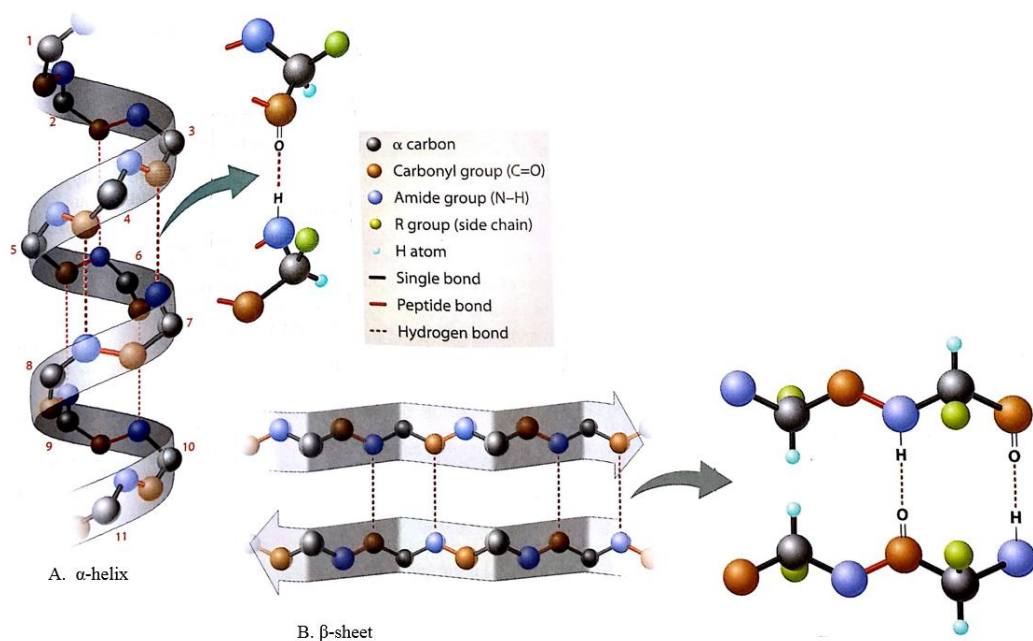


Figure 2. 5: Secondary structure elements: alpha helix (A), and antiparallel beta sheet (B) (Morris, Hartl, Knoll, & Lue, 2013, pp. 4-6).

Hydrogen and ionic bonding may be formed between polar side chains, while between nonpolar groups, hydrophobic interactions occur in the tertiary structure of the protein (Figure 2. 6). Three dimensional structure of one compact polypeptide chain, which is folded into one or more different structures can function biologically and this state is

called the native state. Each domain in the structure may have its own distinctive function. The global structure of a protein is so important that any conformational degeneration of the tertiary structure may cause loss of its function. This, so called denaturation state (loss of the natural shape) of a protein can be initiated by a temperature variation, changes in pH or exposure to some particular chemicals.

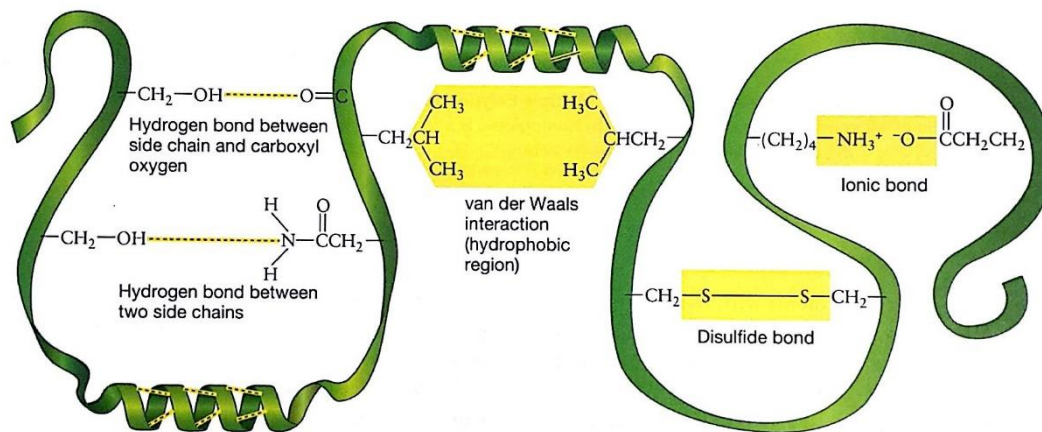


Figure 2. 6: An example demonstrating the interactions within a tertiary structure (Freeman, 2005, p. 59).

Oligomers contain more than one polypeptide chain and each chain has its own tertiary structure. These correctly folded protein clusters can function actively, and the overall form is called the quaternary structure. However, this clustering is not necessary for some proteins. Monomers, dimers, trimers, or polymers are constituted by one, two, three or more polypeptide chains, respectively. In quaternary structure, binding occurs with weak interactions such as hydrogen bonding or van der Waals interactions. The strength in the structure can be gained via the massive number of those interactions.

Protein folding is the process of a linear amino acid chain folding into a stable three dimensional shape. It is mainly stabilized by weak noncovalent bonds (Hydrogen bonding or van der Waals interactions). The last folded state of the protein has the minimum free energy.

2.2 Protein Function

Proteins are the fundamental structures for all living systems from bacteria to mammals and they perform multiple functions. The term *function* is related with the type of biochemical task that a protein performs. Proteins may be divided into six categories depending on their functionalities.

Enzymes as *catalytic proteins* have the function of accelerating biochemical reactions. *Transport proteins* transfer ions, micromolecules and macromolecules across plasma membrane in both ways. Acting as a tunnel, channel proteins transfer water, small molecules and ions such as sodium, potassium and calcium by facilitated diffusion. When molecules bind to carrier proteins, they are transported through the membrane via facilitated diffusion or active transport. *Storage proteins* store molecules. Casein storing amino acids and carbohydrates in milk of mammals, and ferritin storing iron in humans are two examples. *Signaling proteins* (receptors-hormones) can get and transmit chemical signals controlling cell growth and metabolism. *Defense proteins* protect the body by binding to viruses and bacteria and thus inhibiting their function. *Contractile proteins* like actin and myosin, structural proteins like keratin, collagen and silk, and cytoskeletal proteins like tubulin provide structural support.

Proteins are vital since they take part in almost every biochemical task needed for the living system to function. Thus, it is crucial to study protein structure in order to better understand the protein function.

2.3 Porin and Outer Membrane Protein G (OmpG)

2.3.1 Porin

A biological membrane is the boundary of life because it encloses and shields the components inside the cell against the unwanted. Bacteria have a special membrane structure and composition that are designed for survival in their specific conditions. They have to grow, find food and multiply, under adverse effects caused by the environment so they develop a distinguished membrane structure (Silhavy, Daniel, & Walker, 2010, pp. 1-2).

Bacteria have an outer and an inner membrane as their cell wall that are separated by an aqueous environment called periplasm (allow dynamic movement of molecules) (Figure 2. 7).

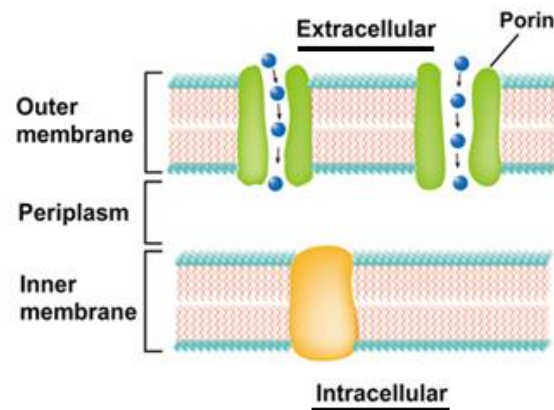


Figure 2. 7: Schematic representation of outer and inner membrane structures (double membrane) of a Gram-negative bacteria (Gram-positive ones do not own an outer membrane). While outside of the cell is called extracellular, inside of the cell is called intracellular. In between the outer and inner membrane, there is periplasm. In the membranes, there are proteins which help molecules enter or exit (PDB:2ODJ, 2009).

Integral (intrinsic) membrane proteins which constitute a relatively great mass in the interior of the membrane may form channels (Figure 2.8). These channels play role for several molecules, chemicals entering and exiting the periplasm. The integral membrane proteins that form channels are porins (transport proteins) found inside the outer-membrane of gram-negative bacteria.

Porins are composed of several connected β -strands in the shape of β -barrels. For example, OmpF is classified as a classical porin, which is a trimer each having 16-stranded β -barrel with pores (Subbarao & Berg, 2006, p. 751). The connections of β -strands are established either by short turns on periplasmic side and long loops on the

outer side. A trimer porin possess a stable structure because of hydrogen bonds in β -barrels, the organized loops being on each side of the membrane and each monomer having active interplay with each other (Haltia & Freire, 1995, p. 20).

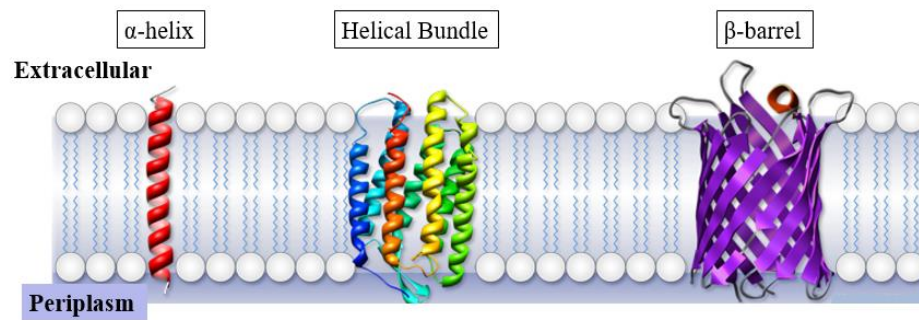


Figure 2. 8: Examples of integral membrane proteins shown in secondary structure forms. They cross the membrane. One α -helix form can function as a receptor protein or a recognition protein or a helical bundle form can be a receptor or a transport protein while a β -barrel form functions as a transporter named porin. The cell membrane contains phosphate groups (hydrophilic-white balls) and phospholipid bilayers (hydrophobic-two tails attached to white balls) (Background, 2009).

2.3.2 OmpG

Another outer-membrane protein is OmpG and it differs from other classical porins (i.e., OmpF) since it is not composed of sub-units. It is a stable monomer with a form of channel and it allows the passive diffusion of ions or molecules, especially sugars with the upper size limit of approximately 900 Da (Fajardo et al., 1998, p. 4452). It has the typical porin form of β -barrel structure consisting of 14 antiparallel β -strands (S1-S14) (Figure 2.9). On the extracellular side of the outer-membrane, seven long loops (L1-L7) connect β -strands. On the periplasmic side, six short turns which are made up of three or five residues (T1-T6) establish the connection between β -strands. Only one turn of an α -helix is on the loop L4.

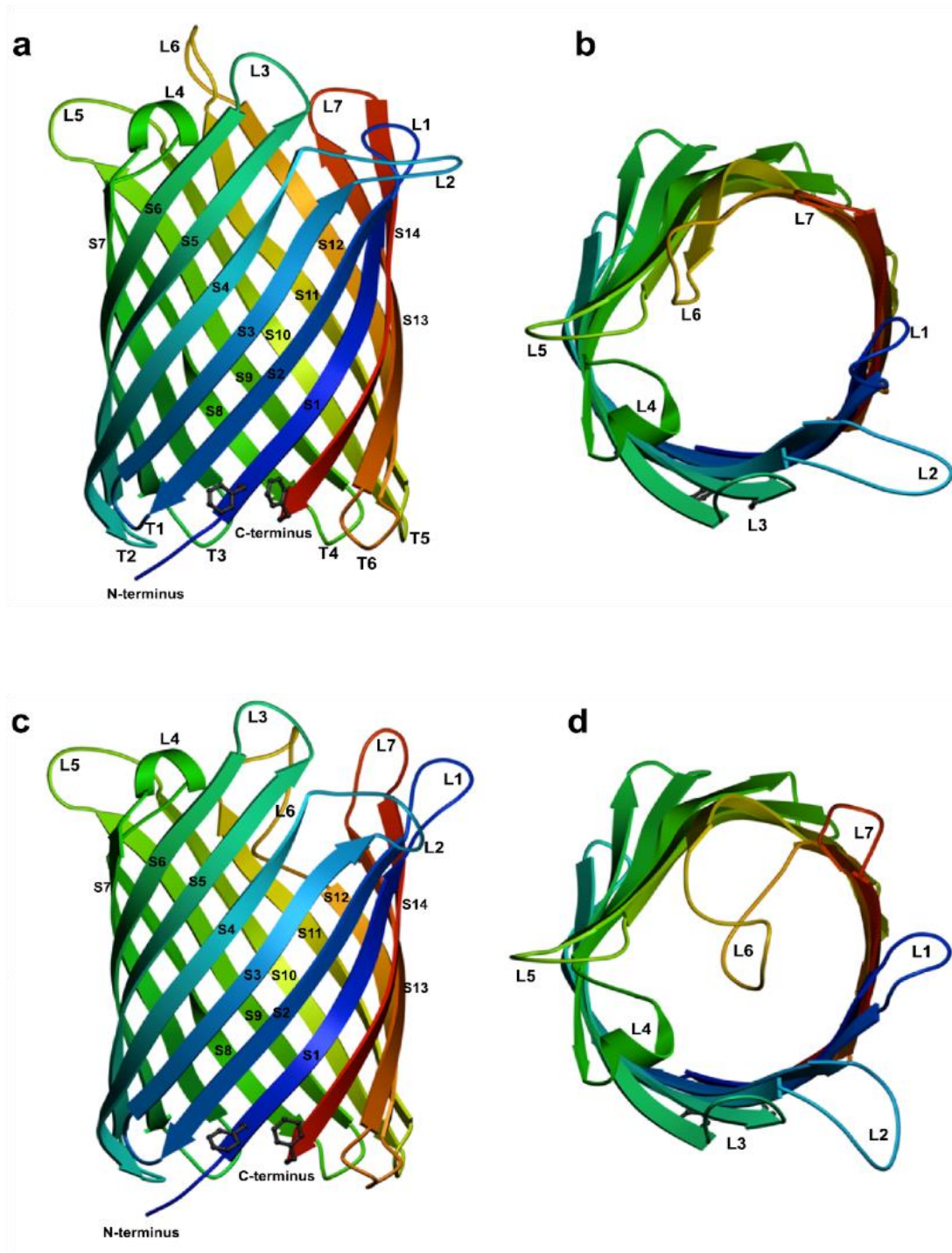


Figure 2. 9: Structure of OmpG in the case of open (a,b) and closed (c,d) conformation. OmpG is demonstrated like positioned inside the membrane (a,c) and viewed from the extracellular part (b,d) (Yildiz et al., 2006, p.3704).

The channel has a belt of aromatic and polar residues in order to stabilize its position within the membrane via interacting with lipid molecules (Yildiz, Vinothkumar, Goswami, & Kühlbrandt, 2006, p. 3703). For the most part of the belt, the polar residues are arginine and glutamine. The aromatic residues (also called greasy slide which is mainly parallel to the axis of the β -barrel) set a line which is assumed to aid the uptake of disaccharides through the channel (Yildiz, et al., 2006, p.3703). Unlike other members of the porin family, the conformation of the channel slightly changes depending on the surrounding buffer pH. While loop L6 extends itself through the extracellular side of the membrane at neutral pH, it folds into the pore channel and blocks the entrance at acidic pH (Figure 2.9). Below pH 6 (closed state of the pore channel), loop L6 (the leading role of the pH gating part of the channel) is led to a conformational change because of the breakage of hydrogen bonds in between S11 and S13 strands. This breakage is caused by the protonated histidines (His231 & His261) (Figure 2. 10) in acidic environment. Thus, loop L6 folds onto the channel at lower pH values (Korkmaz, Köster, Yıldız, & Mantele, 2012, pp. 305,401).

Above pH 6 (open state of the pore channel), the formation of hydrogen bonds between strands S11 and S13 caused by the deprotonated histidines (Figure 2. 10) cause loop L6 to extend through the extracellular medium instead of being bent onto the pore entrance. In the search of a reason for the blockage of loop L6 below pH 6, a sensible argument could be that L6 stops excessive number of protons to enter through the periplasm from an environment which is already acidic (Yildiz et al., 2006, p. 3703).

In general, LamB is the responsible integral membrane protein for the passage of sugars. In the case of deleted genome encoding for LamB, it cannot be expressed. In such cases OmpG replaces the duty of LamB. In other words, OmpG is not a regularly expressed porin in *E.coli*. Besides being a replacement for LamB, the exact function of the OmpG is not clearly defined yet.

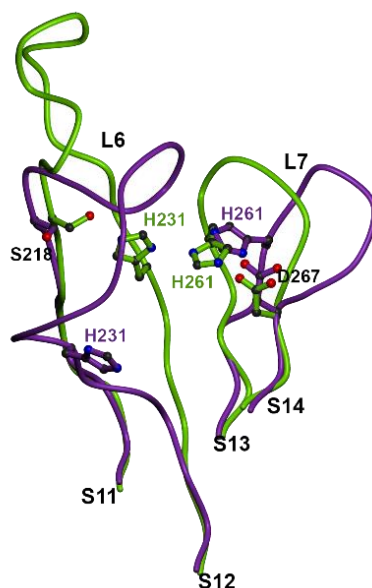


Figure 2. 10: Two strands and their connecting loops are shown in acidic pH (purple) and in neutral pH (green). In the closed state, His231 in S12 and His 261 in S13 repel each other since they are protonated at acidic pH. In the open state (pH above 6), these residues are deprotonated and attract each other, allowing newly constituted H-bonds (Yildiz et al., 2006, p. 3706).

2.4 Fourier Transform Infrared Spectroscopy (FTIR)

2.4.1 Interaction of Electromagnetic Radiation with Matter

Periodic oscillations of charges result in electromagnetic waves. Electric (**E**) and magnetic (**B**) field components of these waves are perpendicular to each other and to the propagation direction of the waves (Figure 2.11). Light is an electromagnetic radiation having wave-particle duality and travelling as photons. Photons can be expressed in frequencies (ν), energies (E), wavelengths or wavenumbers ($\tilde{\nu}$) by the following equations:

$$\nu = \frac{c}{\lambda}$$

$$E = h\nu \quad (2.1)$$

$$\bar{\nu} = \frac{1}{\lambda} = \frac{\nu}{c}$$

where c is the speed of light having value of $3 \times 10^8 \text{ m/s}$, h is Planck's constant having value of $6.63 \times 10^{-34} \text{ J.s}$, and lastly $\bar{\nu}$ is the wavenumber in 1/meters (sometimes can be converted to 1/centimeters).

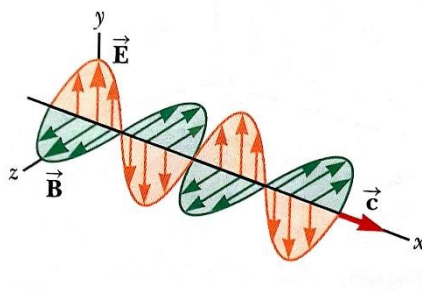


Figure 2.11: Electromagnetic wave propagating in the x-direction with speed of light (Serway & Jewett, 2010, p. 990).

The total energy of a molecule is presented as (Campbell, 1984, p. 15):

$$E_{total} = E_{translational} + E_{rotation} + E_{vibration} \\ + E_{electronic} + E_{electron \text{ spin orientation}} + E_{nuclear \text{ spin orientation}}$$

A molecule can translate, vibrate and rotate when in solution. Any energy related to translation, vibration and rotation is quantized. The quantized energy states have their own characteristics. When a photon strikes a molecule, the charges will begin to oscillate at the frequency of the radiation and exert forces leading to absorption. If the

oscillations occur at the resonant frequency as in Equation 2.1, the molecule will be excited with high probability and electrons in the lowest energy state (ground state) will move to the higher energy states (excited states). There are energy levels (electronic energy states) and sub-levels (vibrational and rotational energy states). These levels are demonstrated in energy level diagram (Figure 2.12).

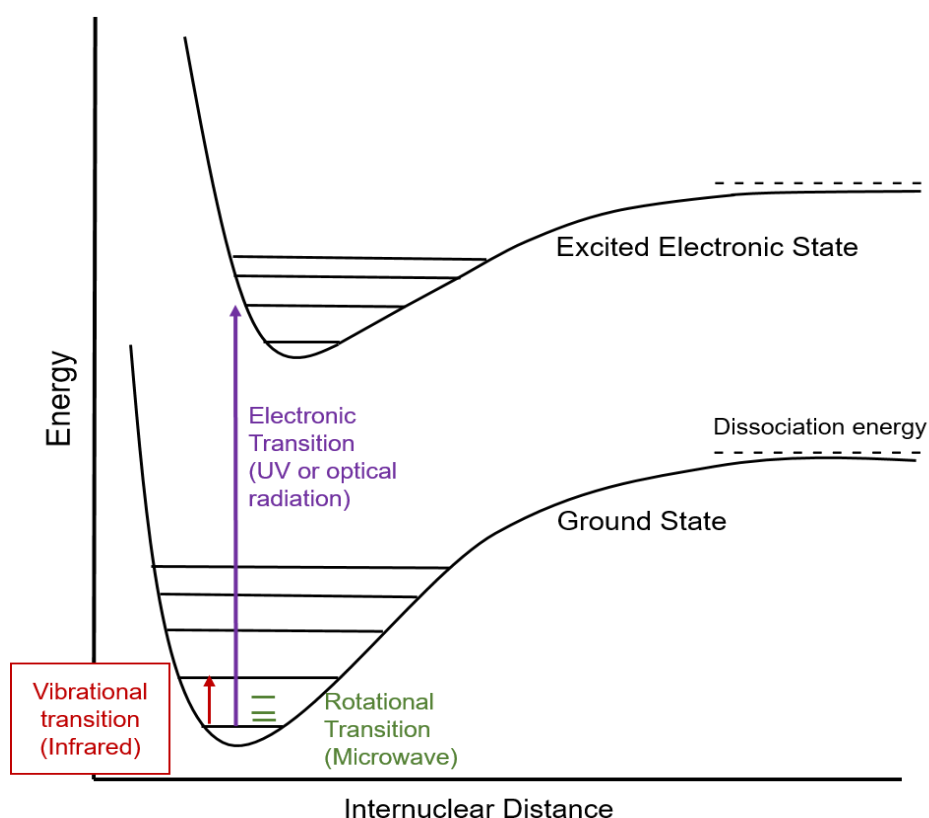


Figure 2. 12: Energy-level diagram demonstrating electronic energy levels and both vibrational and rotational sub-levels. A possible electronic transition is shown by purple arrow and a possible vibrational transition is shown by red arrow.

