PROTEOMIC ANALYSES OF EXCRETORY SECRETORY PROTEINS (ESP) AND MEMBRANE PROTEINS OF *Entamoeba histolytica* HM1:IMSS

JORIM ANAK UJANG

UNIVERSITI SAINS MALAYSIA

2018
PROTEOMIC ANALYSES OF EXCRETORY SECRETORY PROTEINS (ESP) AND MEMBRANE PROTEINS OF *Entamoeba histolytica* HM1:IMSS

by

JORIM ANAK UJANG

Thesis submitted in fulfilment of the requirements for the degree of Master of Science

September 2018
ACKNOWLEDGEMENT

I would like to express my gratitude to Dr Nurulhasanah Othman, my supervisor, and Prof Rahmah Noordin, my co-supervisor. They have poured out their best in moulding me and guiding me in this journey to be a researcher. Not forgetting to mention, my personal development too. For that, I am forever grateful to them.

This research was supported/partially supported by Mr Muhammad Hafiznur Bin Yunus from the Insitute for Research in Molecular Medicine (INFORMM, USM), Dr. Mohd Nazri Ismail and Dr-to-be Sebastian Kwan from the Analytical Biochemistry Research Centre (ABrC, USM) as well as Pn Asmahani Azira Abdu Sani from the Malaysia Genome Institute (MGI, NIBM). Not to forget, all the staff of INFORMM, USM who have supported every facet of this project.

I would like to acknowledge all my friends and my family in INFORMM and at home, Kuching, who have supported me in my ups and downs, and who have tolerated my silliness. Without your moral support, I would not have continued this journey. Especially to mommy and daddy, thank you for your tremendous support in your love and encouragement; and money of course.

Last but not least, a big shout out to God almighty! Thank You for the favour and thank You for all the experiences I gained in this journey.
# TABLE OF CONTENTS

AKNOWLEDGEMENT ...................................................................................... ii  
TABLE OF CONTENTS ............................................................................ iii  
LIST OF TABLES ............................................................................................ ix  
LIST OF FIGURES .......................................................................................... x  
LIST OF SYMBOLS ....................................................................................... xii  
LIST OF ABBREVIATIONS .......................................................................... xiii  
ABSTRAK ................................................................................................. xv  
ABSTRACT ................................................................................................. xvii  

CHAPTER 1 – INTRODUCTION .................................................................... 1  
1.1 An overview .......................................................................................... 1  
1.2 Problem statements and rationale of the study ..................................... 3  
1.3 Objectives of the study ......................................................................... 4  
1.4 Workflow ............................................................................................. 5  

CHAPTER 2 – LITERATURE REVIEW ....................................................... 6  
2.1 Biology of *E. histolytica* ..................................................................... 6  
   2.1.1 Life cycle ....................................................................................... 6  
   2.1.2 Cell morphology .......................................................................... 8  
   2.1.3 Transmission and occurrence ...................................................... 9  
      2.1.3(a) Susceptibility and risk factors ................................................. 9  
      2.1.3(b) Epidemiology ....................................................................... 10  
   2.1.4 Disease, diagnosis and treatment ................................................. 11  
      2.1.4(a) Symptoms ............................................................................ 11  
      2.1.4(b) Diagnosis ............................................................................. 13  
      2.1.4(c) Treatments ............................................................................ 14  
   2.1.5 Pathogenesis .................................................................................. 15
2.2 The proteome of *E. histolytica* ................................................................. 19
2.2.1 Excretory-secretory proteins ................................................................. 21
2.2.2 Membrane proteins .............................................................................. 22
2.3 Tools for proteome discovery ................................................................... 24
2.3.1 Sample preparation .................................................................................. 24
2.3.2 Fractionation and peptide separation .................................................... 27
2.3.3 Protein identification by mass spectrometry ............................................ 28
2.3.3(a) Matrix-assisted laser desorption ionisation and time-of-flight analyser (MALDI-TOF) ................................................................. 29
2.3.3(b) Electrospray ionisation (ESI) ............................................................... 30
2.3.3(c) The advantages and disadvantages of MALDI and ESI .................... 31
2.3.3(d) Mass spectrometry - Peptide fragmentation ...................................... 32
2.3.3(e) Database search .................................................................................. 34
2.3.4 Protein topology prediction ..................................................................... 34
2.3.4(a) Classically and non-classically secreted proteins ......................... 34
2.3.4(b) Transmembrane topology ................................................................. 35

