UNIVERSITI TEKNOLOGI MARA

IDENTIFICATION STUDIES OF *Escherichia coli* USING FTIR PROFILES AND STRAIN TYPING BY PRINCIPAL COMPONENT ANALYSIS

AIMAN BINTI ALWI

Thesis submitted in fulfilment
of the requirements for the degree of
Master of Science
(Molecular Biology)

Faculty of Applied Science

January 2022

ABSTRACT

Escherichia coli (*E. coli*) is the most abundant facultative anaerobic bacteria present in the gastrointestinal tract and is normally harmless until gastrointestinal barriers are violated. Non-pathogenic *E. coli* strains can become pathogenic and cause diseases such as diarrhoea, dysentery, sepsis, and hemolytic-uremic syndrome (HUS). Conventional microbiology practice in determining the exact species of microorganisms include microscopic examination, colony morphology on selective and/or differential media, biochemical tests and serological typing. More precise identification requires application of methods based on molecular biology, which include pulse field gel electrophoresis (PFGE) or Polymerase Chain Reaction (PCR) are laborious, time and cost consuming and require technical expertise. Spectroscopy techniques has been successfully explored for the study of biological materials. Infra-red (IR) spectroscopy has become an accepted tool for the characterization of biomolecules and there is good potential of using IR for differentiating and identifying bacteria but specific tools to identify *E. coli* by using spectroscopy and chemometrics has not been addressed so far. Therefore this research aims to evaluate the potential of Fourier transform mid infrared (FT-MIR) and near infrared (FT-NIR) coupled with Principle Component Analysis (PCA) as chemometrics approach in differentiating pathogenic and non-pathogenic strains of *E. coli*. A selection of 11 species of bacteria from the family Enterobacteriaceae, including 4 strains of pathogenic *E. coli* and 3 strains of non-pathogenic *E. coli*, together with one strain of *Bacillus subtilis* as negative control was subjected to FT-MIR and FT-NIR spectroscopy using whole cells. The FT-MIR spectra analysis indicates that the region from 700 to 1810 cm^{-1} and 2660 to 3665 cm^{-1} were the best region to separate the different strains of *E. coli* due to high concentration of finger-print property of cell components. PCA analysis indicate that a maximum of 97.5% variance can be explained by the first two variable components. The FT-NIR spectra analysis showed the region from 7,200 to 4,000 cm^{-1} to be the best for strain separation and the first two variable components can account for 99.6% of all variations. Thus, FTIR spectroscopy has good potential as a tool in supporting the rapid diagnosis of pathogenic *E. coli* strain.

ACKNOWLEDGEMENT

Foremost, I would like to express my sincere gratitude to my supervisor Profesor Dr Mohd Faiz Foong Abdullah for the continuous support for my Master study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I couldn't have imagined having a better advisor and mentor for my Msc. study.

Beside my supervisor, I would like to thanks my co-supervisor, Dr. Mohd Zuli Jaafar from Universiti of Technology MARA Jasin for his encouragement and insightful comments and guidance in understanding and conduct his precious knowledge expertise in PCA analysis in MATLAB software.

In another part of this thesis, the given support by the collaborator in Photonics Laboratory, MIMOS, Technology Park Malaysia (TPM), Bukit Jalil, Kuala Lumpur, lecturers and friends at Universiti Teknologi MARA Shah Alam and Kuala Pilah. Besides that, special thanks to my research mates for the productive idea, discussions and endless support in my laboratory work.

Finally, I would like to thank my family, my husband, Muhamad Izzat Shafiq Rodi, my parents Puan Noriah Binti Hassan and En. Alwi Bin Muda for their love and support and also to my monetary support from KPT and UiTM.

TABLE OF CONTENTS