A spectrum plotted as absorption energy versus wavelength or frequency is called the absorption spectrum. The difference in possible transition energies is denoted by:

$$\Delta E = E_2 - E_1 = \frac{hc}{\lambda} = h\nu$$

where E_2 and E_1 are the first and ground energy levels, respectively. Moreover, for a possible transition, displacement of charge is necessary. While linear displacement having electric transition dipole moment interact with the electric component of the radiation, rotational displacement having magnetic transition dipole moment interact with the magnetic component of the radiation.

There are two circumstances regarding the absorption:

1. Absorption intensity depends on how many molecules experienced transitions. This gives quantitative information.
2. The broadening of an absorption line depends on various phenomena:
 - i. *Lifetime Broadening*: The line width of the spectrum is determined by the Heisenberg uncertainty principle which is shown as,

$$\Delta E \times \Delta t = \frac{h}{2\pi}$$

where ΔE is the uncertainty in energy and Δt is the uncertainty in lifetime (the time spent for the molecule being in an excited state). In quantum mechanics, when the system changes with time, the energy levels are not determined explicitly. For the lifetime of the molecule Δt , the energy is spread between $E - \Delta E$ and $E + \Delta E$. The $\Delta E = 0$ alternative is impossible since no molecule can reside in an excited state infinitely long. Broad spectral bands instead of sharp lines are resulted from the short-lived molecules in excited energy states.

- ii. *Doppler Broadening*: This is triggered by temperature. When molecules get hotter, they begin to move faster or have random motions. These motions cause frequency shifts of the molecules relative to an observer. Hence, the light emitted is Doppler shifted resulting in

broader lines. Doppler width of a line at half height can be calculated by:

$$\Delta\nu_D = \frac{\nu_0}{c} \sqrt{\frac{2kT}{m}}$$

where ν_0 is the transition frequency, c is the speed of light, k is the Boltzmann constant ($1.38 \times 10^{-23} \text{ JK}^{-1}$), T is the temperature (K) and m is the mass of the molecule. For molecules having very low mass (i.e. Hydrogen), the Doppler width is high.

- iii. *Pressure Broadening*: When molecules collide with each other, the lifetime spent for the molecules on the upper state decreases so that the uncertainty in energy increases. This effect is triggered by both temperature and density of the molecules.
- iv. *Proximity Broadening*: Since molecules get closer, they are affected by each other which cause changes in line width or position. The lines may become so close that they are overlapped and cannot be distinguished exactly.

2.4.2 Infrared (IR) Spectroscopy

The working principle of IR spectroscopy is based on the interaction between the electromagnetic beam and the electric dipole moment of the molecule. Since there has to be a linear displacement of charge (the dipole moment of the molecule must change), IR-active molecules are the ones that vibrate with changing dipole moment. For IR-inactive ones, there will not be a change in dipole moment such as N_2 molecule (no polarity) and hence, will not be seen in the IR spectrum.

IR radiation originates from the energy of the vibrational motions of the molecules. The potential energy curve of a diatomic molecule resembles a parabola that is shown in Figure 2. 12. The low energy part of the parabola indicates the balance of bond lengths of the diatomic molecule. This parabola gives the same potential energy shape of a spring in between the masses of two. Hence, the bond vibrations can be regarded as the simple harmonic motion that the molecule performs.

The vibrational frequency of a diatomic molecule is:

$$\nu_{vib} = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}}$$

where k denotes force constant and μ denotes the reduced mass of the diatomic molecule having masses m_1 and m_2 .

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}$$

A simple vibrational frequency ratio calculation can be examined for different reduced masses. Consider two diatomic molecules O-H and O-²H (or O-D). The reduced mass of O-H molecule is calculated as:

$$\frac{1}{\mu_{O-H}} = \frac{1}{16} + \frac{1}{1} = \frac{17}{16}$$

and thus,

$$\mu_{O-H} = \frac{16}{17}$$

The reduced mass of O-D molecule is calculated as:

$$\frac{1}{\mu_{O-D}} = \frac{1}{16} + \frac{1}{2} = \frac{9}{16}$$

and thus,

$$\mu_{O-D} = \frac{16}{9}$$

The ratio of frequency vibration is then as

$$\frac{\nu_{vib}(O-H)}{\nu_{vib}(O-D)} = \sqrt{\frac{16/17}{16/9}} \sim \frac{1}{\sqrt{2}}$$

where force constants are assumed to be the same and thus, the vibrational frequency differs by $\sqrt{2}$.

The quantized energy of vibrations is denoted as

$$E_{vib} = \left(n + \frac{1}{2}\right) h\nu_{vib} \quad n = 0, 1, 2, \dots$$

where the vibrational quantum number, n takes integer values. The transitions between vibrational energy states must satisfy the selection rule ($\Delta\nu = \pm 1$).

The region of infrared radiation is in between 0.7 μm and 500 μm but the IR region is divided into three sub-regions due to application purposes. These are near (NIR) infrared region ranging from 0.8 to 2.5 μm ; middle (mid-IR) infrared region ranging from 2.5 to 25 μm and far (far-IR) infrared region ranging from 25 to 100 μm . The mid-IR region corresponds to 4000 to 400 cm^{-1} in terms of wavenumbers, which involves mainly the transitions from ground vibrational energy level to the first excited vibrational energy level.

The information gained in the mid-IR region includes changes in chemical structure, redox states, bond strength, distortions and angles, hydrogen bonding, the electric field occurred in the neighborhood of the molecule, and the conformation of the molecule (Barth, 2007, pp. 1074-1075).

An IR spectrum is a plot of absorption versus wavenumber ($\bar{\nu}$). Several normal modes belonging to the same molecular group (such as CO_2 having four different modes) and all the vibration modes of IR-active groups form the absorption bands. These bands are composed of stretching modes of diatomic molecules, and stretching and bending modes of polyatomic molecules. Normal modes can be categorized as symmetric and asymmetric stretch, wagging, rocking, twisting or scissoring. Total degree of freedom of a compound is $3N$ where N is the number of the molecules constituting the compound. While 3 axes are accounted for translational degrees of freedom, 3 other axes are described as rotational degrees of freedom. For linear molecules, there will exist $3N-5$ vibrational normal modes. For nonlinear molecules, $3N-6$ vibrational modes are constituted. For a triatomic molecule like H_2O , there are three normal coordinates, so three modes of vibration are demonstrated in Figure 2.13.

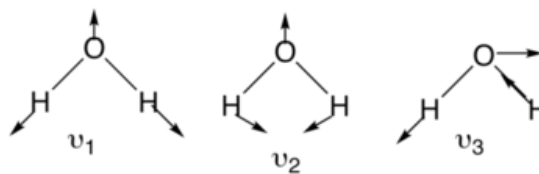


Figure 2. 13: Three modes of vibration of H₂O (in gas phase) where ν_1 (3650 cm⁻¹) is the symmetric stretch, ν_2 (1600 cm⁻¹) is the bending mode and ν_3 (3750 cm⁻¹) is the asymmetric stretch (Kowolik, 2016).

2.4.2.1 The Spectrometer

The FTIR spectroscopy is based on the principle of Michelson interferometer. A heated Silicon Carbide rod or wire (water cooled source, globar) is used as IR radiation source. ETC (Electronic Temperature Control) EverGlo source is an example of globar source, which is also the one used in this study. The beam splitter splits the beam emerging from the source into two halves with the help of a half-reflecting mirror. The beam splitter is made of KBr (Potassium Bromide) and Ge (Germanium) is coated between two KBr materials in order to have half transparency. While one half is reflected back from a fixed mirror, the other half is reflected back from a moving mirror. The reflected beams then recombine either constructively or destructively and reach the detector. This radiation goes through the sample, sandwiched between two CaF₂ (calcium fluoride) windows, which have a low refractive index ($n=1.4339$) and thus, have transparency in the mid-IR region (Fabian & Mantele, 2002, p. 4). During the sample-radiation encounter, some of the beam is absorbed by the sample, while the rest, transmitted portion, travels to the detector. The amount of light absorbed by a given sample is governed by the Beer-Lambert Law equation:

$$A = -\log \frac{I}{I_0}$$

where I corresponds to the intensity of the transmitted light and I_0 is the intensity of the incident light. The radiation reached the detector is converted into electric signals which are amplified, processed, digitalized and finally Fourier transformed. One type

of detector is deuterated triglycine sulfate (DTGS) in which the beam arrives and causes a change in temperature of the dielectric material. Mercury cadmium telluride (MCT) is another type of detector which is the choice in cases of high speed measurements and low IR throughput experiments (Kempfert, Jiang, Oas, & Coffin, 2001).

The components of the Michelson interferometer can be viewed in Figure 2. 14. The fixed and moving mirrors are standing perpendicular to each other. In the spectrometer, the moving mirror moves at a constant velocity and speed values can be adjusted from 0.16 to 0.64 cm/s. A faster speed for the mirror allows fast measurements but will cause a lower signal-to-noise ratio than moving at a slower speed in between the limits. When the splitted beams travel the same distances (this case is called zero path difference-ZPD), the interference of the beams is constructive. If the moving mirror positions at a quarter value of the original wavelength, the interference will become destructive (Figure 2. 15). Of course, these exact recombination of the beams are hard to obtain due to instable temperature so in order to overcome this problem, a reference interference motif is collected by a He-Ne laser.

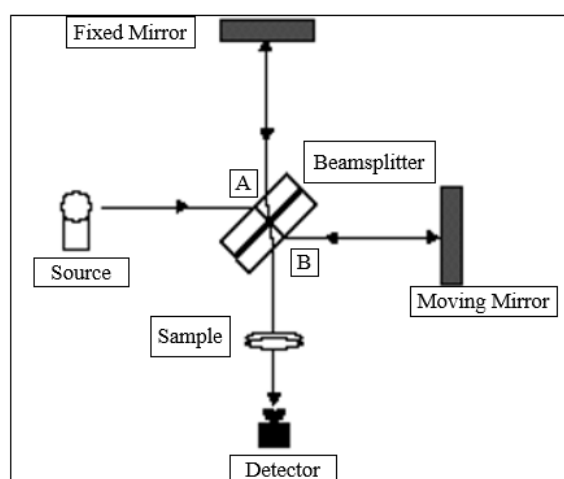


Figure 2. 14: Michelson Interferometer (Experiment 7: IR Spectroscopy, 2007).

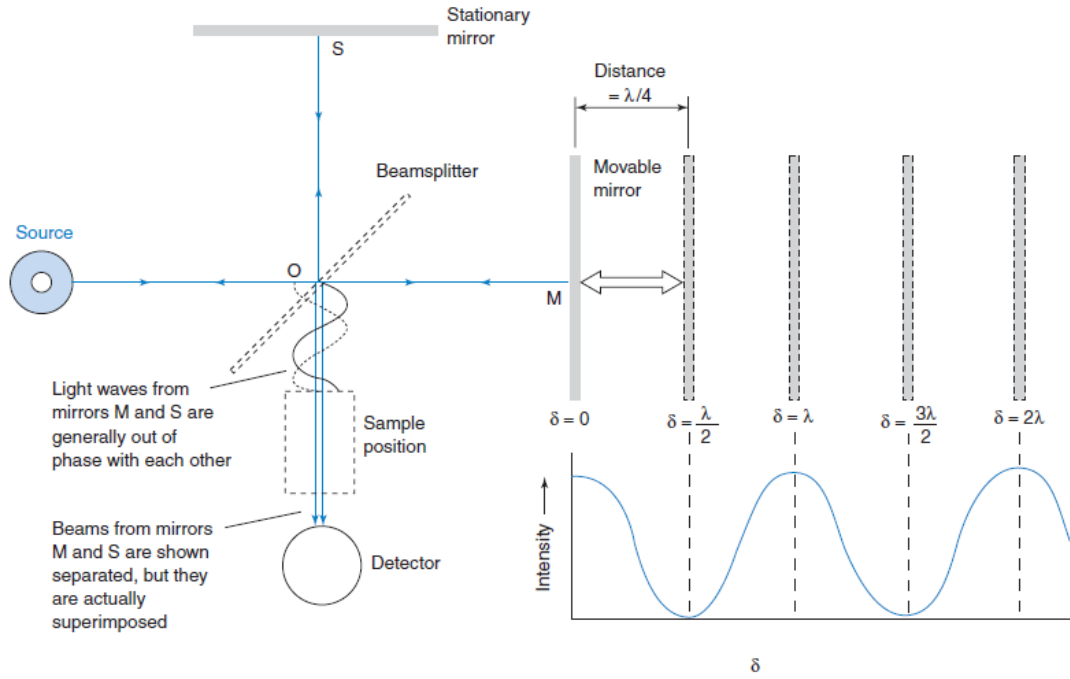


Figure 2. 15: A diagram of monochromatic light passing through the Michelson interferometer and reaching the detector. Constructive interference occurs when there is no path difference ($\delta=0, \lambda, 2\lambda, \dots$) and their corresponding moving mirror distances are $0, \lambda/2, \lambda, \dots$. Destructive interference occurs when there are path differences ($\delta=\lambda/2, 3\lambda/2, \dots$) and their corresponding moving mirror distances are $\lambda/4, 3\lambda/4$, and more (Harris, 2007, p. 443).

The recombined beams that reach the detector may be expressed as a superposition of cosine waves (an interferogram) since the moving mirror travels back and forth. The desired resolution also depends on the interferograms which is measured from ZPD to a maximum optical path difference (OPD, also called retardation). If the desired resolution is 2 cm^{-1} , the maximum OPD must be 0.5 cm and the relation between them can be shown by an equation:

$$\delta\tilde{\nu} = \frac{1}{OPD}$$

where $\delta\tilde{\nu}$ is the wavenumber resolution in cm^{-1} and OPD in cm. The interferogram is a plot of the signal intensity versus OPD and the position of the maximum intensity signal gained is the ZPD (zero retardation). In order to understand what an interferogram looks like, a comparative figure can be seen in Figure 2. 16.

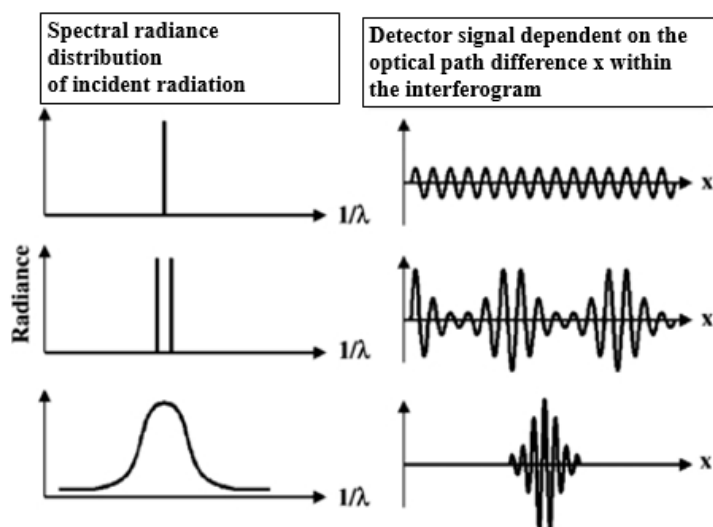


Figure 2. 16: Spectral radiance distribution of incident radiation at three different superposition cases and their corresponding interference patterns (interferograms). The top one is resulted by monochromatic radiation. The middle one corresponds to two different wavelengths having a small wavelength difference. The last one is containing various wavelengths so that a broad continuous spectrum occurs. For example, in the last one, the ZPD is the position of the maximum intensity within the interferogram which is constructed by all the wavelengths interfered constructively and at other retardations, the intensities decrease gradually (Möllman & Vollmer, 2013, p. 126).

Even if the interferogram is plotted as a function of OPD, the signals are actually functions of time. When IR radiation encounters the sample after passing through the mirrors, only IR-active frequencies will be absorbed and the rest will pass the sample and reach the detector. Then, the detector converts the signals to electric signals. Afterwards, the signals are amplified, passed through an electric filter (to filter undesired signals and noise) and lastly, by the help of analog-to-digital (AD) converter they are digitized. Many of the components mentioned above are shown in an FTIR spectrometer layout in Figure 2. 17.

The information of the interferogram is contained in the time domain. If Fourier transform is applied to this interferogram, a mathematical function in the frequency domain and thus, the spectrum is obtained.

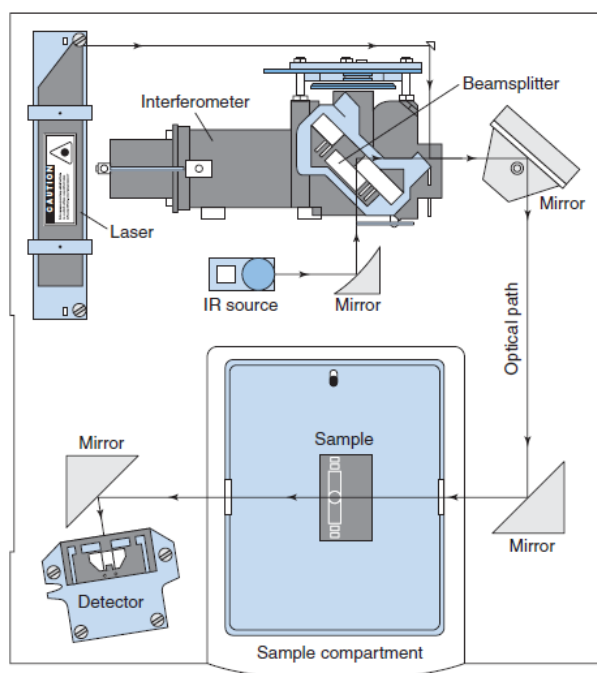


Figure 2. 17: Basic components inside a commercial FTIR spectrometer (Harris, 2007, p. 446).

2.4.2.2 Fourier Transform

Signals are information transmitters which carry information on a wave and they can be written as functions of independent variable(s). The signals mainly vary as continuous-time signals and discrete-time signals. Continuous-time signals occur if the independent variable is continuous. It also commonly can be indicated as analog signal. Discrete-time signal amplitudes get discrete values if the independent variable is discrete. Discrete-time signal is related with digital signal if the signal values are precisely quantized.

The interferogram obtained carries signal intensity information over OPD. With this information, a fringe pattern or phase-shift analysis can be deduced. However, would it be enough to reveal or improve the information that a signal carries? How about if there exists too much noise, a low signal-to-noise ratio, a shape or a function that resembles nothing familiar or too complicated? In the Figure 2. 16, the interferogram

graph at the top has a sinusoidal form, which is very familiar to us, but the graph at the bottom is a combination of many waves. Those waves should be deconstructed so that they can be expressed in their sinusoidal forms with carrying frequency information. The way to gain this is the Fourier transformation. It will also provide valuable information after a signal encounters amplifiers and passes through filters or when more mathematical applications are wished to be done such as derivative calculation. Signals can be expressed as a periodic or non-periodic function of time. A periodic function with a period T has the form of

$$f(t + T) = f(t)$$

Fourier series are composed of periodic functions. In the case of waves, the sum of periodic functions is considered as the sum of sinusoidal functions. Let us write a function $x(t)$ in terms of even and odd functions as a sum. The even function is $\cos(mt)$ and the Fourier cosine series is expressed as

$$x(t) = \frac{1}{\pi} \sum_{m=0}^{\infty} F_m \cos(mt)$$

where F_m represents the coefficients ($m=0,1,2,\dots$). The boundaries of the integration must be in the range of $(-\pi,\pi)$ and the Kronecker delta function is shown as

$$\delta_{mn} = \begin{cases} 1, & \text{if } m = n \\ 0, & \text{if } m \neq n \end{cases}$$

The coefficient F_m is then determined by several steps below:

$$\int_{-\pi}^{\pi} x(t) \cos(m't) dt = \frac{1}{\pi} \sum_{m=0}^{\infty} \int_{-\pi}^{\pi} F_m \cos(mt) \cos(m't) dt$$

while

$$\int_{-\pi}^{\pi} \cos(mt) \cos(m't) dt = \begin{cases} 0, & m \neq m' \\ 1, & m = m' \end{cases} \equiv \pi \delta_{mm'}$$

$$\int_{-\pi}^{\pi} x(t) \cos(m't) dt = \frac{1}{\pi} \sum_{m=0}^{\infty} F_m \pi \delta_{mm'}$$

Only $m = m'$ term survives and thus, putting aside the ' mark:

$$F_m = \int_{-\pi}^{\pi} x(t) \cos(mt) dt$$

is the coefficient for $x(t)$.

The odd function is $\sin(mt)$ and Fourier sine series can be written as:

$$x(t) = \frac{1}{\pi} \sum_{m=0}^{\infty} F'_m \sin(mt)$$

and one can apply the several steps above and find the coefficient as

$$F'_m = \int_{-\pi}^{\pi} x(t) \sin(mt) dt$$

Lastly, Fourier series is in the form of

$$x(t) = \frac{1}{\pi} \sum_{m=0}^{\infty} F_m \cos(mt) + \frac{1}{\pi} \sum_{m=0}^{\infty} F'_m \sin(mt)$$

where F_m and F'_m are the relative weights.

The determined values of interferogram $x(t)$ are Fourier transformed in order to successfully obtain the spectrum. Fourier transform is used for non-periodic signals.

Another function $X(m)$ can be written exponentially.

$$X(m) \equiv F_m - iF'_m = \int_{-\pi}^{\pi} x(t) \cos(mt) dt - i \int_{-\pi}^{\pi} x(t) \sin(mt) dt$$

m is replaced with angular frequency ω and the integral boundaries are from $(-\infty, +\infty)$:

$$X(\omega) = \int_{-\infty}^{+\infty} x(t) \exp(-i\omega t) dt$$

is the Fourier transformed $x(t)$ and the spectrum of $x(t)$.

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega) \exp(i\omega t) d\omega$$

is the inverse Fourier transform.

Euler's formula used in the analysis of complex numbers indicates

$$e^{i\omega t} = \cos(\omega t) + i \sin(\omega t)$$

and

$$e^{-i\omega t} = \cos(\omega t) - i \sin(\omega t)$$

equations. In the Fourier transform and inverse Fourier transform equations can be expressed using Euler's formula. In this form, the $\sin(\omega t)$ part of the equation will disappear when processed under integral since it is an odd and imaginary function.