CHAPTER 3 – MATERIALS AND METHODS ...................................................... 38

3.1 Materials ...................................................................................................... 38
3.1.1 Preparation of Culture medium, Buffers and Solutions ......................... 38
3.1.1(a) Trypticase-Yeast Extract-Iron and Serum Medium, TYI-S-33 .................. 38
3.1.1(b) Heat-inactivated bovine serum .......................................................... 38
3.1.1(c) Potassium phosphate monobasic solution, 0.15 M ............................ 39
3.1.1(d) Potassium phosphate dibasic solution, 0.15 M ................................. 39
3.1.1(e) Phosphate buffer, 0.15 M ................................................................. 39
3.1.1(f) Phosphate Buffer Saline for Amoeba, PBS(A) ...................................... 39
3.1.2 Materials for isolation of excretory-secretory proteins ......................... 40
3.1.2(a) RPMI-C-A ....................................................................................... 40
3.1.2(b) Ammonium bicarbonate, 50 mM ........................................ 40
3.1.2(c) Protease inhibitor (Roche), 7× ..................................... 40
3.1.3 Solutions for the conventional membrane protein extraction ....... 40
  3.1.3(a) Sodium phosphate monobasic, 10 mM .............................. 40
  3.1.3(b) Sodium phosphate dibasic, 10 mM .................................. 41
  3.1.3(c) 10 mM Sodium phosphate buffer, pH 8.0 .......................... 41
3.1.4 Solutions for ProteoPrep® Membrane Extraction Kit (Sigma, USA) ........ 41
  3.1.4(a) Soluble Cytoplasmic and Loosely-bound Membrane Protein Extraction Reagent ........................................ 41
  3.1.4(b) Protein Extraction Reagent Type 4 .................................. 41
3.1.5 Solutions for ProteoExtract® Native Membrane Protein Extraction Kit .......................................................... 41
3.1.6 Materials and buffers for SDS-PAGE ..................................... 42
  3.1.6(a) 10% SDS ........................................................................ 42
  3.1.6(b) Resolving buffer ............................................................ 42
  3.1.6(c) Stacking buffer ............................................................... 42
  3.1.6(d) Ammonium persulfate (APS), 10% .................................. 42
  3.1.6(e) Loading buffer, 5× .......................................................... 42
  3.1.6(f) SDS-PAGE running buffer ............................................... 43
  3.1.6(g) RAMA stain ................................................................. 43
     3.1.6(g)(i) Coomassie Brilliant Blue, CBB R250, 0.05% ............. 43
     3.1.6(g)(ii) Ammonium sulfate, 30% .......................................... 43
     3.1.6(g)(iii) RAMA stain ......................................................... 43
  3.1.6(h) Separating gel and stacking gel ....................................... 44
3.1.7 Materials for sample digestion ............................................. 45
  3.1.7(a) RapiGest ............................................................... 45
  3.1.7(b) 100 mM Dithiothreitol (DTT)........................................ 45
  3.1.7(c) 200 mM Iodoacetamide (IAA) ......................................... 45
3.1.7(d) 1 µg/µL Trypsin ........................................................... 45

3.1.8 Materials for nanoLC-MALDI-TOF/TOF ...................................... 45
3.1.8(a) Buffer A (2% ACN, 0.1% TFA, 97.9% ddH2O) ......................... 45
3.1.8(b) Buffer B (98% ACN, 0.1% TFA, 1.9% ddH2O) ....................... 46
3.1.8(c) Matrix diluent (70% ACN, 0.1% TFA, 29.9% ddH2O) ............... 46
3.1.8(d) Matrix α-cyanohydroxycinnamic acid (5 mg/mL) ................. 46
3.1.8(e) Calibration mix .................................................................. 46

3.1.9 Buffers for LC – ESI (LTQ-Orbitrap Velos Pro system and Dionex UltiMate 3000 system) ......................................................... 47