After Fourier transformed, the signals shown on the interferogram are displayed as spectra (absorption vs. wavenumber), which carry information on the frequency domain. Three examples of Fourier transformed interferograms are shown in Figure 2. 18.

Moreover, in the case of examining a sample, its interferogram and FTIR spectrum after being Fourier transformed can be seen in Figure 2. 19.

All in all, the intensity variation of an FTIR spectrum in the frequency domain which is converted from the time domain by Fourier transform can simply be expressed as:

$$A(\bar{\nu}) = \int_0^{\infty} I(x) \cos(2\pi\bar{\nu} x) dx$$

where A is the intensity variation, I (x) is the source intensity and x is the path difference of beams returning from the mirrors. The overall equation is called one-half Fourier cosine transform.

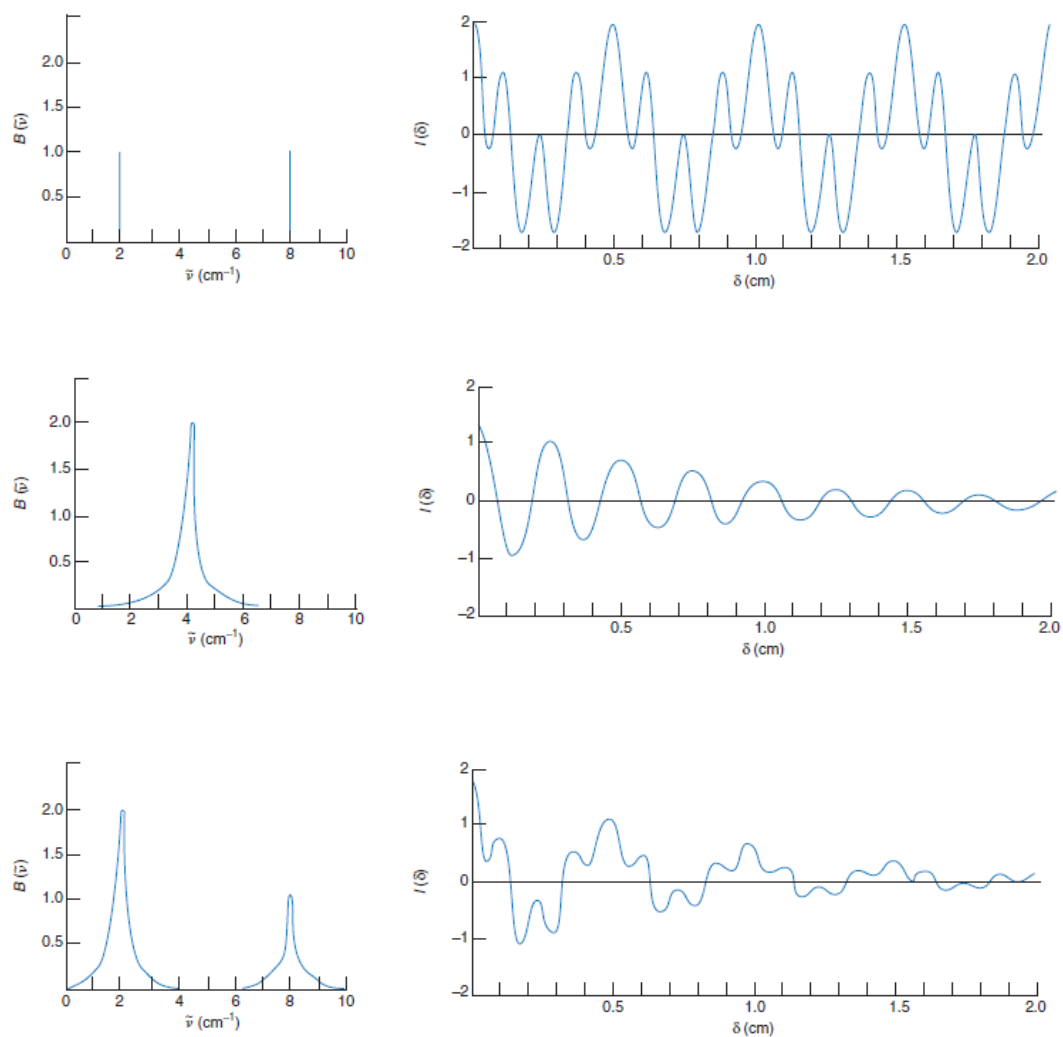


Figure 2.18: The graphs on the left column are the Fourier transformed spectra of the interferograms on the right column. At the top, the interferogram and its corresponding spectrum contain information of two monochromatic waves (waves with wavenumbers 2 and 8 cm^{-1}). Graphs in the middle and at the bottom are composition of many wavelengths. The spectrum in the middle demonstrates that at wavenumber 4 cm^{-1} , the absorption band is centered. At the bottom, there are two different absorption bands where one is centered at 2 cm^{-1} and the other is centered at 8 cm^{-1} (Harris, 2007, p. 444).

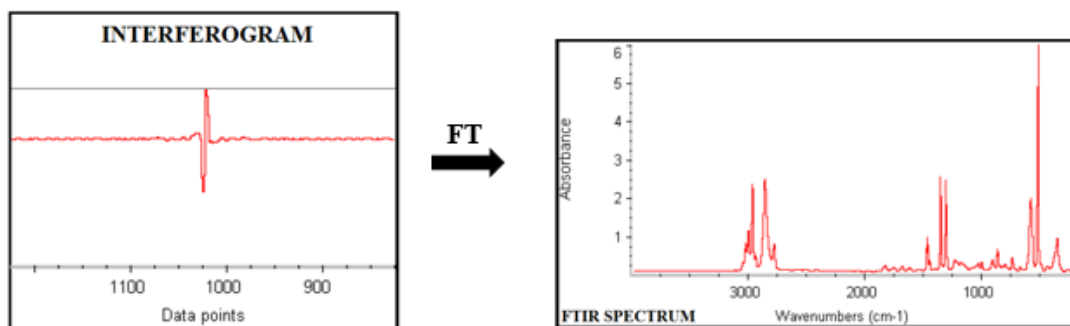


Figure 2. 19: An interferogram of a sample over data points and its FTIR spectrum (absorbance vs. wavenumbers) (Nicolet FTIR User's Guide, 2004).

2.4.3 FTIR Vibrational Frequencies of a Protein

A meaningful but a less informative deduction can be made at first glance on an FTIR spectrum of a protein because of the bands arising at different wavenumbers. For example in the Figure 2. 19, FTIR spectrum has characteristic bands at around 3000, 1300 and 600 cm^{-1} . Characteristic bands arisen in an FTIR spectrum correspond to characteristic vibrational frequencies (or wavenumbers) and some of these will correspond to different secondary structure components inside the examined protein. The characteristic bands of a protein spectrum are named as amide bands and they all correspond to a molecular group or groups. These amide bands correspond to 9 amide modes (vibrational modes) called from amide I, amide II to amide VII, and amide A and B. However, amide I, amide II and amide A bands are the ones used in the analysis of protein structure determination.

Amide I band seen in 1600-1700 cm^{-1} region, consists mostly of C=O stretching vibration (80%) with lesser contribution from out-of-phase C-N stretching vibration and N-H in-plane bending (Goormaghtigh, Cabiaux, & Ruyschaert, 1994, p.332; Barth, 2007, p.1080; Jackson & Mantsch, 1995, p. 96). The amide I band gives information about the secondary structure profile of the protein backbone.

The amide II absorption band seen in 1500-1600 cm^{-1} region, is formed by out-of-phase N-H in-plane bending (40-60%), C-N stretching vibrations (18-40%) and C-C stretching modes (nearly 10%). The amide A band seen in 3270-3310 cm^{-1} region, is mainly made up of N-H stretching vibration. These three bands are shown in Figure 2. 20.

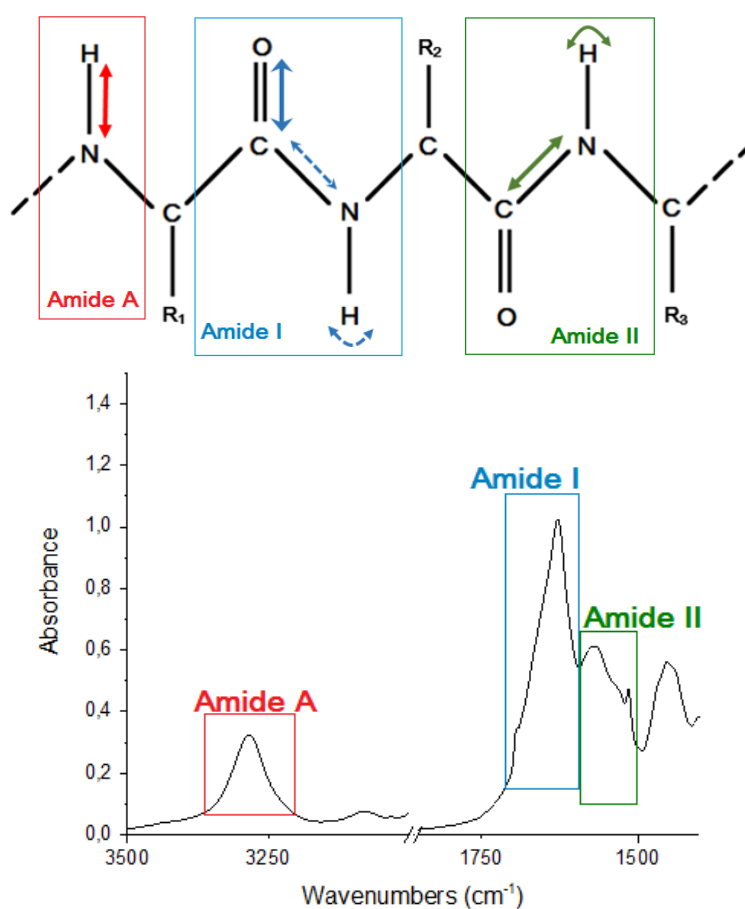


Figure 2. 20: A representative spectrum showing amide bands and their corresponding molecular groups shown in boxes on the protein backbone in accordance with color coding.

The α -helix structures in the amide I region is located around 1655 cm^{-1} . However, if β -sheet structures are dominant inside a protein, a strong band close to 1630 cm^{-1} and a weak band close to 1685 cm^{-1} will appear and these will correspond to antiparallel β -sheet structures. While turns are positioned at around 1670 cm^{-1} , unordered structures are located close to 1660 cm^{-1} . Moreover, the H-O-H bending mode of water molecules overlap with the amide I band at around 1640 cm^{-1} , which makes the study of proteins in solution a challenge. At first sight, the assigned bands give few information about the structure by just determining the wavenumbers and secondary structures corresponding to those wavenumbers. However, this approach is not supported because there might be molecular groups having close characteristic wavenumbers in the spectrum and their bands may overlap. This is why, more informative approaches (mathematical calculations) must be performed such as subtracting bands, taking derivatives, and applying curve-fit methods to overlapping bands.

CHAPTER 3

MATERIALS AND METHODS

3.1 Materials

3.1.1 Chemicals

Table 3. 1: Chemicals used in the study.

Chemicals	Company Name
Na ₂ HPO ₄ (di-sodium hydrogen phosphate) Molecular weight (MW) = 141.96 g/mol	Merck
NaH ₂ PO ₄ (sodium phosphate monobasic) MW = 119.98 g/mol	Sigma-Aldrich
D ₂ O (deuterium oxide) MW = 20.03 g/mol	Merck
C ₁₄ H ₂₈ O ₆ (n-octyl β-D-glucopyranoside)	Glycon
PTFE-Paste	Carl Roth

3.1.2 Buffers

In buffer preparation, according to the desired pH value, percentages of Na₂HPO₄ and NaH₂PO₄ were calculated by Henderson-Hasselbalch equation. The calculation is as followed:

$\text{Na}_2\text{HPO}_4 \rightarrow \text{A}^-$ (proton acceptor; conjugate base of HA)

$\text{NaH}_2\text{PO}_4 \rightarrow \text{HA}$ (proton donor; weak acid)

and the Henderson-Hasselbalch equation is

$$\text{pH} = \text{pK}_a + \log\left(\frac{[\text{A}^-]}{[\text{HA}]}\right) \quad (3.1)$$

where pK_a is 6.86 for phosphate and the summation of the concentrations of proton acceptor and donor is always equal to 1.

$$[\text{A}^-] + [\text{HA}] = 1 \quad (3.2)$$

Solving equations 3.1 and 3.2 together gives the concentrations of $[\text{HA}]$ and $[\text{A}^-]$, and their percentages. Later, in order to prepare a sodium phosphate (NaP_i) buffer, the weights of Na_2HPO_4 and NaH_2PO_4 in grams must be calculated. Molarity equation is:

$$\text{Molarity} = \frac{\text{moles of solute (mol)}}{\text{volume (L)}}$$

$$\text{where moles of solute} = \frac{\text{mass (g)}}{\text{molar mass (g/mol)}}$$

Therefore, the mass of one substance is calculated as:

$$\text{Mass} = \text{Molarity} \times \text{Volume} \times \text{Molar mass} \times \text{Percentage of that mass}$$

For instance, let us calculate the mass of NaH_2PO_4 for pH 7.5 with 0.1 M concentration in 5 ml volume:

$$\text{Mass} = 0.1 \text{ M} \times 5 \times 10^{-3} \text{ L} \times 119.98 \frac{\text{g}}{\text{mol}} \times 0.1864 = 0.0575 \text{ g}$$

In this study, 0.1 M NaP_i pH 5.5, pH 7.5, p^2H 5.5 and p^2H 7.5 buffers were prepared and used as environment for the proteins. The masses of Na_2HPO_4 and NaH_2PO_4 are present in Table 3. 2 for both pH and p^2H values.

In order to gain the desired pH or p^2H value, for 0.1 M NaP_i buffer, the calculated masses (Table 3. 2) of both Na_2HPO_4 and NaH_2PO_4 were mixed. When calculating the masses of Na_2HPO_4 and NaH_2PO_4 for p^2H values, all of the calculations were done according to $\text{p}^2\text{H} = \text{pH} + 0.4$ formula (pH 5.1 was calculated for obtaining a buffer of p^2H 5.5). Buffers also contained %0.5 n-octyl β -D-glucoside (OG) weighed for 2 and

5 ml volumes. In general, OG is a nonionic detergent which mimics the lipid bilayer surrounding, solubilize membrane proteins without damaging their structures, and can be easily removed (Seddon, Curnow, & Booth, 2004, p. 108).

Table 3. 2: Calculated masses of Na_2HPO_4 and NaH_2PO_4 in double distilled water and deuterated water for two different pH and p^2H values.

pH	Mass of Na_2HPO_4 (g)	Mass of NaH_2PO_4 (g)
5.5	0.0030	0.0575
7.5	0.0577	0.0112
p^2H		
5.5	0.0005	0.0236
7.5	0.0180	0.0088

3.1.3 FTIR Spectra of H_2O and $^2\text{H}_2\text{O}$

As mentioned in section 2.4.2, water molecules have three vibrational modes: O-H (O-D) symmetric stretching (ν_1), O-H (O-D) asymmetric stretching (ν_3) and H-O-H (D-O-D) bending (ν_2) modes. According to the IR spectrum of H_2O (Figure 3. 1), a strong broad band at $\sim 3400 \text{ cm}^{-1}$ corresponds to (ν_1) mode, overtone of bending ($2\nu_2$) and (ν_3). Another band at $\sim 2132 \text{ cm}^{-1}$ is a combination of (ν_2) mode and libration (a collection of normal modes). The absorbance of the third band's peak belonging to (ν_2) mode is placed at 1645 cm^{-1} . This H-O-H bending mode obscures the analysis of amide I band when monitoring a protein's FTIR spectrum in H_2O buffer. To avoid the obscurity of amide I band, an alternative approach is to replace H_2O buffer with $^2\text{H}_2\text{O}$ buffer. As the figure shows (Figure 3. 1), when the O-H molecules are replaced with O- ^2H molecules, the IR spectrum is shifted to lower wavenumbers due to the heavier mass of ^2H . This downshift leads the amide I part of the spectrum to become free of water absorbance. In the IR spectrum of $^2\text{H}_2\text{O}$, the weak band at 3406 cm^{-1} is resulted from the O-H stretching vibrations of the existing H-O- ^2H molecules. The peak at 2508

cm^{-1} is a complex band containing an overtone of bending ($2\nu_2$), (ν_1) and (ν_3). A combination of (ν_2) and libration modes give rise to the band at 1463 cm^{-1} . The H-O-H bending mode at 1645 cm^{-1} in H_2O shifts down to 1210 cm^{-1} , which corresponds to (ν_2) mode in $^2\text{H}_2\text{O}$ (Venjaminov & Prendergast, 1997, pp. 236-237).

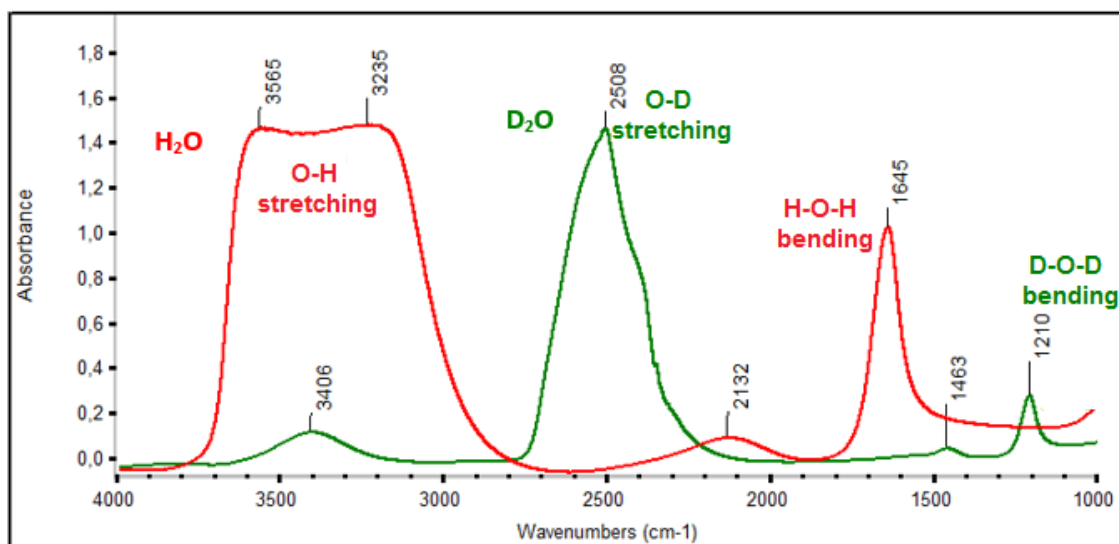


Figure 3. 1: The FTIR absorbance spectra of H_2O (red) and $^2\text{H}_2\text{O}$ (green).

3.1.4 OmpG-WT and Mutants

OmpG is made up of a polypeptide containing 301 amino acids. Wild type OmpG (OmpG-WT) used in this study lacks its signal sequence and contains 281 amino acids. A mutant, OmpG-16S is designed to have a form of 16-stranded β -barrel and it contains 319 amino acids. The amino acid sequence in S13 and S14 strands is repeated with the intention to form S15 and S16 strands. In order to achieve this, 38 amino acids are added to the original amino acid chain. Another mutant, OmpG-16SL is also created in order to have a 16-stranded β -barrel form. Seven lysines in the OmpG-16S structure are exchanged with arginines in OmpG-16SL, in order to get channel

electrostatic properties similar to OmpG-WT. Loop L6 chain is also increased by 6 amino acids so the final structure consists of 325 amino acids in total.

OmpG-WT and mutants were received as stock solutions from Dr. Özkan Yıldız and researcher Katharina van Pee at the Department of Structural Biology, Max Planck Institute of Biophysics, Frankfurt am Main, Germany. The properties of these stock solutions can be viewed in Table 3. 3. Later, in 25 µl aliquots, they were exchanged to 0.1 M NaPi buffers with pH 5.5-7.5 and p²H 5.5-7.5 by centrifugation.

Table 3. 3: Stock solution properties of OmpG-WT and mutants.

OmpG-WT	Concentration of 30 mg/ml in 1% OG, 25 mM tris buffer with pH 8
OmpG-16S	Concentration of 10 mg/ml in 0.5 % OG, 25 mM tris-HCl buffer with pH 8
OmpG-16SL	Concentration of 12-18 mg/ml in 0.5 % OG, 25 mM tris buffer pH 8

3.2 Methods

3.2.1 FTIR Spectroscopy

With the use of CaF₂ windows, IR spectra of proteins were collected in the FTIR transmission mode. Proteins in liquid form (both H₂O and ²H₂O) were placed between the two CaF₂ windows. One of the windows has a cavity specifying the optical path length as 10 µm. The choice on CaF₂ windows being used in the transmission mode is feasible because in the mid-IR part of the spectrum they are transparent. For protein solutions in H₂O (both 0.1 M NaPi pH 5.5 and pH 7.5), 5 µl of the solution was pipetted onto the window which has the cavity and was dried under a gentle stream of N₂ gas. Then, the same procedure was applied with the added 5 µl solution. Later, the windows were closed onto each other and transmission spectra were collected.

For protein solutions in $^2\text{H}_2\text{O}$ (both 0.1 M NaPi p ^2H 5.5 and p ^2H 7.5), 10 μl of the solution was placed onto the window and dried for 5 min. Then, 5 μl of $^2\text{H}_2\text{O}$ was added to the sample and the windows were closed for data collection. The collection of transmission spectra was done by Thermo Nicolet 6700 FTIR spectrometer outfitted with DTGS detector.

Windows were sealed with a water impermeable paste in temperature ramp (t-ramp) experiments so that no water was lost during heating. 128 scans were performed with 2 cm^{-1} resolution. For each mutant protein, in H_2O and $^2\text{H}_2\text{O}$, in acidic and neutral environment, 45 t-ramp spectra were recorded during heating from 22 $^{\circ}\text{C}$ to 106 $^{\circ}\text{C}$ with increments of 2 $^{\circ}\text{C}$. One such experiment lasted approximately 20 hours. A fresh background spectrum was collected before every step of temperature rise using a sample shuttle. For every sample, this procedure was applied until an exact phase transition temperature value was assessed. The heating was controlled by a microprocessor, which circulated a current wired around the sample holder.