3.2 Methods .................................................................................. 47
3.2.1 Axenic culture of *E. histolytica* trophozoites ............................... 47
3.2.1(a) Expansion of the culture ...................................................... 48
3.2.1(b) Mass Culture .................................................................... 48
3.2.2 Protein isolation and extraction ................................................ 48
3.2.2(a) Isolation of excretory-secretory proteins .............................. 48
3.2.2(b) Extraction of membrane proteins ....................................... 49
   3.2.2(b)(i) ProteoExtract® Native Membrane Protein Extraction Kit (Calbiochem, Germany) ............... 49
   3.2.2(b)(ii) ProteoPrep® Membrane Extraction Kit (Sigma, USA) ......................................................... 50
   3.2.2(b)(iii) Conventional Membrane Protein Extraction Method (Texeira) ........................................ 51
3.2.3 Protein concentration and buffer exchange ................................ 51
   3.2.3(a) Protein concentration ...................................................... 51
   3.2.3(b) Acetone precipitation ..................................................... 52
3.2.4 Gel electrophoresis .................................................................. 55
3.2.5 Protein digestion ..................................................................... 55
3.2.6 Mass spectrometry analysis ...................................................... 56
3.2.6(a) LC–MALDI–TOF/TOF .............................................. 56
3.2.6(b) LC–ESI–MS/MS (LTQ-Orbitrap Velos Pro) .................. 57
3.2.6(c) LC–ESI–MS/MS (Orbitrap Fusion) ........................... 58
3.2.7 Data analysis ...................................................................... 59
3.2.7(a) Prediction of excretory-secretory proteins ................. 59
3.2.7(b) Prediction of membrane protein ............................... 60
3.2.7(c) Functional annotation .............................................. 60

CHAPTER 4 – RESULTS .............................................................................. 61
4.1 ES proteins .............................................................................. 61
  4.1.1 Protein profile of ES proteins ......................................... 61
  4.1.2 ES proteins identification by LC–MALDI–TOF/TOF and LC–ESI–MS/MS .... 61
  4.1.3 Prediction of classical and non-classical secretion .............. 66
  4.1.4 Functional annotation of ES proteins ............................. 66
4.2 Membrane proteins ................................................................. 69
  4.2.1 Comparison of protein yield between ProteoExtract® kit, ProteoPrep® kit and conventional method .................. 69
  4.2.2 SDS-PAGE protein profiles of the three extraction methods .......... 69
  4.2.3 E. histolytica protein identification of the membrane fractions extracted using the ProteoExtract® kit, ProteoPrep® kit and conventional method ......................................................... 72
  4.2.4 Evaluation of the membrane extraction methods selectivity by comparing the identified membrane fraction and cytosolic fraction proteins ......................................................... 85
  4.2.5 Evaluation of the membrane extraction methods specificity by comparing the predicted membrane proteins among the identified membrane and cytosolic fraction proteins ......................... 89
  4.2.6 Functional annotation of membrane proteins .................. 92

CHAPTER 5 – DISCUSSION ................................................................. 95
5.1 The excretory-secretory (ES) proteins of E. histolytica .............. 95
  5.1.1 Isolation of the ES proteins ......................................................... 95
5.1.2 ES proteins identification and prediction of the classical and non-classical secretion pathways ................................................................. 97

5.1.3 Functional protein classification ................................................................. 99

5.2 The membrane proteins of *E. histolytica* .................................................................. 102

5.2.1 Comparison of membrane protein extraction between ProteoExtract® kit, ProteoPrep® kit and conventional method ........................................... 102

5.2.2 Assessment of method’s selectivity and specificity................................. 103

5.2.3 Functional annotation ..................................................................................... 106

5.3 The complementary use of LC-MALDI-TOF/TOF and LC-ESI-MS/MS system .......................................................................................... 109

CHAPTER 6 – SUMMARY AND CONCLUSION .............................................. 112

REFERENCES ............................................................................................................. 115