3.2.2 Data Process of the Spectra in order to Analyze Protein Structure

The infrared spectra of proteins are mixtures of band compositions of multiple structural elements since proteins are intrinsically complex molecules. The existence of intermolecular and intramolecular interactions inside the proteins make the amide I band complicated and featureless so that these structural elements are not distinguished immediately (Arrondo, Muga, Castresana, & Goni, 1993, p. 36). The overlapped bands each belonging to a type of secondary structure element are needed to be resolved. The data processing part of this study contains the necessary mathematical methods to resolve the overlapped bands. Therefore, after the collection of spectra and successful buffer subtraction from the spectra, data contained in each were analyzed by several mathematical operations processed on the software Origin Pro 2017 and by curve-fitting processed on the software OPUS 7.0.

3.2.2.1 Buffer Subtraction

The absorbance spectra of the buffers used in this study were collected under the same conditions when spectra were collected for proteins in buffers. The absorbance spectra of buffers were subtracted from the absorbance spectra of proteins in buffers with a suitable subtraction factor. When deciding on the suitable choice of the subtraction factor, the spectra must result in a straight line in the region around 1900 cm^{-1} . For cases in $^2\text{H}_2\text{O}$, along with a straight line near 1900 cm^{-1} region, the Amide A band must be clearly distinguished from the water band ($3600\text{-}3200\text{ cm}^{-1}$). After a successful buffer subtraction was performed, the final spectrum was considered to be the spectrum of the protein solely in principle.

3.2.2.2 Second Derivative

This mathematical interpretation is frequently used for overlapped bands and it provides information about the secondary structure elements contained in the protein. If there are maxima (peak positions, even the hidden ones) in the original absorbance spectrum, they will appear as minima in the second derivative spectrum. All secondary structure elements emerge under amide I band ($1700\text{-}1600\text{ cm}^{-1}$) and the second derivative applied in that region will reveal the number and position of sub-bands. However, during the application of second derivative on the original spectrum, noise can be generated. In order to avoid such noise, smoothing is applied on the spectrum. Smoothing is modifying data points in order to create a mean of those points. The smoothing applied on the original spectrum was chosen to be Savitzky-Golay with a factor of 9. Thus, one point was substituted for 9 estimated adjacent data points in the spectrum. The quantitative information contained in the second derivative of the original absorbance spectrum can reveal the hidden maxima of the original spectrum. A representative second derivative spectrum can be seen in Figure 3. 2.

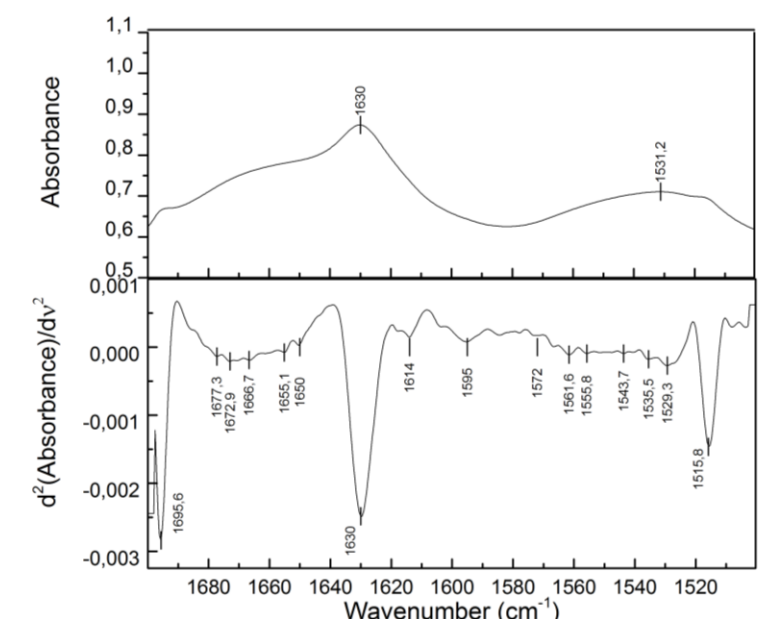


Figure 3. 2: The original absorbance spectrum (upper graph) of OmpG-16S in pH 5.5 and its second derivative spectrum (bottom graph) in the region 1700-1500 cm^{-1} . There are hidden peaks revealed in the second derivative analysis, which cannot be viewed in the original spectrum.

3.2.2.3 Side Chain Subtraction

Before curve-fitting, the absorbances of amino acid side chains under the Amide I and II bands are needed to be subtracted too. The amino acids present in the structure were subtracted from the absorbance spectrum of the protein via a Matlab-based software named Kinetics. The essential part of the procedure in Kinetics is that one needs to know the exact number of amino acids and their protonation states at the buffer condition of that particular experiment. For the proteins used in this study, the amino acid information was taken from Protein Data Bank with an ID number of 2IWW (2IWW, 2006). When subtracting the absorbances of amino acid side chains, the software takes tyrosine peak as reference since its absorbance at $\sim 1515 \text{ cm}^{-1}$ (due to C-C stretching mode of aromatic ring) can be selected easily in the spectrum. Therefore, the software scales the amino acid subtraction factor and spectra without amino acid contribution will be obtained. An example of amino acid subtraction performed in Kinetics software can be observed in Figure 3. 3.

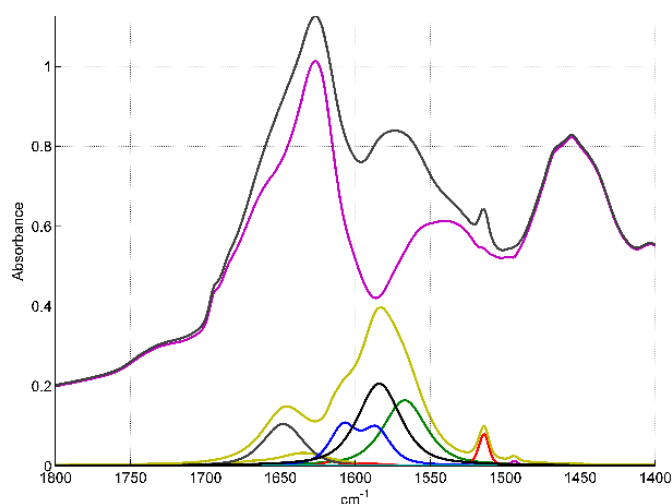


Figure 3. 3: The absorbance spectrum (black) and the amino acid subtracted spectrum (purple) of OmpG-16SL in p²H 7.5. The relatively small bands under the absorbance spectra are the subtracted absorbance bands of the amino acids present in the OmpG-16SL's structure. The total subtracted absorbance spectrum of all amino acids is the cyan band. Subtracted amino acids are aspartic acid (black), arginine (navy blue), asparagine (dark grey), glutamine (cyan-small scaled), tyrosine (red), glutamic acid (green), phenylalanine (purple). Additionally, histidine, lysine, amino end and carbon end are also subtracted. However, since they are overlapped with other amino acid bands they are not picked out.

3.2.2.4 Curve-fitting

The overlapped bands which shape the experimental absorption spectrum in the amide I region are decomposed by a mathematical method named curve-fitting. It is also called band decomposition or band-fitting. OPUS 7.0 is the software program used in the application of this technique and it is based on the non-linear least squares method. Before practising the curve-fit procedure, one needs to set the preliminary parameters for each band. These parameters are band position, band intensity, band shape and full width at half height of the band. The band position is set by the minima revealed in the second derivative of the original absorption spectrum. The band shapes which compose the overall absorption spectrum are either Lorentzian, Gaussian or Voigt. A Lorentzian line shape function is shown as if the radiation distribution is presented in the x-intercept:

$$L(x; x_0, \gamma) = \frac{1}{\pi\gamma \left[1 + \left(\frac{x-x_0}{\gamma} \right)^2 \right]} \quad (3.3)$$

where x_0 is the peak location of the radiation distribution and γ is the half-width at half-maximum. A Gaussian line shape function is then expressed as:

$$G(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\sigma^2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (3.4)$$

where μ is the mean value of the peak location and σ^2 is the variance. In Figure 3. 4, a graph can be seen comparing a Lorentzian and Gaussian function.

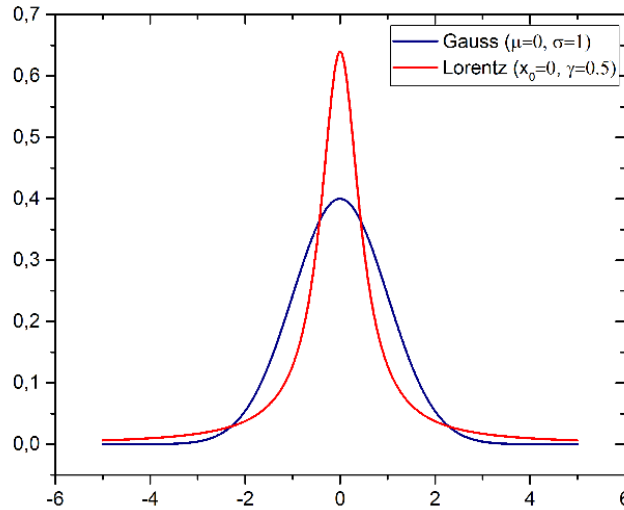


Figure 3. 4: Lorentzian and Gaussian functions. From the equations 3.3 and 3.4 above with parameters chosen $x_0=0$, $\gamma=0.5$ for Lorentz and $\mu=0$, $\sigma=1$ for Gauss respectively, the functions are drawn in Origin Pro 9.2 in between $-5 \leq x \leq 5$.

The last kind of the line shapes is Voigt which is a mixture of Lorentzian and Gaussian functions and its function is the convolution of those functions which is demonstrated in the following:

$$V(x; \sigma, \gamma) = \int_{-\infty}^{+\infty} G(x'; \sigma) * L(x - x'; \gamma) dx'$$

where x shows the frequency shift of the line center. In the equation above, if the condition is $\sigma = 0$, the Voigt function becomes only Lorentzian and if the condition is $\gamma = 0$, it only becomes Gaussian function. In the case of vibrational spectra of liquids, Voigt profile will be more suitable rather than only Lorentzian or Gaussian profile (Bradley, 2007).

In the curve-fit, the aim is to build up a spectrum (fitted spectrum) which is very similar to the original absorption spectrum. After setting parameters for several bands, during the procedure, the differences between the original and the fitted spectra will be minimized. First of all, mainly in the region 1720-1500 cm^{-1} , the numerical values for band positions and intensities which were revealed by the mathematical operations mentioned above, were introduced to the program by the program user. Secondly, band shapes were chosen to have Voigt profile and the program automatically iterated the curve-fitting and optimized these values. Then, the optimization was carried on by the program and the user until a successful match between the original and the fitted spectra is achieved. However, this match may not be successful enough if the fitted spectrum is made up of too many bands. Therefore, it may not be possible to decide whether a spectrum is finely fitted or not. In order to overcome this, a better and credible approach is initially introduced by Korkmaz (2009, p.42), then named as co-fitting by Baldassare et al. (2015, p.12600). The studies conducted on different proteins revealed that co-fitting presents substantial advantages over only absorption spectrum fitting (Baldassare, Li, Eremina, Goormaghtigh, & Barth, 2015). Co-fitting is that while the user tries to obtain similar original and fitted spectra, also tries to gain similar second derivative spectra of the original and the fitted spectra. For some proteins used in this study after applying curve-fit, secondary structure elements were assigned to the appropriate band peak positions in the amide I region. Lastly, the band areas were calculated so that the percentages of the secondary structure elements were found out.

The root-mean-square error (RMSE) of the curve-fitting was calculated by the software between the original and the fitted spectrum in order to verify the accuracy. The RMSE function is:

$$RMSE(bp) = \sqrt{\frac{1}{Q} \sum_{j=1}^Q (bp_j^{true} - bp_j^{calc})^2}$$

where bp shows either position, area or FWHH, Q indicates the number of bands, bp_j^{true} is the true value of the bp in the original spectrum and bp_j^{calc} is the calculated value in the fitted spectrum (Lorentz-Fonfria & Padros, 2004, p. 2707).

3.2.3 ATR-FTIR Spectroscopy

The main reason of working with ATR-FTIR spectroscopy instead of FTIR transmission spectroscopy is that collecting background spectrum just for once is enough. Water bands appear less in the spectrum and very little amount of samples standing horizontally can be handled even without sample preparation. Samples used can be liquids, solids or powders and, for liquid samples, their dynamic nature can be monitored by changing environmental conditions (temperature, pH) of the sample.

The IR beam is pointed onto a material having a high refractive index (the material is called internal reflection element-IRE, which is a diamond in our setup). The sample is placed on the top of this IRE material (Figure 3. 5). The radiation must be pointed at an angle above the critical angle so that it is totally internally reflected. The radiation experiences multiple reflections (in our setup it performs 10 attenuated total reflections) after interacting with the crystal and these multiple total internal reflected waves establish evanescent waves and they penetrate the sample (within a few microns) since the sample is in close contact with the IRE (Figure 3. 6). These evanescent waves will be attenuated (decay exponentially) when the sample absorbs its energy. Then, the decrease in the radiation intensity reach the detector and the corresponding absorption spectrum is generated by the computer.

The established evanescent wave will penetrate a depth in the sample part and that depth of penetration is determined by the incident angle of the IR beam and the refractive indices of the IRE and the sample (n_1 and n_2).

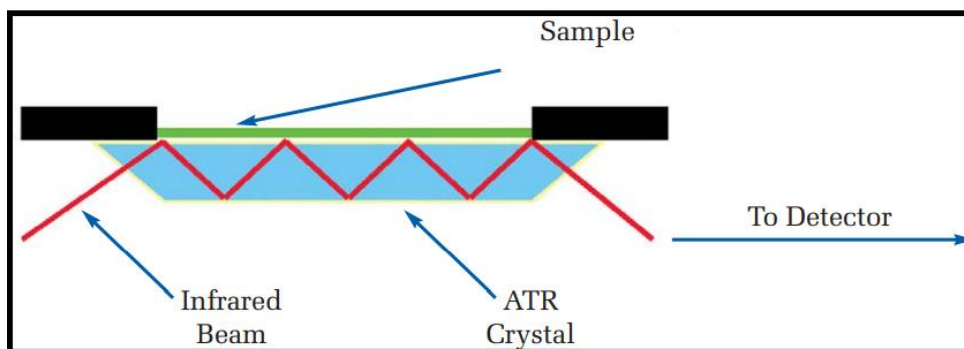


Figure 3. 5: An ATR sample cell and the path of the IR beam (Perkin Elmer, 2005).

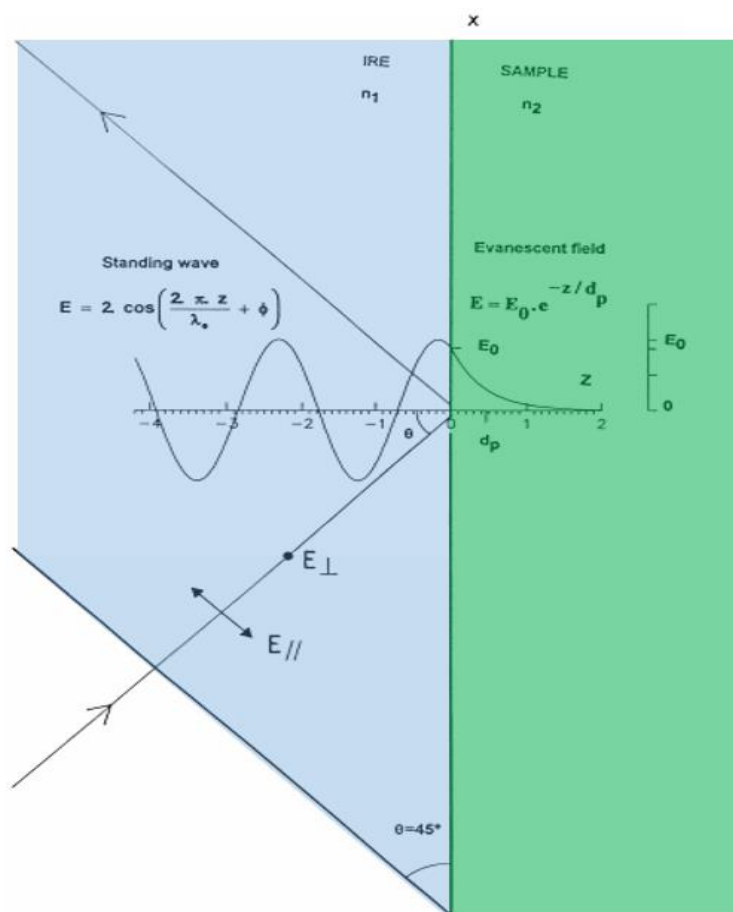


Figure 3. 6: The figure demonstrates the total internal reflection of an incoming IR wave penetrating into the IRE and the evanescent wave established by that reflection (Goormaghtigh, Raussens, & Ruyschaert, 1999, p. 108).

The penetration depth is expressed as

$$d_p = \frac{\lambda}{2\pi n_1 ((\sin \theta)^2 - \left(\frac{n_2}{n_1}\right)^2)^{1/2}}$$

where λ is the wavelength of the IR radiation, n_1 is the refractive index of the IRE, θ is the incident angle and n_2 is the refractive index of the sample. Diamond, ZnSe, germanium or silicon are the possible IRE elements since they all have high refractive indices and are transparent to IR light.

The ATR-FTIR cell used in the experiments is developed by taking an ATR-IR microdialysis cell as reference (Dzafic, Klein, Screpanti, Hunte, & Mantele, 2008, p. 103). This ATR-IR microdialysis cell (Figure 3. 7) is utilized for monitoring *in situ* changes in the structure of OmpG-WT and mutants during buffer exchange ($^1\text{H}/^2\text{H}$ exchange).

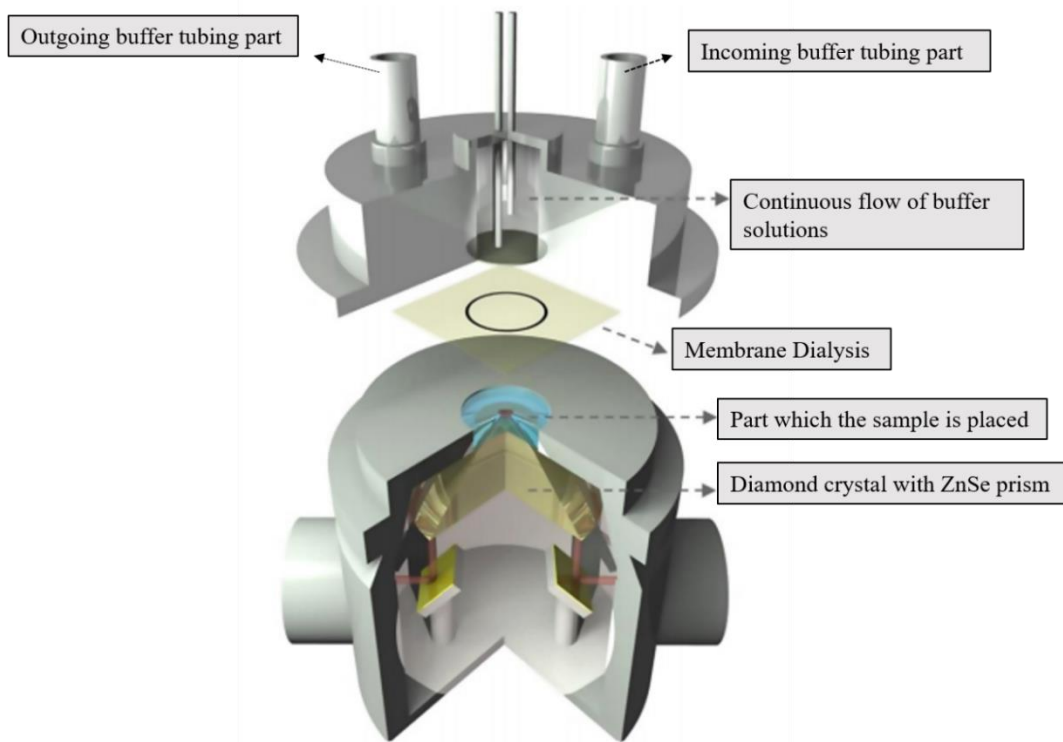


Figure 3. 7: ATR-IR microdialysis cell which is mounted into the FTIR spectrometer (Dzafic et al., 2008, p.103).

In preparation, 15 μ l (determined by checking the absorbance signal of the protein) of protein sample is placed on top of the diamond crystal. The sample compartment is covered by a membrane dialysis having 25 kDa pore size which is a semi-permeable synthetic membrane. Water, ions and anything smaller than 25 kDa size can freely pass through but larger molecules are blocked. In our application, it enables the buffer solution to enter the sample compartment but blocks the sample (protein) being swept away by the rush of buffer solution. Then, the flow-through cell compartment is tightly closed, so that the buffer solution perfuses sample while the sample is scanned at certain intervals. The continuous flow of buffer solution is performed by a peristaltic pump which has a constant speed of 5 rpm.

3.2.3.1 $^1\text{H}/^2\text{H}$ Exchange

ATR-IR spectroscopy measures almost only the absorbance of the sample instead of the bulk water since the sample has direct contact with the ATR crystal. This surface layer technique is used for determining which molecular groups are accessible by the buffer solution during hydrogen/deuterium ($^1\text{H}/^2\text{H}$) exchange. The measurements gained after this exchange experiment provide information on dynamics of the protein sample and the solvent accessibility (Raschke & Marqusee, 1998, p. 80). In the process of exchange, the deuteration of the amide hydrogens will be monitored. The $\text{N}-^1\text{H}$ groups that are located on the surface of the protein sample will exchange with D_2O and become $\text{N}-^2\text{H}$ immediately after being exposed to D_2O buffer (since they are accessible by the solvent). The $\text{N}-^1\text{H}$ groups located at the flexible (unstable) regions will exchange slowly than the ones on surface of the protein. The $\text{N}-^1\text{H}$ groups located in the hydrophobic core of the protein will exchange very slowly or will not exchange at all.