APPENDICES
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1</td>
<td>Preparation of SDS-PAGE gel for one small gel - 1.0 mm Mini-PROTEAN</td>
<td>44</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Examples of <em>E. histolytica</em> ES proteins identified by LC-ESI-MS/MS</td>
<td>63</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Examples of <em>E. histolytica</em> ES proteins identified by LC-MALDI-TOF/TOF</td>
<td>64</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Examples of proteins extracted using the ProteoExtract® kit and identified by LC-MALDI-TOF/TOF</td>
<td>74</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>Examples of proteins extracted using the ProteoExtract® kit and identified by LC-ESI-MS/MS</td>
<td>75</td>
</tr>
<tr>
<td>Table 4.5</td>
<td>Examples of proteins extracted using ProteoPrep® kit and identified by LC-MALDI-TOF/TOF</td>
<td>76</td>
</tr>
<tr>
<td>Table 4.6</td>
<td>Examples of proteins extracted using ProteoPrep® kit and identified by LC-ESI-MS/MS</td>
<td>77</td>
</tr>
<tr>
<td>Table 4.7</td>
<td>Examples of proteins extracted using the conventional method and identified by LC-MALDI-TOF/TOF</td>
<td>78</td>
</tr>
<tr>
<td>Table 4.8</td>
<td>Examples of proteins extracted using the conventional method and identified by LC-ESI-MS/MS</td>
<td>79</td>
</tr>
<tr>
<td>Table 4.9</td>
<td>List of all the predicted membrane proteins exclusively identified by LC-MALDI-TOF/TOF</td>
<td>84</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Comparison of the protein yield, time, cost and the need of ultracentrifugation of each method</td>
<td>103</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Flowchart of the study.</td>
<td>5</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>The life cycle of <em>Entamoeba histolytica</em>.</td>
<td>7</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>A compilation of studies by Norhayati et al. (2003) on the prevalence of amoebiasis among the orang asli communities in West Malaysia.</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Pathogenesis of intestinal amoebiasis (Lejeune et al., 2009)</td>
<td>18</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>A bottom-up proteomics workflow.</td>
<td>26</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>A standard curve constructed from a serial dilution of known BSA concentrations.</td>
<td>54</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Protein profiles of <em>E. histolytica</em> ES proteins from three biological replicates.</td>
<td>62</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Venn diagram represents the <em>E. histolytica</em> ES proteins identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS.</td>
<td>65</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Venn diagram represents the prediction analysis of classically and non-classically secreted proteins between the <em>E. histolytica</em> protein database and the ES proteins identified by both mass spectrometry systems.</td>
<td>67</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>The protein classes of <em>E. histolytica</em> excretory-secretory proteins.</td>
<td>68</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Protein yields from the three extraction methods.</td>
<td>70</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>The protein profiles of <em>E. histolytica</em> membrane fractions extracted using the ProteoExtract® kit, conventional method and ProteoPrep® kit.</td>
<td>71</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>The Venn diagrams represent the number of proteins identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS from the membrane fractions.</td>
<td>80</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>The number of predicted membrane proteins versus non-membrane proteins identified from the membrane fractions using the three extraction methods.</td>
<td>81</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>A combination of the identified <em>E. histolytica</em> membrane fraction proteins extracted by the three extraction methods.</td>
<td>82</td>
</tr>
<tr>
<td>Figure 4.10</td>
<td>The Venn diagram represents the predicted membrane proteins identified in at least two replicates from the membrane fractions of the three extraction methods.</td>
<td>83</td>
</tr>
</tbody>
</table>
Figure 4.11  The number of proteins identified from the membrane fractions and the cytosolic fractions of the ProteoExtract® kit, ProteoPrep® kit and conventional method ............................................... 87

Figure 4.12  The Venn diagrams compare the proteins identified using LC-ESI-MS/MS from the membrane and cytosolic fractions. .................. 88

Figure 4.13  The number of predicted membrane proteins versus non-membrane proteins identified from the membrane fraction (MF) and cytosolic fraction (CF) of the three extraction methods. ............... 90

Figure 4.14  The Venn diagrams compare the predicted membrane proteins from the membrane and cytosolic fractions. ................................. 91

Figure 4.15  The proteins classes of E. histolytica membrane fraction proteins. .................................................................................................. 93