For the experiment, Thermo Nicolet 6700 FTIR spectrometer equipped with an MCT detector that is cooled with liquid N_2 is used. The background spectrum is scanned before placing any protein onto the ATR plate and used for all spectra. Firstly, the sample is perfused by 0.1 M NaP_i H_2O buffer with 0.5 % OG (pH 5.5 or pH 7.5) until obtaining the same absorbance spectrum over and over again. Next, the buffer is

changed to 40 ml of 0.1 M NaPi $^2\text{H}_2\text{O}$ buffer with 0.5 % OG (p ^2H 5.5 or p ^2H 7.5). 35 ml of that buffer is used up until the exchange is complete inside cell. 5 ml of the buffer is left for closed circulation. The exchange (circulation) is continued for each sample for 18 hours. In order to scan the sample at certain intervals a Macro (an example in Figure 3. 8) is written in Omnic 8.0 program. For the first hour of the exchange, 120 spectra were recorded in total with 32 scans averaged for 2 cm^{-1} resolution for each spectrum. In the second hour of the exchange 12 spectra were recorded, each with 128 scans and 2 cm^{-1} resolution. For the third hour, 6 spectra were recorded, each with 256 scans and 2 cm^{-1} resolution. After that, during 12 hours, 16 spectra were recorded with the same properties of the previous scan. Moreover, the buffers were kept inside a box filled with dry ice so that the degradation of the protein sample was prevented over the course of buffer exchange.

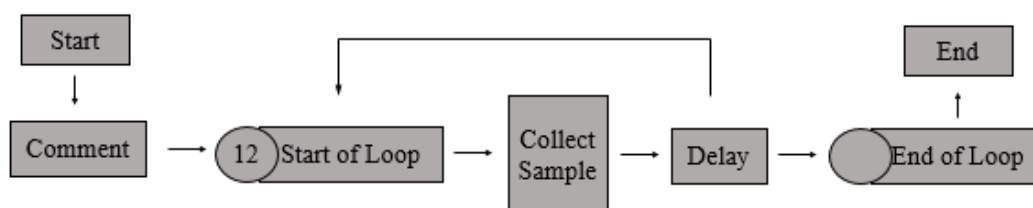


Figure 3. 8: A representative loop written for the second hour of the exchange in Macro and used for automated scanning.

3.2.3.2 Analysis of Spectra in the $^1\text{H}/^2\text{H}$ Exchange Experiment

During the exchange, there are two varying conditions that can be monitored on the spectra. First of all, in comparison of the sequential absorbance spectra, the water bands gradually disappear and the deuterium bands appear instead over time. Second of all, N-H in-plane bending vibrations positioned in the amide II region (1600-1500 cm^{-1}) causes C-N stretching vibrations to experience a downshift by 100 cm^{-1} after deuteration and that newly appearing band is called amide II' band (centered at 1450

cm⁻¹) (Figure 3. 9). In the analysis, 150 spectra for each protein sample were buffer subtracted. At the moment that the ¹H/²H exchange is triggered, there will be several spectra having bending vibrations of H-O-²H which position at 1450 cm⁻¹. These spectra must be carefully buffer corrected with the buffer spectra that differ in the ratio of ¹H to ²H. Then, the area under the amide I and amide II bands were calculated and the ratio of those areas (Equation 3.5) was used for determining the hydrophilic percentage of the protein.

$$\text{Ratio} = \frac{\int \text{Amide II}}{\int \text{Amide I}} \quad (3.5)$$

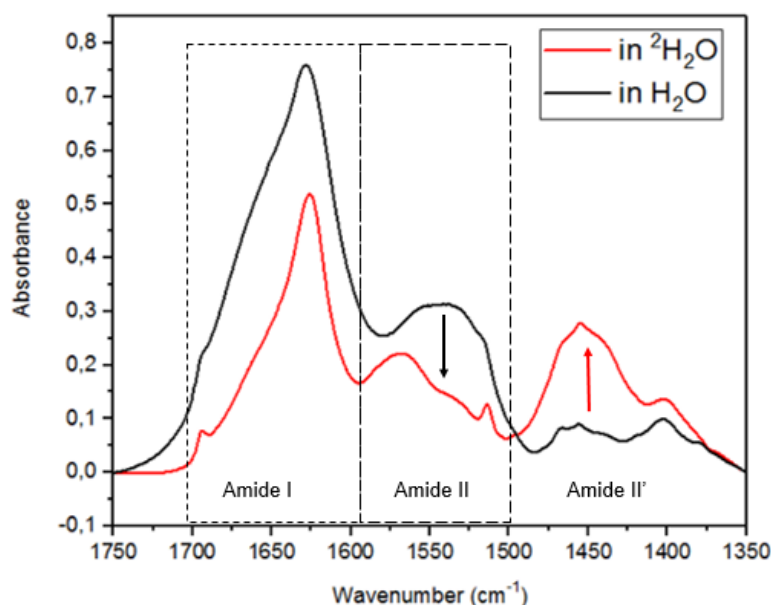


Figure 3. 9: The absorbance spectrum of OmpG-16S in pH 7.5 (black line) and in p²H 7.5 (red line). The spectrum in ²H₂O buffer is recorded after an hour of exchange. While amide II band intensity decreases, the amide II' intensity increases after deuteration.

CHAPTER 4

RESULTS AND DISCUSSION

In 1989, OmpG was synthesized by a mutation in *Escherichia coli* K-12. Isolation, and biochemical and genetic characterization of OmpG revealed that it resembles a porin-like protein and allows the passage of small hydrophilic molecules (Misra & Benson, 1989, p. 4105). OmpG is different than other membrane proteins in sequence and has more negatively charged residues in its sequence (Fajardo et al., 1998, p. 4454). OmpG can be overexpressed, and purification and refolding can be applied on it (Conlan, Zhang, Cheley, & Bayley, 2000, p. 11851).

2-D projection maps of two crystallized conformations of OmpG taken by electron cryo-microscopy with 6 Å resolution revealed that the structure of OmpG is β -barrel with a diameter of nearly 2.5 nm and the size of the barrel suggests that OmpG is made up of 14 strands (Behlau, Mills, Quader, Kühlbrandt, & Vonck, 2001, p. 75). X-ray crystallography study showed that loop L6 in OmpG plays an active role in the opening and closing mechanism of the pore in response to pH changes. At neutral pH, two histidine residues His231 and His261 that are exposed to solvent are uncharged and loop L6 extends out into the extracellular medium, leaving the pore entrance open for transport. At acidic pH (below pH 6), these residues become positively charged, and on account of repelling each other, the hydrogen bonds between strands S11 and S13 are unzipped. Thus, the loop L6 bends onto the pore and makes the pore closed (Yildiz et. al., 2006, p. 3703).

4.1 Band Assignment on FTIR Absorbance Spectra

An absorbance spectrum example of OmpG-16S mutant in p^2H 5.5 is demonstrated in Figure 4. 1. Band positions in wavenumbers and their corresponding functional groups

are shown with color coding.

Protein OmpG is a membrane protein, thus it needs a hydrophobic environment to protect its own hydrophobic parts. During purification of the protein, such a medium is provided by detergent molecules. Detergents (and also lipids) form hydrophobic cavities/layers in water. CH₂ and CH₃ asymmetric and symmetric stretching modes shown in Figure 4. 1 are due to the carbon chain in the structure of the detergent OG, forming the hydrophobic region. These molecular groups also have other modes of vibrations in the region 1500-1200 cm⁻¹; however, bands are relatively smaller than the bands of other groups arising in the same region. The PO₂⁻ antisymmetric and symmetric stretching modes are originating from the sodium phosphate buffer (NaPi_i).

The most significant bands that are used for the structure analysis of protein samples are amide I and amide II bands. Amide I band (1700-1600 cm⁻¹ region) is made up of mainly C=O stretching. C-N stretching, C-CN deformation and very few in-plane N-H bending vibrations originating from protein structure also contribute to amide I band in lesser amounts.

Each secondary structure element and, their corresponding position in wavenumbers both in water and deuterium are given in (Table 4. 1). These structures have different characteristics due to the H-bonding strength between the C=O and N-H groups, thus corresponding bands are located at different positions under the amide I band.

Amide II band is the second most important band in protein structure studies. It is located next to the amide I band (1600-1500 cm⁻¹ region). Amide II band shape and components change dramatically depending on the buffer conditions. In water, the band is mainly composed of N-H in-plane bending with less contribution from C-N and C-C stretching modes. These modes are mainly due to the protein secondary structure but there are side chain contributions too. However, after deuteration, N-H modes leave the amide II band due to heavier reduced mass upon deuteration. C-N mode also leaves the amide II band and shifts down to ~ 1455 cm⁻¹, forming the amide II' band. As a result, amide II band in heavy water is a smaller band than the same band in water. How much smaller it is, depends on the specifics of the protein, i.e., the proportion of the backbone in hydrophilic environment to that in hydrophobic

environment. This fact is used as the basis for the $^1\text{H}/^2\text{H}$ exchange experiments. On the other hand, amide II band is generally not a good choice for protein structure investigation due to its inherent complexity.

Lastly, there are amino acid side chain absorbances that are hidden but they are located in amide I and amide II region and are listed in Table 4. 2.

In the light of these references and previous studies in literature mentioned thus far, the objective of this study is to compare OmpG-WT protein, OmpG-16S and OmpG-16SL mutants among each other and reveal the resemblances and differences in conformations in their structure with respect to pH changes, thermal stabilities, ratio of $^1\text{H}/^2\text{H}$ exchange rates.

Table 4. 1: Protein secondary structure elements and their position ranges in wavenumbers compiled from the literature (Kong & Yu, 2007, p. 552) (Arrondo et al., 1993, p. 35) (Surewicz, Mantsch, & Chapman, 1993, p. 391) (Barth, 2007, p. 1083). An average position is also shown in parenthesis for each structure.

Secondary structures	in H_2O (cm^{-1})	in $^2\text{H}_2\text{O}$ (cm^{-1})
α -helix	(1654) 1648-1655	(1652) 1648-1655
Antiparallel β -sheet	1623-1641 1674-1695	1620-1638 1682-1696
Parallel β -sheet	1630 1645	1632 1648
Turns	1667-1685	1681-1665
Random Coil	(1654) 1642-1660	(1645) 1639-1660

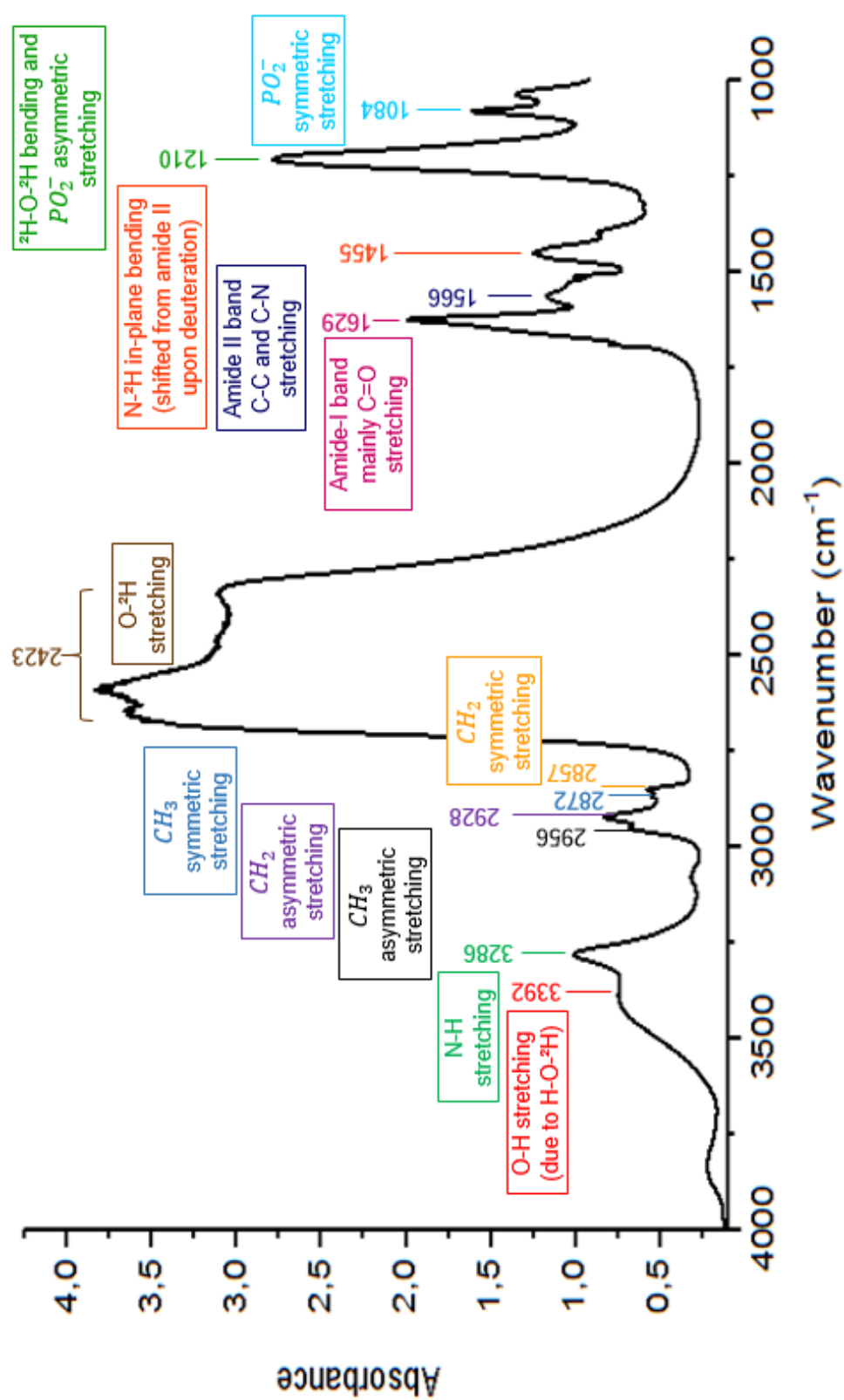


Figure 4. 1: A characteristic absorbance spectrum of OmpG-16S mutant in deuterated buffer.

Table 4. 2: Amino acids side chain positions in wavenumbers (Barth, 2000, pp. 145-152). The bands appear in the range 2551-724 cm^{-1} in H_2O and 1849-865 cm^{-1} in $^2\text{H}_2\text{O}$ but this table specifically focuses on amide I and II regions.

Side Chain (Its Mode)	Spectrum Ranges (cm^{-1})	
	in H_2O	in $^2\text{H}_2\text{O}$
Asn (C=O stretching)	1677-78	1648
Arg (CN_3H_5^+ asymm. stret.)	1672-73	1608
Gln (C=O stret.)	1668-1687	1635-1654
Arg (CN_3H_5^+ symm. stret.)	1633-1636	1586
HisH₂⁺ (C=C stret.)	1631	1600,1623
Lys (NH_3^+ asymm. bending)	1626-1629	1201
Tyr-OH (C-C stret. & C-H bending)	1614-1621	1612-1618
Trp (C=C & C-C stret.)	1622	1618
Tyr-O⁻ , (C-C stret.)	1599-1602	1603
Tyr-OH (C-C stret.)	1594-1602	1590-91
Gln (NH_2 bending)	1586-1610	1163
HisH (C=C stret.)	1575,1594	1569,1575
Asp (COO^- asymm. stret.)	1574-1579	1584
Glu (COO^- asymm. stret.)	1556-1560	1567
Lys (NH_3^+ symm. Bending)	1526-27	1170
Tyr-OH (C-C stret. & C-H bending)	1516-1518	1513-1517

4.2 pH-Dependent Conformational Changes and Secondary Structures of OmpG-WT, OmpG-16S and OmpG-16SL

The structure investigation of OmpG-WT and mutants were carried out by analyzing their FTIR spectra in the 1720-1600 cm^{-1} region. In the four graphs shown below (Fig. 4.2-5), the IR absorbance spectra are shown for OmpG-WT, OmpG-16S and OmpG-16SL in the extended region between 1720-1350 cm^{-1} in order to generally observe how the spectra look like in different buffer conditions and how they provide information about the structure of these proteins.

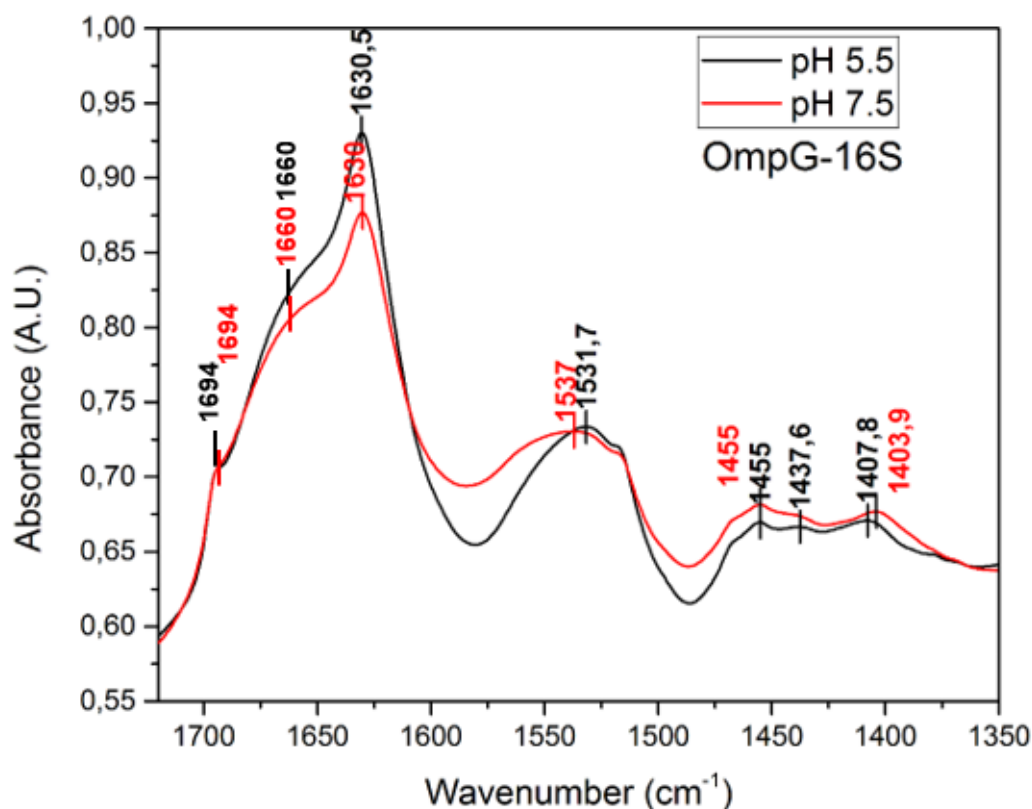


Figure 4. 2: Absorbance spectra of OmpG-16S at pH 5.5 and pH 7.5 at $\sim 22^\circ\text{C}$.

Absorbance spectra of OmpG-16S at pH 5.5 (black line) and at pH 7.5 (red line) in H₂O buffer are shown in Figure 4. 2. Amide I band has peaks at ~ 1630 cm⁻¹ (low frequency β -sheet signal) and ~ 1694 cm⁻¹ (high frequency β -sheet signal) in H₂O. There is a bump around 1660 cm⁻¹. The band peaked at ~ 1535 cm⁻¹ is the amide II band. Since the absorbance spectra profiles of OmpG-WT and OmpG-16SL in H₂O share similar profiles and positions with OmpG-16S, data are not shown. The main obstacle of measurements in H₂O buffer is that the protein bands and water band due to H-O-H bending mode are located very close to one another and thus overlap. Therefore, any change in the protein band like a spectral shift becomes difficult or even impossible to observe. As a solution to this problem, same experiments were also performed in heavy water routinely for all the samples.

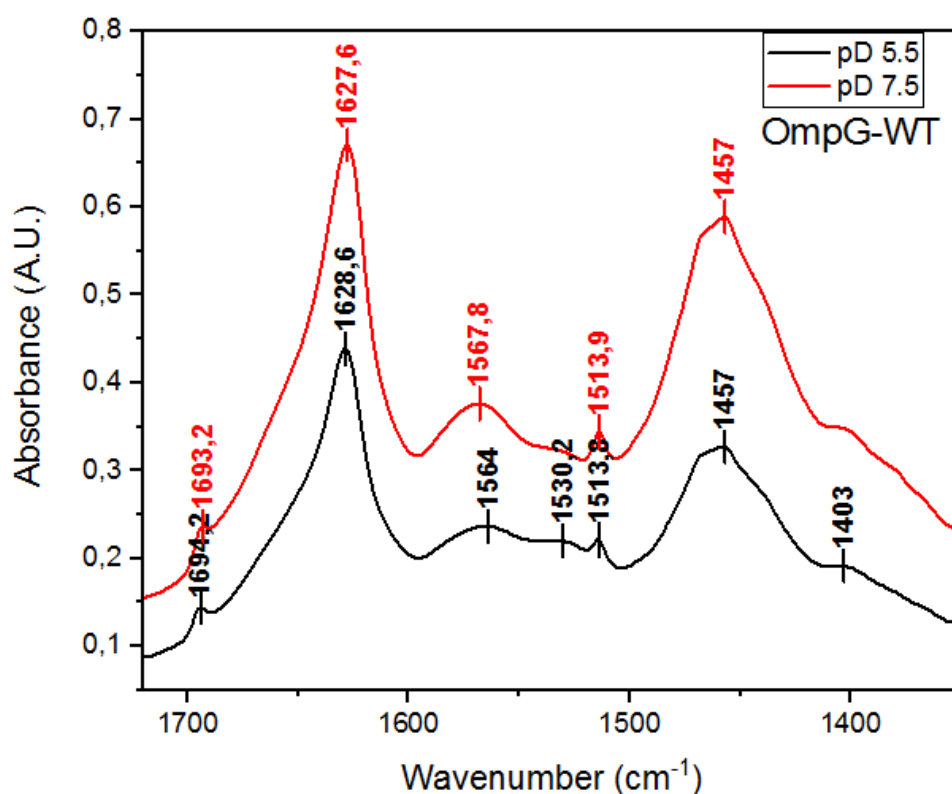


Figure 4. 3: Absorbance spectra of OmpG-WT at p²H 5.5 and p²H 7.5 at ~ 22 °C.