Figure 4.16  The E. histolytica membrane fraction proteins that are classified according to the cellular compartment. ........................................ 94
## LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>degree Celsius</td>
</tr>
<tr>
<td>µL</td>
<td>microlitre</td>
</tr>
<tr>
<td>µg</td>
<td>microgram</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>g</td>
<td>earth’s gravitational acceleration/relative centrifugal force</td>
</tr>
<tr>
<td>L</td>
<td>litre</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>mL</td>
<td>millilitre</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
</tbody>
</table>
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>1-dimensional</td>
</tr>
<tr>
<td>2DE</td>
<td>2-dimensional electrophoresis</td>
</tr>
<tr>
<td>ACN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>ALA</td>
<td>Amoebic liver abscess</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium persulfate</td>
</tr>
<tr>
<td>CHCA</td>
<td>α-Cyano-4-hydroxycinnamic acid</td>
</tr>
<tr>
<td>CID</td>
<td>Collision induced dissociation</td>
</tr>
<tr>
<td>CLS</td>
<td>Cyst-like structure</td>
</tr>
<tr>
<td>CSA</td>
<td>Crude-soluble antigen</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>ECD</td>
<td>Electron-capture dissociation</td>
</tr>
<tr>
<td>ES</td>
<td>Excretory-secretory</td>
</tr>
<tr>
<td>ESA</td>
<td>Excretory-secretory antigen</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionisation</td>
</tr>
<tr>
<td>ETD</td>
<td>Electron-transfer dissociation</td>
</tr>
<tr>
<td>FA</td>
<td>Formic acid</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>IAA</td>
<td>Iodoacetamide</td>
</tr>
<tr>
<td>LC</td>
<td>Liquid chromatography</td>
</tr>
<tr>
<td>m/z</td>
<td>mass-to-charge ratio</td>
</tr>
<tr>
<td>MALDI</td>
<td>Matrix-assisted laser desorption ionisation</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>MS/MS</td>
<td>Tandem mass spectrometry</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>NaOH</td>
<td>Sodium hydroxide</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffer saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>pI</td>
<td>Isoelectric point</td>
</tr>
<tr>
<td>PMF</td>
<td>Peptide mass fingerprinting</td>
</tr>
<tr>
<td>PPDK</td>
<td>Pyruvate phosphate dikinase</td>
</tr>
<tr>
<td>RCDC</td>
<td>Reducing agent and detergent compatible</td>
</tr>
<tr>
<td>RT</td>
<td>Real time</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetrakis(methylene) diamine</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TOF</td>
<td>Time-of-flight</td>
</tr>
<tr>
<td>TOF/TOF</td>
<td>Tandem time-of-flight</td>
</tr>
</tbody>
</table>
ANALISIS PROTEOMIK PROTEIN PERKUMUHAN PEREMBES (ESP)
DAN PROTEIN MEMBRAN DARIPADA Entamoeba histolytica HM1:IMSS

ABSTRAK

PROTEOMIC ANALYSES OF EXCRETORY SECRETORY PROTEINS (ESP) AND MEMBRANE PROTEINS OF *Entamoeba histolytica* HM1:IMSS

ABSTRACT

*Entamoeba histolytica* is a protozoan parasite that causes amoebiasis. Infection of this parasite may lead to amoebic dysentery and amoebic liver abscess, which is fatal if left untreated. Until now, understanding of the pathogenesis of amoebiasis is limited. Hence, in this study, proteomic analyses were performed on the excretory-secretory (ES) and the membrane sub-proteomes of *E. histolytica* trophozoites. A total of 209 ES proteins were identified in which 191 and 97 proteins were detected by LC–ESI–MS/MS and LC–MALDI–TOF/TOF, respectively. Of the 209 proteins, 79 were identified by both mass-spectrometry systems, while 112 and 18 proteins were detected exclusively by LC–ESI–MS/MS and LC–MALDI–TOF/TOF respectively. Subsequently, the secretome prediction analyses were performed whereby 8 and 31 out of 209 total proteins were identified as classically and non-classically secreted proteins, respectively. Functional annotation classification showed that the largest ES protein class, which is 23%, is the oxidoreductase. The second part of this study involved the comparison of three membrane protein extraction methods: two commercial kits (ProteoExtract® from Calbiochem and ProteoPrep® from Sigma), and a conventional laboratory method. The results showed that the ProteoExtract® kit and the conventional method extracted higher protein yields compared to the ProteoPrep® kit. The combined data from LC-MALDI-TOF/TOF and LC-ESI-MS/MS identified 490, 492, and 587 proteins extracted using the ProteoExtract®, ProteoPrep®, and conventional methods,
respectively. *In-silico* analysis predicted 109 (22%), 237 (48%) and 182 (31%) membrane proteins from the ProteoExtract®, ProteoPrep® and conventional method extracts, respectively. Furthermore, the identification of the cytosolic and membrane protein fractions showed that the ProteoPrep® extraction kit was the most selective and specific for the extraction of the membrane proteins. In conclusion, the results revealed 39 and 249 *E. histolytica* ES and membrane proteins, respectively. Furthermore, this study confirmed that the use of two types of mass spectrometers enhances proteome coverage. The data generated has increased the understanding on the types of proteins that are excreted-secreted by *E. histolytica* and also the proteins that reside at the parasite’s membrane. The identified proteins will be useful for further studies in understanding the pathogenesis of amoebiasis and the roles the proteins play in the host-parasite interactions.
CHAPTER 1 – INTRODUCTION