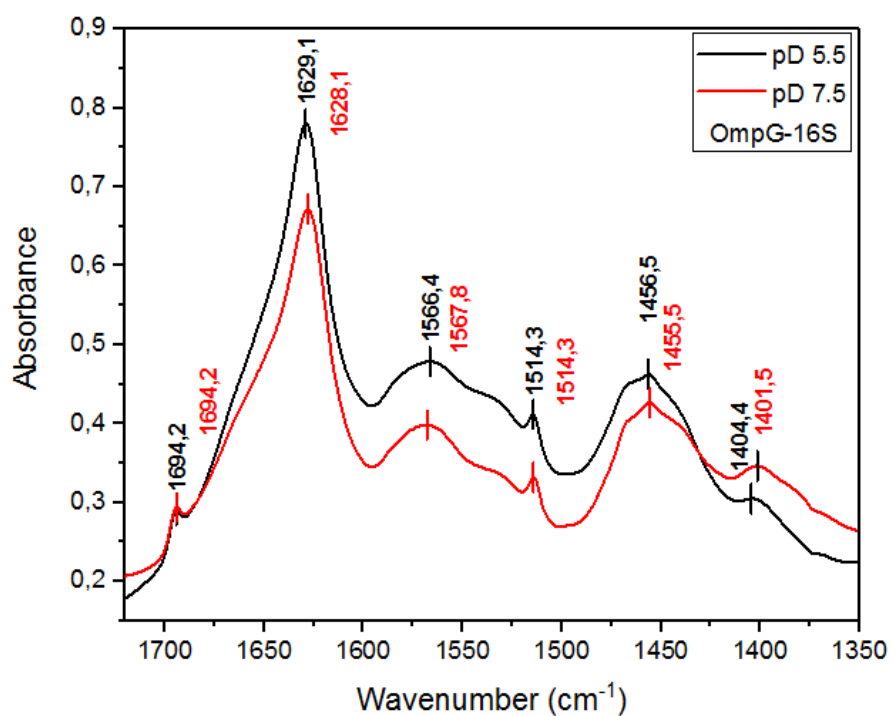


Figure 4. 4: Absorbance spectra of OmpG-16S at p²H 5.5 and p²H 7.5 at ~ 22 °C.

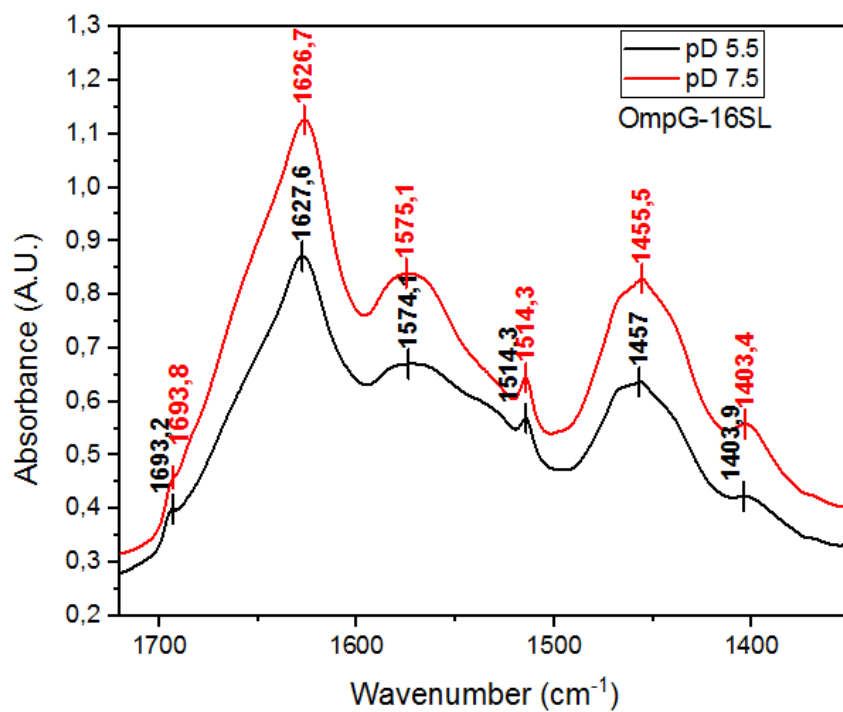


Figure 4. 5: Absorbance spectra of OmpG-16SL at p²H 5.5 and p²H 7.5 at ~ 22 °C.

The overall band in between 1700 and 1600 cm^{-1} is the amide I absorption band having two peaks at high and low frequencies (Figure 4. 3, Figure 4. 4, and Figure 4. 5). While the peak around 1694 cm^{-1} shows antiparallel β -sheet signal at high frequency, the peak around 1628 cm^{-1} indicates the low frequency component of the same structure. The low frequency β -sheet signal in $^2\text{H}_2\text{O}$ is shifted by $\sim 2 \text{ cm}^{-1}$ as the proteins experience medium change from H_2O to $^2\text{H}_2\text{O}$ (Figure 4. 2). Moreover, the bump seen at 1660 cm^{-1} in H_2O disappears in $^2\text{H}_2\text{O}$ because the unordered structure band position in H_2O is at 1660 cm^{-1} , while it is at 1640 cm^{-1} in $^2\text{H}_2\text{O}$. Besides, while arginine side chain signal appears at 1680 cm^{-1} in H_2O , it downshifts to 1580 cm^{-1} in $^2\text{H}_2\text{O}$ and makes the amide I region less complex. As can be deduced here, the amide I band in H_2O is much more complex than in $^2\text{H}_2\text{O}$. Even if the protein in H_2O is dried, there will be remains of water due to bound water molecules to protein, which will appear in the IR spectra.

Amide II band is also affected by the buffer change and is now called amide II'. Absorbance of the band is observed to be decreased due to the reasons explained in the previous section. A significant peak is now seen at 1514 cm^{-1} that is due to the C-C ring stretching mode of tyrosine side chain absorbance.

The overall IR spectra profile of proteins (both in H_2O and in $^2\text{H}_2\text{O}$) look similar. However, a shift is captured when examining the spectra only in $^2\text{H}_2\text{O}$. When the environment of OmpG-WT and mutants is changed from acidic to neutral (5.5 \rightarrow 7.5), the low frequency signal of amide I band downshifts by $\sim 1 \text{ cm}^{-1}$. This downshift is linked to the formation of new H-bonds; thus, indicates an increase in β -sheet content as the channel opens at neutral environment (Korkmaz-Özkan, Köster, Kühlbrandt, Mantele, & Yildiz, 2010, p. 57). For OmpG-WT, it is reported that as the His231 and His261 deprotonates at neutral pH, hydrogen bonding is formed between the strands S11 and S13, which causes the opening of loop L6.

OmpG-WT, OmpG-16S and OmpG-16SL show that kind of downshift in their low frequency β -sheet position in the amide I region meaning that they all respond to the pH change in similar manner.

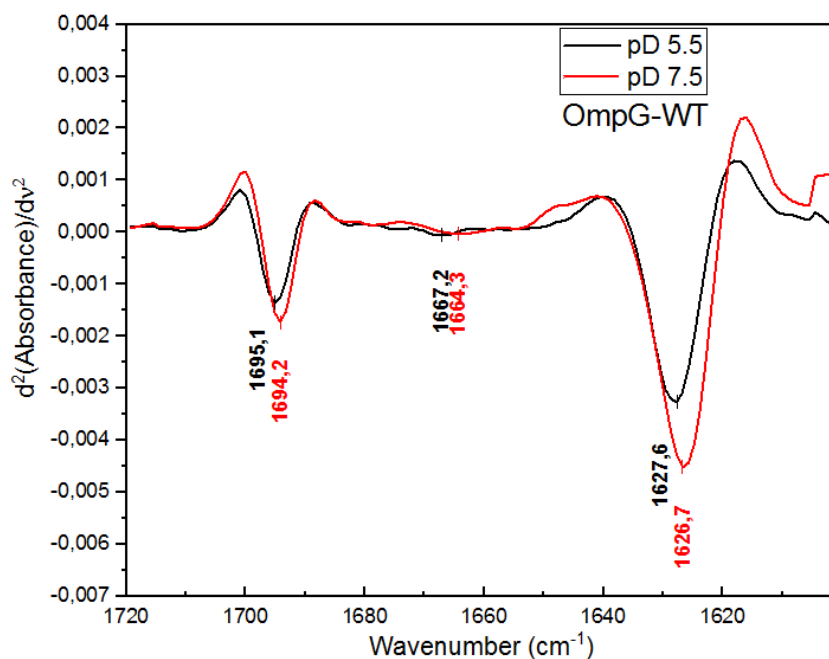


Figure 4. 6: Second derivative of OmpG-WT absorbance in p²H 5.5 and p²H 7.5 at ~ 22 °C shown in the region 1720-1600 cm⁻¹.

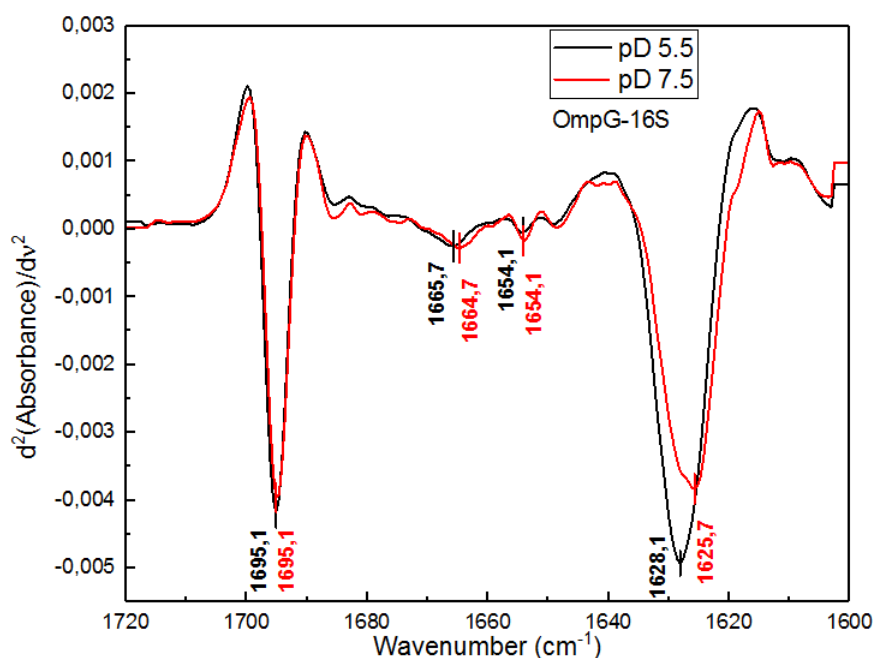


Figure 4. 7: Second derivative of OmpG-16S absorbance in p²H 5.5 and p²H 7.5 at ~ 22 °C shown in the region 1720-1600 cm⁻¹.

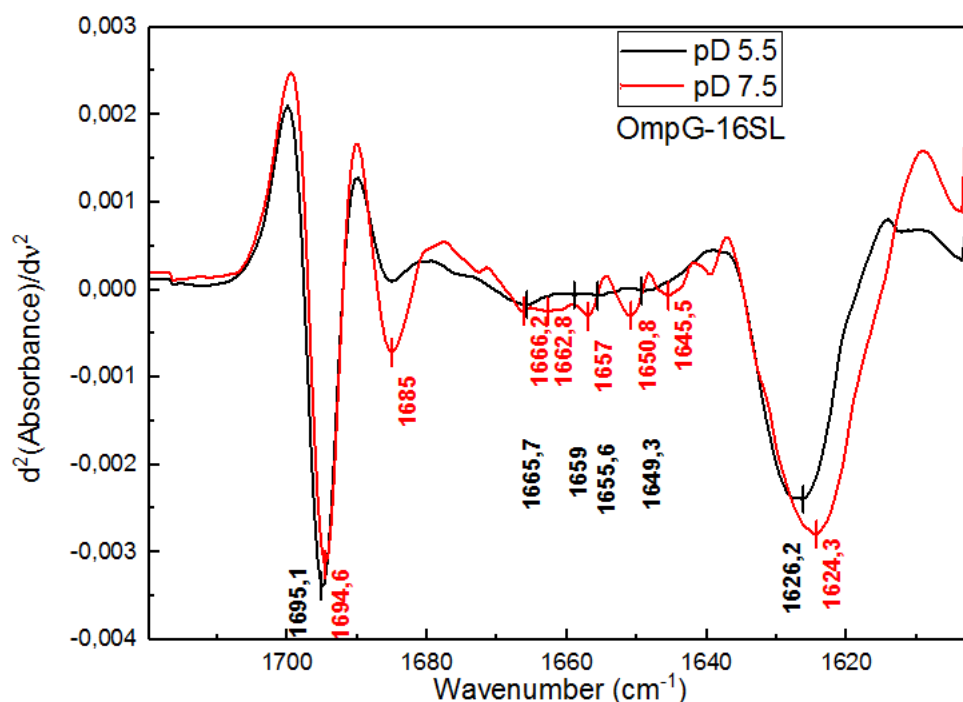


Figure 4. 8: Second derivative of OmpG-16S absorbance in p²H 5.5 and p²H 7.5 at ~ 22 °C shown in the region 1720-1600 cm⁻¹.

The downshift of the low β -sheet frequency is much more obvious in the second derivative profiles of the proteins upon pH increase (Figure 4. 6-8). These processed spectra reveal hidden differences and peaks in the amide I region but specifically between 1685-1645 cm⁻¹. While OmpG-WT and OmpG-16S present minor dissimilarities in that region, OmpG-16SL shows much more structural alterations due to pH increase. Since there is only one individual turn of α -helix in the structure of the of OmpG-WT and mutants, these differences in the region between 1685-1645 cm⁻¹ are caused possibly by turns, loops and unordered structures.

In the amide I region (1700-1600 cm⁻¹), there are overlapping bands, which contain information on secondary structure elements. However, there are many bands representing various parts of a protein structure, all located in this 100 cm⁻¹ window of the spectrum. This brings a natural complexity. There are some Mathematical tools to resolve a highly overlapped band like amide I. Second derivative calculation,

Fourier self-deconvolution and curve-fitting are the most frequently used tools in this regard. In this study, the percentages of the secondary structure elements in the protein were revealed by curve-fitting method as described in the previous chapter.

Samples only in $^2\text{H}_2\text{O}$ are treated under curve-fitting method owing to two reasons. The first reason is the very high water absorbance in the amide I region. In order to gain meaningful and unambiguous information under this region, one must subtract water absorbance very carefully. Even if this can be achieved, the subtraction will not be considered successful since the bound water molecules surrounding the samples will never be fully subtracted. Hence, these will create additional overlapping bands lowering the accuracy of the method. However, in heavy water, these problems do not exist and a careful buffer subtraction will lead to a successfully subtracted spectrum. The second reason is that when in $^2\text{H}_2\text{O}$ buffer, Arg side chain modes downshift to the intersection of amide I and II ($\sim 1600\text{ cm}^{-1}$) and let the user apply the method under a less intricate region. In addition to all these, when in $^2\text{H}_2\text{O}$ buffer, there are side chains located in the hydrophobic parts of the protein that are not affected from the buffer and some of them located in the hydrophilic regions. Because of that, amino acid side chain subtraction procedure (described in the previous chapter) is also applied to the amide I and II' regions.

For a successful curve-fitting, the RMS value must be as low as possible along with a similar co-fitting profiles of the fitted and the original absorption spectrum. For OmpG-WT and OmpG-16S, successful co-fitting profiles with a low RMS value were not achieved. However, for OmpG-16SL, fitting with an RMS of 0.000091 for p^2H 5.5 and 0.000141 for p^2H were achieved along with their similar second derivative profiles. Since no successful curve-fitting results were attained for OmpG-WT and OmpG-16S, the percentages of the secondary structure elements were taken from literature in order to compare them with OmpG-16SL (Table 4. 3).

The overlapping bands of OmpG-16SL in the amide I region are shown with the percentages of their relative secondary structure elements (Table 4. 4). The curve-fitting overlapping bands and second derivative comparisons can also be viewed graphically (Figure 4.9 and 4.10).

Table 4. 3: The percentages of the secondary structure elements of OmpG-WT and OmpG-16S in p²H 5.5 and p²H 7.5 obtained from curve-fitting results are taken from literature (Korkmaz, van Pee, & Yildiz, 2015, p. 76).

% Structural Weight				Secondary Structure Elements
OmpG-WT		OmpG-16S		
p ² H 5.5	p ² H 7.5	p ² H 5.5	p ² H 7.5	
66	71	58	59	β -sheet
25.5	20.4	28.4	22.6	Unordered structure
8.4	8.6	13.4	18.5	Turns

Table 4. 4: The curve-fitting results of OmpG-16SL in the amide I region. The positions of overlapping bands, the integrated areas of these bands and the secondary structure element assignments according to the band positions are shown.

p ² H 5.5		p ² H 7.5		Secondary Structure Elements
Band position(cm ⁻¹)	Band area (%)	Band position(cm ⁻¹)	Band area (%)	
1627	51.3	1623	39.0	β -sheet (low frequency component)
-	-	1634.5	11.0	
1640	2.4	1644	13.4	Loops/coils/helix
1646	14.1	1655	9.5	
1657	12.4	-	-	
1666	6.1	1665	7.8	Turns
1675	7.3	1674	7.3	
-	-	1684	5.2	β -sheet (high frequency component)
1688	5.1	1688	0.6	
-	-	1692	0.6	
1695	1.2	1694	1.3	

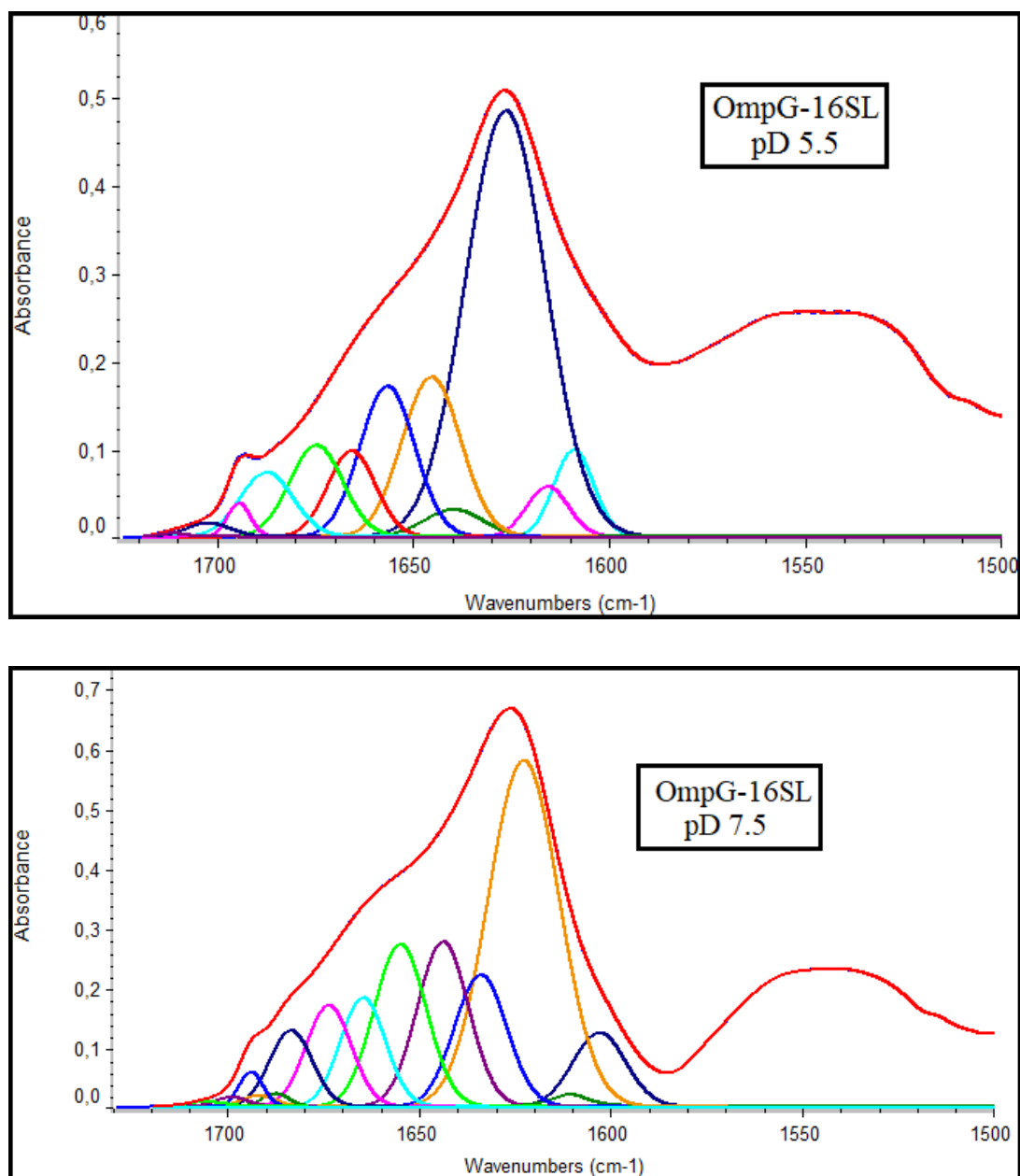


Figure 4. 9: Curve-fitting results of OmpG-16SL (1720-1500 cm⁻¹ region). The original absorption spectrum (blue intersects with the red spectrum) and the fitted spectrum (red) together with the overlapping bands belonging to the fitted spectrum are demonstrated for both p²H values.