1.1 An overview

Amoebiasis was first reported as a deadly disease in 1873 by Hippocrates who examined a patient suffering from bloody dysentery (Tanyuksel and Petri, 2003). Two years later, *Entamoeba histolytica* trophozoite was identified by Fedor Aleksandrovich Lösch in a farmer who suffered from a fatal case of dysentery (Marshall et al., 1997). Further investigation by inoculating the stool of the patient into the rectum of a dog caused a similar manifestation (Marshall et al., 1997). A significant milestone was achieved with the characterisation of *E. histolytica* as the causative agent for amoebic colitis and amoebic liver abscess (ALA) in the 1890s by Sir William Olser and his colleagues (Tanyuksel and Petri, 2003). Subsequently, the identification of cyst as an infectious stage was confirmed by Walker and Sellards in 1913, and followed by the establishment of the *E. histolytica* life cycle by Dobell in 1925 (Tanyuksel and Petri, 2003).

Most patients infected with *E. histolytica* are asymptomatic or only suffered from mild diarrhoea (Hankenson et al., 2003). Meanwhile, only 10% of the patients presented classic amoebic symptoms such as stomach cramps and bloody diarrhoea (Farthing, 2006). *E. histolytica* was not immediately associated as the causative agent of amoebiasis because most amoebic infections cases were asymptomatic. However, subsequent studies found that the infectious and the non-infectious amoeba were not similar (Fotedar et al., 2007). Since then, *E. histolytica* was reclassified into two species namely the infectious species, *E. histolytica* and the non-infectious species, *E. dispar* (Fotedar et al., 2007).
In 1997, amoebiasis was ranked second as death-causing parasitic infection, after malaria (World Health Organization, 1997). Approximately 40,000 to 100,000 deaths occurred annually, which include 1.9% to 9% of amoebic colitis patients (Aristizábal et al., 1991). Death occurrence in amoebic liver abscess (ALA) cases have decreased to 1 – 3% due to the effective medical intervention. Nonetheless, the mortality rate caused by the late detection resulting in the sudden intraperitoneal rupture occurred in 2 – 7% of the patients (Stanley Jr, 2003).

The results of previous studies have contributed to the advancement on many aspects in the management of amoebiasis. This includes a better way of diagnosis whereby the detection of pathogenic *E. histolytica* could be accurately distinguished from the morphologically similar but non-pathogenic *E. dispar* (Fotedar et al., 2007). Although many attempts have been made to improve the management of amoebiasis, the disease remains prevalent in underdeveloped countries of warmer climate (Walsh and Ravdin, 1988). Furthermore, the combination of poor sanitation and bad water quality provides the optimum breeding ground for this parasite (Walsh and Ravdin, 1988).

Large amounts of information on *E. histolytica* genome were made available since it was sequenced in the year 2005 (Loftus et al., 2005). In tandem, the advent of proteomic technologies has allowed proteomic studies on amoebiasis to be conducted. Early studies focused on analysing the subcellular expression profiles of trophozoites under various conditions (Davis et al., 2006, Tolstrup et al., 2007, Perdomo et al., 2015). Tolstrup et al. (2007) used 2-dimensional gel electrophoresis (2-DE) followed by mass spectrometry (MS) on 400–1500 *E. histolytica* protein spots. Therein, 63 proteins were identified and found related to cytoskeleton, surface, metabolic, the ubiquitin-proteasome system and signalling associated proteins. The
application of differential protein expression analysis using 2-DE and subsequent MS analysis was also performed to compare the proteome of Rahman and HM1:IMSS strains of *E. histolytica*. The results showed six proteins were found differentially expressed between the two strains (Davis et al., 2006). In 2015, a study on the *E. histolytica* trophozoite ER and Golgi apparatus using LC-MS/MS identified over 1,500 proteins of which are involved as trafficking machinery and GTPases (Perdomo et al., 2015). Hence, with the advancement of proteomic technologies and a complete *E. histolytica* protein database, high-throughput studies on the proteome of *E. histolytica* can be conducted.