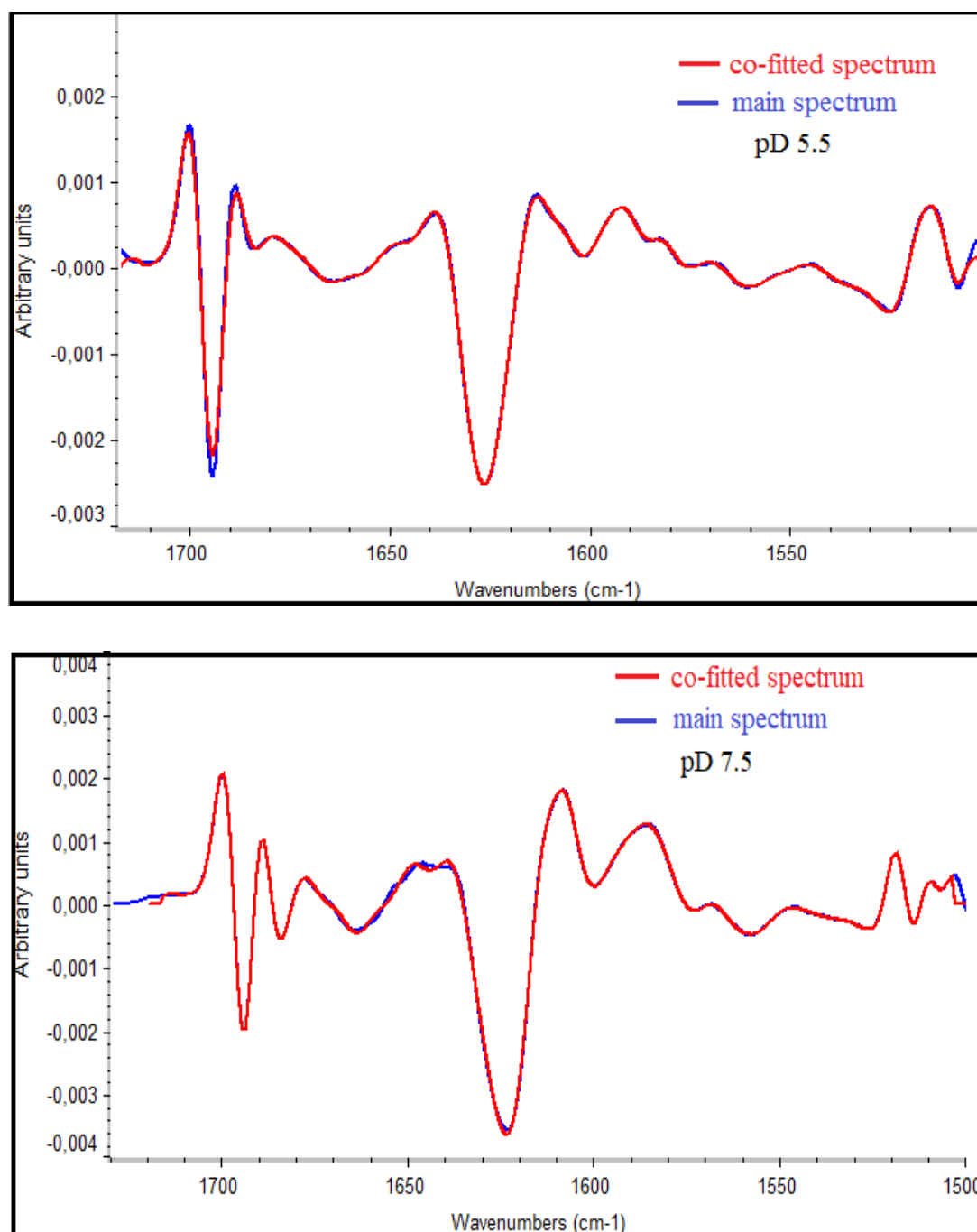


Figure 4. 10: The second derivative comparisons of the original absorption (main) spectrum and the co-fitted spectrum (1720-1500 cm⁻¹ region) of OmpG-16SL.

As previously mentioned, OmpG-WT, OmpG-16S and OmpG-16SL proteins are composed of 281, 319 and 325 amino acids, respectively. The curve-fitting results in Table 4. 3 signifies that 66% of OmpG-WT structure in p²H 5.5 and 71% in p²H 7.5 are constituted of β -sheet structure. In curve-fitting results of OmpG-16S, 58% in p²H 5.5 and 59% in p²H 7.5 belong to β -sheet. It seems that the β -sheet structure percentages drop in OmpG-16S. However, after calculating the number of amino acids involved in β -sheet structure for both proteins and conditions, it is possible to see that similar number of residues form this structure. While 185 residues are involved in the β -sheet structure in p²H 5.5, 199 residues take part in p²H 7.5 for OmpG-WT. For OmpG-16S, 185 and 188 residues take part in β -sheet structure in p²H 5.5 and p²H 7.5, respectively. This result suggests that in the neutral environment, additional 38 residues introduced to OmpG-16S primary sequence mostly participate in the unordered structure and turns rather than β -sheet.

As can be calculated from Table 4. 4; 57.6% and 57.7% of OmpG-16SL compose β -sheet structure, %28.9 and %26.9 compose unordered structure and %13.4 and %15.1 compose turns at p²H 5.5 and p²H 7.5, respectively. 187 and 188 residues take part in β -sheet structure in p²H 5.5 and p²H 7.5, respectively. Results suggest that the β -sheet structure in two mutants are very similar in terms of both the proportion and the behavior at different pH conditions. Moreover, the percentage of turns in OmpG-16SL is the same with OmpG-16S in acidic environment but decreases according to OmpG-16S in neutral environment.

When comparing these three proteins, the β -sheet content shows a slight or no increase in OmpG-16S and OmpG-16SL in neutral pH conditions. Unordered structure content decreases but turn content increases (a very slight increase in OmpG-WT) as the environment changes from acidic to neutral p²H. Hence, OmpG-16S and OmpG-16SL show similar content-wise profile with OmpG-WT at pH 5.5.

Data at hand is difficult to interpret for neutral pH conditions for the mutants. The fact that the three proteins respond to the p²H changes were already evidenced by the shift of the amide I band peak positions in absorption spectra and in the second derivative profiles.

For OmpG-WT, this shift has been reported as the indication of newly formed H-bonds at neutral p²H (Korkmaz et. al, 2015, p. 75), which in turn stabilizes the β -barrel. A similar shift is observed for both mutants suggesting the formation of new H-bonds; however, curve-fitting results do not support the idea of more number of residues taking part in β -sheet structure at neutral pH.

For OmpG-16SL, there are two high frequencies (at 1688 and 1695 cm⁻¹) and one low frequency (at 1627 cm⁻¹) β -sheet component in p²H 5.5 as indicated in Table 4. 4. Slightly different folding types of β -sheet structures can occur, which are caused by the different distances of hydrogen bonding. These types correspond to different modes and thus, result in different but close frequencies like two high frequency β -sheet components. In p²H 7.5, this time there exists two low frequencies (at 1623 and 1634.5 cm⁻¹) of β -sheet components. The 1627 cm⁻¹ band at p²H 5.5 splits into two in p²H 7.5. While %39 of the band at 1627 cm⁻¹ is shifted down by 4 cm⁻¹, %11 is shifted up by 7.5 cm⁻¹. The band having a high percentage at 1623 cm⁻¹ is ascribed to strongly interacting, tightly packed β -sheet structure and the weakly interacting-newly formed hydrogen bonds (Yildiz et. al., 2006, p. 3703) are assigned to the band at 1634.5 cm⁻¹ (Korkmaz et. al., 2015, p. 76). It has been reported that this kind of splitting and H-bond interaction occur in a similar manner both in OmpG-WT and OmpG-16S.

Switching from p²H 5.5 to 7.5, the β -strand content increases by %5 (corresponding to 14 residues) in OmpG-WT and ~%1 (3 residues) in OmpG-16S but there is almost no increase in OmpG-16SL as displayed in Table 4. 3 and Table 4. 4. Even though in OmpG-16S and OmpG-16SL there is scarcely any increase, the splitting of the low frequency β -sheet signals in switching from p²H 5.5 to 7.5 is evident for newly formed H-bonds for all three proteins.

4.3 Thermal Stability of OmpG-WT, OmpG-16S and OmpG-16SL

Increasing the temperature alters the conformation of a protein causing its denaturation. By tracing the changes of the positions of its amide I band components, denaturation of a protein caused by increasing temperature can be followed (Chenin, et al., 1999, pp. 1525-1526). With increasing temperature, the transition temperature

values are indicative parameters related with the thermal stability of proteins. Proteins that undergo transitions due to increasing temperature can be compared among each other by looking at the changes in their overall IR spectra and their transition temperature values. Transitions occurring at higher temperatures will correspond to more stable proteins. Proteins containing high β -sheet structure and make transitions at relatively high temperatures are known as stable proteins (Korkmaz-Özkan et. al., 2010, p. 59). Since OmpG-WT, OmpG-16S and OmpG-16SL contain high β -sheet structure, they are also expected to be stable proteins too.

For each of OmpG-WT, OmpG-16S and OmpG-16SL in $^2\text{H}_2\text{O}$ buffers, temperature ramp (T-ramp) IR spectra were collected from 22 °C to 106 °C with 2 °C increments. T-ramp IR spectra of OmpG-WT in p^2H 5.5 is investigated in the region 1720 – 1500 cm^{-1} (Figure 4. 11). While the black lines indicate the absorbance before transition, red lines show the absorbance in the course of transition and beyond. At 22 °C (the black line at the top), while the high frequency β -sheet signal is at 1695 cm^{-1} , the low frequency β -sheet peak is at 1629 cm^{-1} . At the highest temperature gained (the red line at the bottom), while the high frequency β -sheet peak position is at 1685 cm^{-1} , the low frequency β -sheet peak position is at 1620 cm^{-1} . By increasing the temperature, the peak positions belonging to high and low frequency β -sheets shift down by $\sim 10 \text{ cm}^{-1}$.

The second derivative of the absorbance spectra of OmpG-WT in p^2H 5.5 are demonstrated in the region 1720-1500 cm^{-1} (Figure 4. 12). The second derivative graph also clearly indicates the downshift of the high and low frequency β -sheet signals by 10 cm^{-1} . For OmpG-WT in p^2H 7.5 buffer, OmpG-16S and OmpG-16SL in $^2\text{H}_2\text{O}$ buffers, the IR spectra and the second derivative graphs of T-ramp experiments also show the same downshift (data is not shown here). Therefore, temperature stability graphs of OmpG-WT, 16S and 16SL in pH 5.5, pH 7.5 and in p^2H 5.5, p^2H 7.5 are drawn by following only the position of Amide I peak as the temperature is increased from 22 to 106 °C.

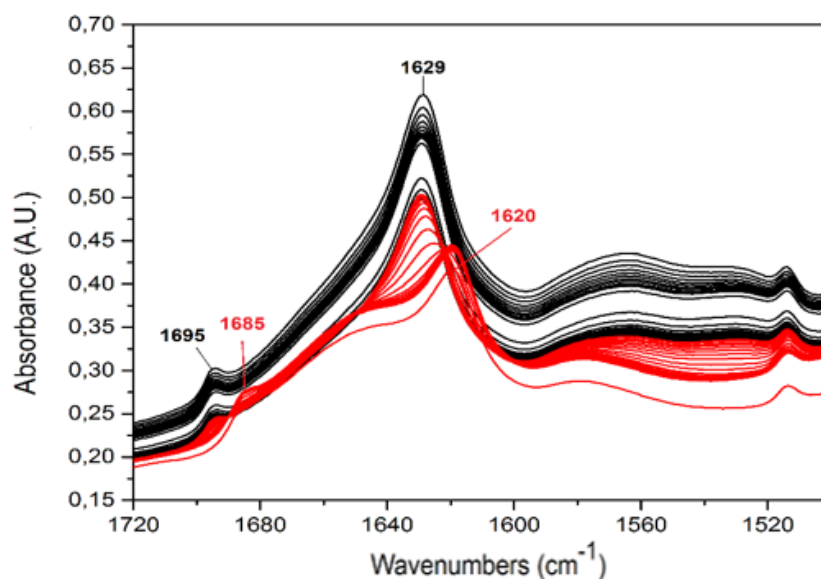


Figure 4. 11: IR spectra of OmpG-WT in p²H 5.5 in the region 1720-1500 cm⁻¹ during a temperature ramp. All recorded spectra are shown (from 22 °C to 106 °C). Black lines indicate the spectra before transition. Red lines indicate the spectra during and after the transition state.

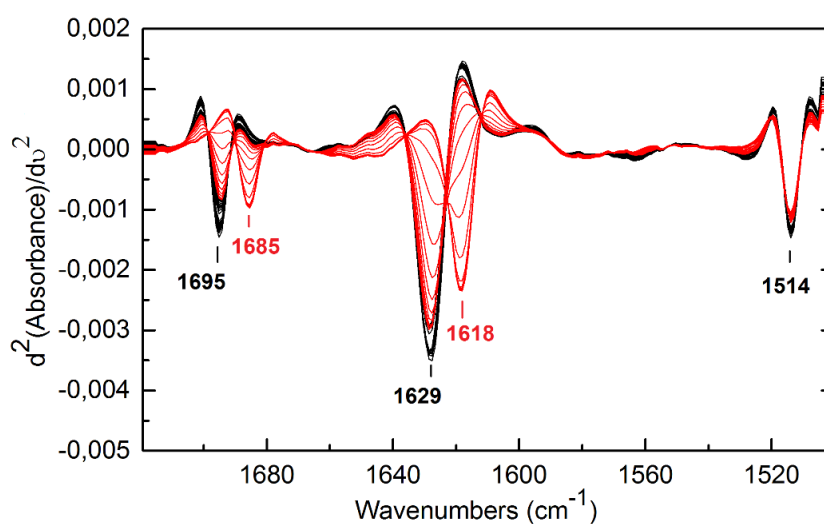


Figure 4. 12: The second derivatives of the absorbance IR spectra of OmpG-WT in p²H 5.5 (IR spectra in Figure 4. 11). While the black lines show the second derivatives of the IR spectra before transition, the red lines show the second derivatives during and after the transition.

In the course of temperature increase, newly occurred peaks between wavenumbers 1620 and 1615 cm^{-1} are correlated with the unfolding of the proteins, which are resulted from aggregated structures (Haris & Severcan, 1999, p. 216). As observed in Figure 4. 11 and Figure 4. 12, 10 cm^{-1} downshift of high and low frequency β -sheet signals show changes in the secondary structure which is attributed to the unfolding of the protein due to high intermolecular contacts (Arrondo & Goni, 1999, p. 372).

Since OmpG-WT, OmpG-16S and OmpG-16SL in $^2\text{H}_2\text{O}$ buffers demonstrate that downshift, they all aggregated at increasing temperatures.

Additionally, a 3-D plot is drawn to clearly demonstrate the diminishing intensities of amide I and II bands (Figure 4. 13). The intensity decrease of the amide II band is the sign of the deuteration increase (the protein structure becomes more flexible) and its tertiary structure is disturbed and broken down upon temperature increase (Dzafic, Klein, Goswami, Kühlbrandt, & Mantele, 2009, p. 734).

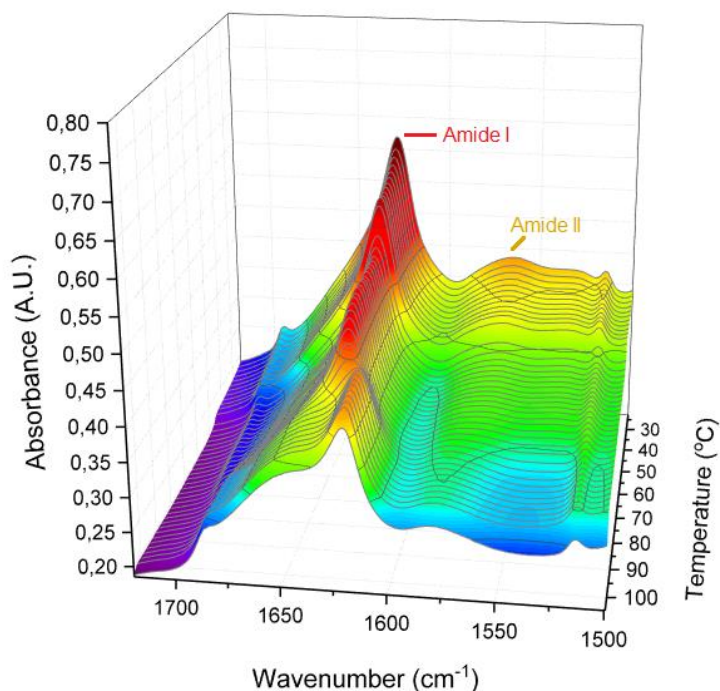


Figure 4. 13: T-ramp 3-D plot of protein OmpG-WT in p^2H 5.5 in the region 1720-1500 cm^{-1} . All IR spectra are shown from 22 to 106 $^{\circ}\text{C}$.

In a transition temperature graph (Figure 4. 14), where band position vs temperature is plotted, in order to comment on the changes in the conformation of a protein, three separate regions are expected to show up. If OmpG-WT graph is reviewed for both environments, three separate regions can be distinguished immediately (Figure 4. 14). In the first region, from 20 °C to 75 °C, there is no change in the amide I peak position of the protein, which is indicative of no change in the conformation of the protein and thus, the protein maintains its native state. In the second region between 75 °C and 85 °C, the transition is initiated which leads to conformational changes. After 85 °C, the amide I peak position is stabilized, which means the total aggregation of the protein.

As can be viewed from the figure, they all undergo temperature induced transitions, which differ between 64 and 86 °C (Figure 4. 14). The exact transition temperature values (which is the midpoint temperature value in the transition region- T_t) of three proteins in $^2\text{H}_2\text{O}$ buffers are listed in Table 4.5.

Table 4. 5: Transition temperature values of the three proteins at both p^2H values.

Proteins	Transition Temperatures	
	p^2H 5.5	p^2H 7.5
OmpG-WT	82 °C	86 °C
OmpG-16S	82 °C	74 °C
OmpG-16SL	74 °C	64 °C

By analyzing the graphs shown in Figure 4. 14, all three proteins in acidic and neutral p^2H environment undergo transitions at different temperature values. The transition profiles and temperature values of OmpG-WT in acidic and neutral environment are similar to each other.

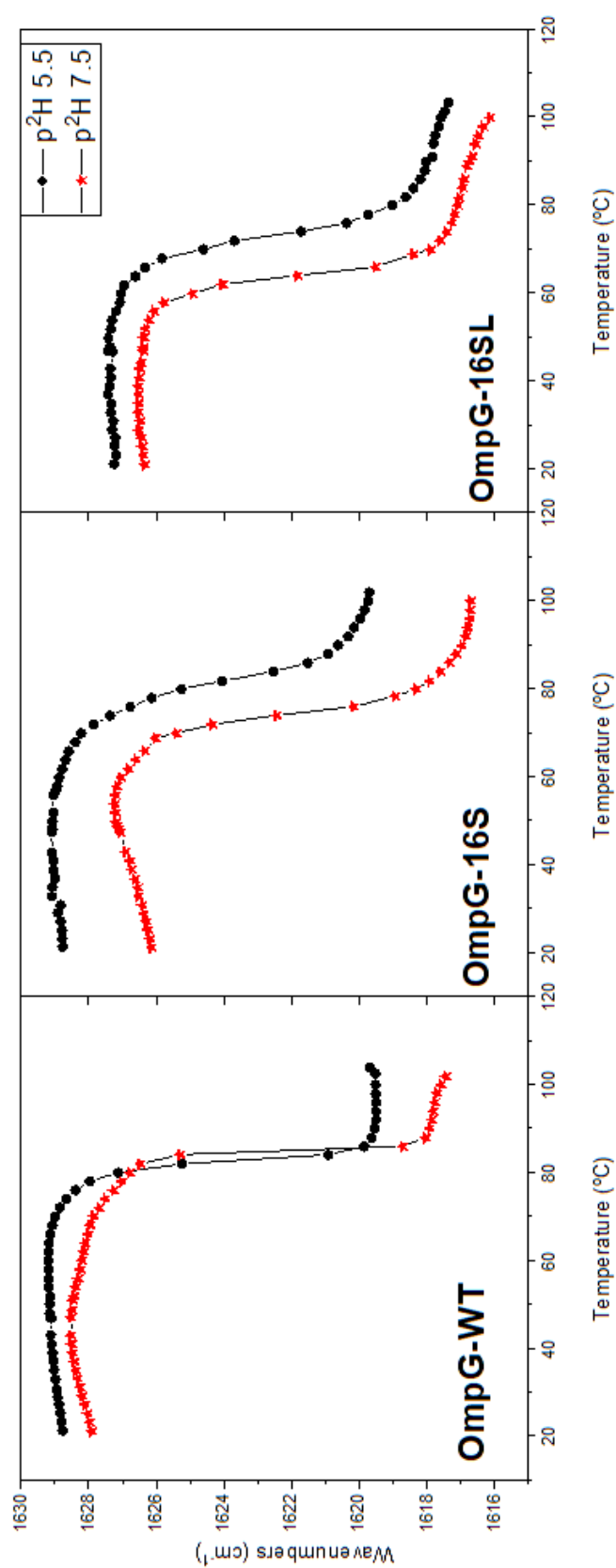


Figure 4. 14: Comparative thermal stability spectra of OmpG-WT and mutants in both p²H values. Amide I peak position followed from absorbance spectra and plotted against temperature at acidic and neutral p²H conditions.

OmpG-16S and OmpG-16SL are more stable in acidic environment than in neutral environment. OmpG-WT in acidic and neutral surroundings and OmpG-16S in acidic surrounding have close transition temperature values, and they show the highest thermal stability profiles in p²H environment. While OmpG-16S in neutral environment and OmpG-16SL in acidic environment have the same transition temperature, OmpG-16SL in neutral environment indicates the least stable case among all. In brief, OmpG-16S and OmpG-16SL are more stable in the closed state than in the open state. Close transition temperature values of OmpG-WT and OmpG-16S in acidic environment verify that they fold in a similar way.

4.4 ¹H/²H Exchange Properties of OmpG-WT, OmpG-16S and OmpG-16SL

When the OmpG protein is subjected to buffer solution, the hydrophilic regions (loops and turns) will be exposed to the buffer solution. Remaining residues buried inside the detergent micelle constitute the hydrophobic core (β -barrel) of the protein. The protein backbone accessible to the buffer solution is immediately affected by the exchange of buffer from water to heavy water. The ¹H/²H exchange experiments performed inside an ATR microdialysis cell provide information on the ratio of the buffer accessible parts to inaccessible parts of proteins. For this purpose, the investigation of spectra is carried out in the amide II region since the C-N stretching modes downshift by 100 cm⁻¹ after the sample is deuterated. Thus, the solvent accessibility of OmpG-WT, OmpG-16S and OmpG-16SL are determined by the ratio of amide II/amide I band areas, which corresponds to the exchanged protein backbone percentage at the time of the particular measurement. This way, it is possible to follow the exchange dynamics of a protein on the scale of minutes.

The OmpG-WT and mutants are exposed to ¹H/²H exchange at pH 5.5 and 7.5 conditions separately. The kinetics of deuteration of the protein backbone can be viewed during 18 hours of constant deuteration conditions (Table 4. 6).

Table 4. 6: The percentages of $^1\text{H}/^2\text{H}$ exchange of the three proteins for pH 5.5 and 7.5 after 18 hours.

Proteins	Closed State (pH/p²H 5.5) (%)	Open State (pH/p²H 7.5) (%)
OmpG-WT	26	38
OmpG-16S	40	39
OmpG-16SL	40	41

The increase of $^1\text{H}/^2\text{H}$ exchange percentages of the first 180 minutes are illustrated in Figure 4. 15. Vertical scale gives a measure of the deuterated backbone percentage with respect to the whole structure. As can be deduced from the graphs, the exchange occurs in two phases: fast and slow exchange. In the first 20 minutes, the fast phase of the exchange is completed. In this stage, the protons on the hydrophilic parts of the proteins (loops and turns) are immediately deuterated. Loops and turns are already in the buffer solution; thus, they are the first structures affected from water to heavy water switch. After 20 minutes, the exchange is very slow (nearly at a constant rate) until the end of the exchange process, which indicates that the protons on the buried parts of the proteins (less flexible and stable parts: i.e., β -sheet) are not in immediate contact with heavy water. At the end of 180 minutes, the maximum exchange is around 40% for the three proteins or in other words, 40% of the protein backbone is accessible to buffer and the remaining 60% is not. This technique is particularly useful when the protein in question changes conformation upon an external trigger. In such cases, buffer accessibility of the backbone also changes and FTIR spectroscopy is one of the unique techniques to provide information on the dynamics of proteins.

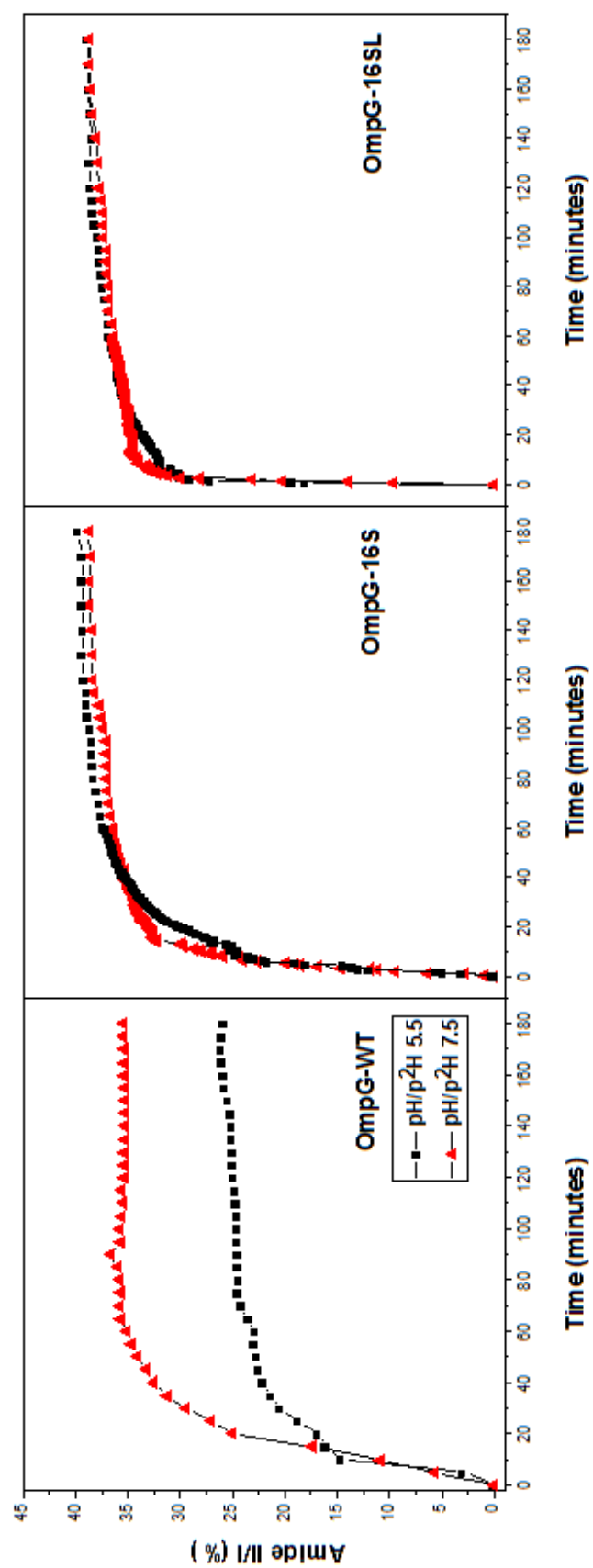


Figure 4. 15: $^1\text{H}/^2\text{H}$ exchange percentage profiles of the three proteins.

For the wild type protein, 26% of the backbone is accessible to buffer at acidic pH, which corresponds to the closed state of the protein. At neutral pH, the protein channel becomes accessible to buffer too, and thus, the exchange rate is increased to 38%. This is a spectral proof that the protein attains two forms as open and closed and these two forms are different in terms of the buffer accessibility. According to high resolution X-ray crystallography structure of the protein, loop L6, which blocks the protein channel entrance at acidic pH (closed state) but allows water and maltodextrin entrance at neutral pH (open state) is the reason of this increase in exchange rate. (Korkmaz-Özkan et. al., 2010, p. 61). On the other hand, increased exchange rate is only seen in OmpG-WT but not in mutants.

When comparing these three proteins, the exchange rate of OmpG-16SL is the fastest among all. The differences in conformations (loops and turns) may have resulted in different fast phase exchanges. However, the exchange ratio of OmpG-WT and mutants are very close to each other except OmpG-WT in acidic environment (closed state). The very close exchange rates of OmpG-16S and OmpG-16SL for both buffer conditions indicate that the buffer entered the interior of the protein channel immediately as also shown in open state of OmpG-WT.

The main structure of OmpG-16S and OmpG-16SL are nearly the same but with some exceptions. The difference in OmpG-16SL structure originates from 6 added amino acids to loop L6 and 7 replaced arginines (located inside the pore, which are replaced by lysines). The reason of replacing lysines with arginines is that arginine has a longer side chain than lysine and thus, the side chains of arginines extend through the center of the enlarged pore. However, the exchange ratios of OmpG-16SL at p²H 5.5 and p²H 7.5 shows that the effects of these added and replaced amino acids are not monitored since their ratio are very close to the exchange ratios of OmpG-16S.

The total amino acid numbers that were exchanged with deuterium can be computed since how many amino acids compose each protein in total is known. By using the exchange percentages, for OmpG-WT, 73 amino acids are deuterated at acidic pH (closed state) and 107 amino acids are deuterated at neutral pH (open state). A similar calculation yields 128 amino acids at acidic pH and 124 amino acids at neutral pH for

OmpG-16S. For OmpG-16SL, 130 amino acids are deuterated at acidic pH and 133 amino acids are deuterated at neutral pH. All these amino acids that are exposed to water and then heavy water, thus exchanged, are suggested to be positioned in the hydrophilic regions of the proteins. It seems while for OmpG-16S and OmpG-16SL amino acid numbers that are exposed and deuterated are the same for two buffer conditions, for OmpG-WT more than 30 amino acids deuterated in p^2H 7.5 compared to acidic pH.

The increase in the number of the deuterated amino acids at p^2H 7.5 in OmpG-16S and OmpG-16SL is caused by the added amino acids. 17 of the added 38 amino acids to the OmpG-16S structure are deuterated meaning that half of the added amino acids are located in the hydrophilic region.

CHAPTER 5

CONCLUSION

In this study, pH-dependent structural differences and functional properties of genetically engineered mutant proteins, OmpG-16S and OmpG-16SL were investigated and the results were compared to those of the wild type protein OmpG.

For these proteins that were surrounded by acidic and neutral environment both in water and heavy water, transmission FTIR spectroscopy technique was used to reveal the changes in the structure of these proteins and to determine their thermal stabilities. For monitoring their dynamic properties, ATR-FTIR spectroscopy technique was utilized. In the latter technique, a new spectrometer attachment (placed onto the ATR diamond) was developed, which enabled in situ $^1\text{H}/^2\text{H}$ exchange. During each experiment of exchange, the collection of mass amount of spectra was successfully maintained by an automated spectra collection computer code written in Omnic 8.0 software.

According to the absorbance spectra, OmpG-WT, OmpG-16S and OmpG-16SL share similar profiles and close wavenumber positions at both pH values. In heavy water, their overall spectra also look similar. When the p^2H was switched from acidic to neutral, positions of their amide I bands downshifted, which indicates that new H-bonds were formed in all three proteins; hence, β -sheet content increased. This shift and similarities in spectral profile point out that all proteins respond to pH change but at different degrees.

The second derivative profiles of proteins showed that as the pH increases, the structural differences among them mostly occurred due to unordered structures, loops and turns. However, they also showed that OmpG-16SL is the one, which goes through more structural alterations.

Curve-fitting results demonstrated that at acidic environment, all proteins contain same number of residues participated in β -sheet structure. However, at neutral environment, even though the β -sheet content of OmpG-WT was increased, there was only a slight or no increase in mutants. On the other hand, in absorbance spectra, all proteins had shown that they responded to pH change and thus, new H-bonds formed, which corresponds to β -sheet content increase. This increase did not show up in the curve-fitting results of OmpG-16SL. When considering β -sheet content, two mutants showed similarities in behavior and proportion at both conditions. The proportions of unordered structures and turns of OmpG-16SL at acidic environment are almost the same with OmpG-16S but differs at neutral pH. In OmpG-16SL, there are much more unordered structures formed at neutral pH when compared with OmpG-16S. Moreover, at neutral pH, the splitting of the low frequency β -sheet signals also was seen in all proteins, which is the indicative of a more rigid β -barrel. This splitting effect also seen in OmpG-16SL may indicate that there might be new H-bonds formed at some small degree.

The results of thermal stability experiments showed that OmpG-WT at neutral pH is slightly stable than at acidic pH due to increased β -sheet content. Even if OmpG-16S is as stable as OmpG-WT at acidic pH, it is less stable at neutral pH. Thus, the structure of OmpG-16S is similar with OmpG-WT at acidic pH but differs at neutral pH. At both buffer conditions, especially at neutral pH, OmpG-16SL is less stable than OmpG-16S and OmpG-WT. Since OmpG-16SL contains more unordered structure, this result is expected and it confirms the results of the curve-fitting analysis.

$^1\text{H}/^2\text{H}$ exchange experiments showed that a different exchange rate was only observed in OmpG-WT not in mutants. This is the sign of the pH-dependent gating mechanism of the channel. However, at both buffer conditions, buffer penetrated deep into the pores of both mutants immediately just as in OmpG-WT at open state.

The close exchange ratios of mutants and OmpG-WT at neutral pH suggest that the pores of OmpG-16S and OmpG-16SL were enlarged even if at some small degree. The number of residues deuterated at both pH conditions is nearly the same. This leads to the conclusion that 7 arginines (lysines were replaced with arginines) in OmpG-16SL structure were not enough to create an electrostatic barrier so that it did not prevent the

excess water to enter the pore at acidic pH.

The role of loop L6 in OmpG-16SL structure to block and unblock the gate entrance could not be understood in this study. Although the length of L6 is extended by six additional amino acids to be able to block a pore with enlarged diameter, the function of the loop is controlled by the H-bonds established between the strands S11 and S13. According to the curve-fitting results, the number of these H-bonds are not enough when compared to the wild type protein. Although the transmission-IR experiments showed a shift in amide I band peak position upon pH exchange, we do not have enough evidence to evaluate this shift as the sign of formation new β -sheet structure because if the strands were able to completely zip and unzip through H-bonding, loop L6 would do its normal function unless there are other factors that are restraining the mobility of the loop.

In conclusion, the intention of creating an enlarged pore with same structural stability and gating behavior is only partially fulfilled. The second attempt in this line of motivation, OmpG-16SL, has less ordered structure and thus, shows less structural stability compared to the wild type. The gating mechanism is only partially working. The globular structure goes through a conformational change towards opening/closing; however, the number and degree of changes are less than required to attain a fully closed form.

REFERENCES

- 2IWW. (2006, 08 14). Retrieved from RCSB PDB: www.rcsb.org/pdb/explore/explore.do?structureId=2IWW
- Arrondo, J. L., & Goni, F. M. (1999). Structure and dynamics of membrane proteins as studied by infrared spectroscopy. *Progress in Biophysics & Molecular Biology*, 72, 367-405.
- Arrondo, J. L., Muga, A., Castresana, J., & Goni, F. M. (1993). Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy. *Progress in Biophysics and Molecular Biology*, 59, 23-56.
- Background. (2009, 04 06). Retrieved from Warwick: <https://www2.warwick.ac.uk/fac/sci/chemistry/research/dixon/dixongroup/members/msrhar/research/background/>
- Baldassare, M., Li, C., Eremina, N., Goormaghtigh, E., & Barth, A. (2015). Simultaneous Fitting of Absorption Spectra and Their Second Derivatives for an Improved Analysis of Protein Infrared Spectra. *Molecules*, 20, 12599-12622.
- Barth, A. (2000). The infrared absorption of amino acid side chains. *Progress in Biophysics and Molecular Biology*(74), 141-173.
- Barth, A. (2007, June 19). Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta*, 1767, 1073-1101.
- Behlau, M., Mills, D. J., Quader, H., Kühlbrandt, W., & Vonck, J. (2001). Projection structure of the monomeric porin OmpG at 6 Å resolution. *Journal of molecular biology*, 305, 71-77.

- Bradley, M. (2007). *Curve Fitting in Raman and IR Spectroscopy: Basic Theory of Line Shapes and Applications*. Retrieved from Thermo Fisher: https://www.thermofisher.com/content/dam/tfs/ATG/CAD/CAD%20Documents/Application%20&%20Technical%20Notes/Molecular%20Spectroscopy/Raman/Raman%20Instruments/AN50733_E.pdf
- Campbell, I. (1984). *Biological Spectroscopy*. Menlo Park, California: Benjamin/Cummings Pub.
- Chenin, R., Iloro, I., Marcos, M. J., Villar, E., Shnyrov, V. L., & Arrondo, J. L. (1999). Thermal and pH-induced conformational changes of a B-sheet protein monitored by infrared spectroscopy. *Biochemistry*, 38(5), 1525-1530.
- Conlan, S., & Bayley, H. (2003). Folding of a monomeric porin, OmpG, in detergent solution. *Biochemistry*, 42, 9453-9465.
- Conlan, S., Zhang, Y., Cheley, S., & Bayley, H. (2000). Biochemical and biophysical characterization of OmpG: a monomeric porin. *Biochemistry*, 39(39), 11845-11854.
- Dzafic, E., Klein, O., Goswami, P., Kühlbrandt, W., & Mantele, W. (2009). Infrared spectroscopic study of the structural and functional properties of the Na⁺/H⁺ antiporter MjNhaP1 from *Methanococcus jannaschii*. *Biochimica et Biophysica Acta*, 1787, 730-737.
- Dzafic, E., Klein, O., Screpanti, E., Hunte, C., & Mantele, W. (2008, August 20). Flexibility and dynamics of NhaA Na⁺/H⁺-antiporter of *Escherichia coli* studied by Fourier transform infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*(72), 102-109.
- Experiment 7: IR Spectroscopy*. (2007). Retrieved from Columbia University Intensive Seminars In Modern Chemistry: www.columbia.edu/cu/chemistry/ugrad/hssp/EXP_7.html
- Fabian, H., & Mantele, W. (2002). Infrared Spectroscopy of Proteins. In J. M. Chalmers, & G. Peter, *Handbook of Vibrational Spectroscopy*. John Wiley & Sons.

- Fajardo, D. A., Cheung, J., Ito, C., Sugawara, E., Nikaido, H., & Misra, R. (1998). Biochemistry and Regulation of a Novel Escherichia coli K-12 Porin Protein, OmpG, Which Produces Unusually Large Channels. *Journal of Bacteriology*, 180(17), 4452-4459.
- Freeman, S. (2005). Protein Structure and Function. In *Biological Science* (2nd ed., p. 59). Upper Saddle River, New Jersey: Pearson, Prentice Hall.
- Goormaghtigh, E., Cabiaux, V., & Ruyschaert, J.-M. (1994). Determination of Soluble Membrane Protein Structure by Fourier Transform Infrared Spectroscopy. I. Assignment and Model Compounds. *Subcellular Biochemistry*, 23, 329-361.
- Goormaghtigh, E., Raussens, V., & Ruyschaert, J.-M. (1999, March 19). Attenuated total reflection infrared spectroscopy of proteins and lipids in biological membranes. *Biochimica et Biophysica Acta*(1422), 105-185.
- Haltia, T., & Freire, E. (1995). Forces and factors that contribute to the structural stability of membrane proteins. *Biochimica et Biophysica Acta* , 1228, 1-27.
- Haris, P. I., & Severcan, F. (1999). FTIR spectroscopic characterization of protein structure in aqueous and non-aqueous media. *Journal of molecular catalysis B: enzymatic* , 7, 207-221.
- Harris, D. C. (2007). *Quantitative Chemical Analysis* (7th ed.). New York: W.H. Freeman and Company.
- Jackson, M., & Mantsch, H. H. (1995). The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure. *Critical Reviews in Biochemistry and Molecular Biology*, 30(2), 95-120.
- Kempfert, K. D., Jiang, E. Y., Oas, S., & Coffin, J. (2001). *Thermo Nicolet Application Note*. Retrieved from Thermo Fisher Scientific: <http://mmrc.caltech.edu/FTIR/Nicolet/Nicolet%20Tech%20Notes/DetectorsforFTIR1204.pdf>
- Kong, J., & Yu, S. (2007). Fourier transform infrared spectroscopic analysis of protein secondary structures. *Acta Biochimica et Biophysica Sinica*, 39(8), 549-559.

- Korkmaz, F. (2009). *Structure-Function Analysis of Membrane Proteins by Infrared Spectroscopy: Porin OmpF, Porin OmpG and Betaine Transporter BetP*. Frankfurt am Main.
- Korkmaz, F., Köster, S., Yıldız, Ö., & Mantele, W. (2012). In situ opening/closing of OmpG from E.coli and the splitting of beta-sheet signals in ATR-FTIR spectroscopy. *Spectrochimica Acta Part A*, 91, 395-401.
- Korkmaz, F., Pee, K. v., & Yildiz, Ö. (2015). IR-spectroscopic characterization of an elongated OmpG mutant. *Archives of Biochemistry and Biophysics*(576), 73-79.
- Korkmaz-Özkan, F., Köster, S., Kühlbrandt, W., Mantele, W., & Yildiz, Ö. (2010). Correlation between the OmpG secondary structure and its pH-dependent alterations monitored by FTIR. *Journal of molecular biology*, 401, 56-67.
- Kowolik, K. (2016, November 26). *Normal Modes*. Retrieved from ChemWiki: http://chemwiki.ucdavis.edu/Physical_Chemistry/Spectroscopy/Vibrational_Spectroscopy/Vibrational_Modes/Normal_Modes
- Lorentz-Fonfria, V. A., & Padros, E. (2004, January 18). Curve-fitting of Fourier manipulated spectra comprising apodization, smoothing, derivation and deconvolution. *Spectrochimica Acta Part A*(60), 2703-2710.
- Misra, R., & Benson, S. A. (1989). A novel mutation, cog, which results in the production of a new porin protein (OmpG) of Escherichia coli K-12. *Journal of Bacteriology*, 171(8), 4105-4111.
- Morris, J., Hartl, D., Knoll, A., & Lue, R. (2013). Translation and Protein Structure. In *Biology: How Life Works* (1st ed., pp. 4-6). New York: Macmillan.
- Möllman, K.-P., & Vollmer, M. (2013, October 22). Fourier transform infrared spectroscopy in physics laboratory courses. *European journal of physics*(34), 123-137.
- Nicolet FTIR User's Guide*. (2004). United States of America: Thermo Electron Corporation.

- PDB:2ODJ*. (2009, 05 17). Retrieved from Protein Data Bank Japan: https://pdj.org/eprts/index_en.cgi?PDB%3A2ODJ
- Perkin Elmer. (2005). FT-IR Spectroscopy ATR. USA. Retrieved June 5, 2017, from http://www.uts.utoronto.ca/~traceslab/ATR_FTIR.pdf
- Raschke, T. M., & Marqusee, S. (1998). Hydrogen exchange studies of protein structure. *Current Opinion in Biotechnology*(9), 80-86.
- Seddon, A. M., Curnow, P., & Booth, P. J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta* 1666, 105-117.
- Serway, R., & Jewett, J. J. (2010). *Physics for Scientists and Engineers* (8th ed.). Canada: Nelson Education.
- Silhavy, T. J., Daniel, K., & Walker, S. (2010, May). The Bacterial Cell Envelope. *Cold Spring Harbor Perspectives in Biology*, 2(5).
- Subbarao, G. V., & Berg, B. v. (2006). Crystal Structure of the Monomeric Porin OmpG. *Journal of Molecular Biology*(360), 750-759.
- Surewicz, W., Mantsch, H. H., & Chapman, D. (1993). Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment. *Biochemistry*, 32(2), 389-394.
- Venjaminov, S. Y., & Prendergast, F. G. (1997). Water (H₂O and D₂O) Molar Absorbivity in the 1000-4000 cm⁻¹ Range and Quantitative Infrared Spectroscopy of Aqueous Solutions. *Analytical Biochemistry*(248), 234-245.
- Yildiz, Ö., Vinothkumar, K. R., Goswami, P., & Kühlbrandt, W. (2006). Structure of the monomeric outer-membrane porin OmpG in the open and closed conformation. *The EMBO Journal*, 25(15), 3702-3713.

APPENDIX

CHEMICAL BONDING

1. Intramolecular Forces (Forces within a molecule)
1.1. Ionic Bond Formed by donation of electrons between atoms (i.e. Na-Cl)
1.2. Covalent Bond Formed by sharing pair of electrons 1.2.1. Polar Covalent Formed by atoms with different electronegativities (i.e. H-Cl) 1.2.2. Nonpolar Covalent Formed by atoms with similar electronegativities (i.e. Cl-Cl)
2. Intermolecular Forces (Forces between molecules)
2.1. Dipole-Dipole Interactions Interaction of two dipolar molecules (i.e. H-Cl-----H-Cl)
2.2. Ion-Dipole Interactions (i.e. Na ⁺ -----H ₂ O)
2.3. Hydrogen Bonding Electrostatic attraction is formed between a hydrogen atom (bound to a more electronegative atom) and another more electronegative atom in the close vicinity
2.4. Van der Waals Interactions Attraction and repulsion between instantaneous dipoles