### 1.2 Problem statements and rationale of the study

The shift of analysing a single protein to larger sets of proteins such as the excretory-secretory (ES) proteins and membrane proteins is made possible with advances in the proteomics technologies. Currently, there is no proteome report on the ES proteins of *E. histolytica*. Other than the study of *E. histolytica* cell surface membrane proteins by Biller et al. (2014), no other study has been performed on the membrane proteome of *E. histolytica*. Thus, this study aimed to identify the ES and membrane proteomes of *E. histolytica*. The identification of these proteins and their functions could add to the knowledge in understanding amoebic pathogenesis.
1.3 Objectives of the study

This study was conducted with the following objectives:

1. To perform proteomic analysis using LC-MALDI-TOF/TOF and LC-ESI-MS/MS and functional classification via PANTHERDB for the identified *E. histolytica* excretory-secretory proteins.

2. To compare *E. histolytica* membrane protein extraction methods: two commercial kits namely ProteoExtract® (Calbiochem), ProteoPrep® (Sigma), and a conventional laboratory method.

3. To perform proteomic analysis using LC-MALDI-TOF/TOF and LC-ESI-MS/MS and functional classification via PANTHERDB for the identified *E. histolytica* membrane proteins.
1.4 Workflow

Figure 1.1   Flowchart of the study.
CHAPTER 2 – LITERATURE REVIEW

2.1 Biology of *E. histolytica*

2.1.1 Life cycle

*E. histolytica* exists in two distinct stages namely trophozoites and cysts. The simple life cycle begins with the consumption of the tainted fluid containing *E. histolytica* cyst (Hankenson et al., 2003). The cyst withstands harsh environment such as the gastric acid. Upon reaching a conducive environment such as the small intestine, a single cyst ex-cysts to form 8 trophozoites. These blood ingesting trophozoites then colonise the colon and cause dysentery. Trophozoites are unable to live in an unconducive environment outside the host or the host’s gastric acids unless quadrinucleate cysts are formed again through a process known as encystation. Humans and primates are the only natural hosts for *E. histolytica* (Rivera et al., 2010, Stanley Jr, 2003). Figure 2.1 depicts the life cycle of *E. histolytica*. 
Figure 2.1  The life cycle of *Entamoeba histolytica*. 

Image source: 
2.1.2 Cell morphology

*E. histolytica* is a unicellular eukaryotic organism. The cyst form of *E. histolytica* is round in shape and is enclosed within a refractile wall which protects them from harsh conditions such as the stomach acid. It is responsible for the transmission of the disease. A mature cyst is 10-15 μm in size and consists of four nuclei. From a single quadrinucleate cyst, eight uninuclear trophozoites are formed through a process known as excystation. The trophozoite form of *E. histolytica* is 10-50 μm in size and consists of one nucleus. It is actively motile with finger-shaped pseudopodia and responsible for tissue invasion and damage. It is also responsible for causing tissue damage to the host (Stanley Jr, 2003).

In the trophozoite form, it contains a single nucleus and multiplies by binary fission. It is an endoparasite whereby it ingests nutrients from the host and can alter its shape for various purposes such as locomotion and evasion of the host immune responses (Espinosa-Cantellano et al., 1992, Markiewicz et al., 2011).

Other amoeba species such as *E. dispar* and *E. moshkovskii* share the same physical features with *E. histolytica*, thus causing difficulty in differentiating them from *E. histolytica* under the microscope (World Health Organization, 1997, Fotedar et al., 2007). The need to distinguish *E. histolytica* from other non-pathogenic *Entamoeba* species is important to avoid misdiagnosis and wrong treatment.
2.1.3 Transmission and occurrence

2.1.3(a) Susceptibility and risk factors

Typically, amoebiasis is acquired through the faecal-oral route, whereby food or water contaminated by the cyst form of *E. histolytica* is ingested by the host. Transmission can also occur through oral and anal sex as well as contaminated enema apparatus (Istre et al., 1982). According to Hankenson et al. (2003), the communicability of the disease is high as asymptomatic carriers can be a source of further infection. Furthermore, common household pest like flies and cockroaches can help spread the cyst form of *E. histolytica*. Adults and infants have similar chances of acquiring amoebiasis. However, according to a report by Hung, Chang & Ji (2012), men who have sex with men have a higher risk of being infected with *E. histolytica*.

Amoebiasis is still a major health problem especially among the aboriginals and communities living in the remote areas of Malaysia (Tengku and Norhayati, 2011). A study on the prevalence of *E. dispar/E. histolytica* among school children in the interior of Sabah showed that 83.8% of them had the infection (Mahsol et al., 2008). One of the leading factors causing high intestinal parasitic infections including *E. histolytica* in remote communities involves water sources (Duc et al., 2011). For instance, the transmission of parasites occurs in settings where a river contaminated with human and animal excretion is used interchangeably for agriculture, socio-economic and personal hygiene (Duc et al., 2011).
2.1.3(b) Epidemiology

Amoebiasis commonly occurs in populations living in tropical areas that lack proper sanitation. The disease is prevalent in developing countries such as Mexico, India, Africa and Malaysia (Centers for Disease Control and Prevention, 2015). In western Nepal, amoebic infection was ranked second after giardiasis (Mukhopadhyay et al., 2007). The prevalence of *E. histolytica* infection in the different regions of Brazil from the year 2001 to 2014 ranged between 6.8% and 46.3% (Silva et al., 2014). In Pakistan, the prevalence of *E. histolytica* was reported to be as high as 23.1%, whereby the most susceptible age group was found between 6 to 10 years old (Zeb et al., 2018).

In Malaysia, the prevalence of amoebiasis among the Orang Asli was found between 1% and 14% (Norhayati et al., 2003) (Figure 2.2). An outbreak of amoebiasis among the orang asli communities in the year 2004 reported 13.2% of 28 diarrhoea patients were infected by *E. histolytica* (Noor Azian et al., 2006). In 2005, 72.4% out of 58 liver abscess patients admitted to Hospital Universiti Sains Malaysia (HUSM) were found to be caused by *E. histolytica* (Zeehaida et al., 2008). Furthermore, between 2008 and 2009, 76.7% out of 30 liver abscess cases admitted to HUSM were also positive for *E. histolytica* DNA (Othman et al., 2010).

Despite being prevalent in tropical countries, human to human transmission can still occur regardless of climate and high sanitation standards. For example, in a temperate country such as Japan, mass *E. histolytica* infection at an institution for the mentally disabled in the Yamagata Prefecture of Japan reported 5 to 10% of people infected were symptomatic, while 90 to 95% of infected subjects were asymptomatic (Haghigi et al., 2003). In a report by Vreden et al. (2000), an amoebiasis outbreak in
the Netherlands demonstrated that *E. histolytica* can remain dormant for 13 years in their climate.

### 2.1.4 Disease, diagnosis and treatment

#### 2.1.4(a) Symptoms

In most *E. histolytica* infections, symptoms are either not present or very mild (Stanley Jr, 2003). The majority of asymptomatic patients excrete cysts for a short period and are clear from the infection within 12 months of infection (van Hal et al., 2007). Only a small percentage of people infected with *E. histolytica* develops clinical symptoms. Patients with symptomatic amoebiasis often suffer from amoebic colitis and amoebic liver abscess (ALA) (Stanley Jr, 2003).

Patients with amoebic colitis commonly present a history of persistent abdominal pain and diarrhoea with the presence of blood and mucus in the stool. As amoebiasis is often neglected, a study reports that common inappropriate symptomatic treatment using corticosteroid has led to toxic megacolon complication in about 0.5% of patients (Ackers et al., 1997). Furthermore, when left untreated, the resulting gut perforation, exhaustion, and extraintestinal amoebiasis will lead to death (Hankenson et al., 2003).

ALA is the most common extraintestinal manifestation of amoebiasis (van Hal et al., 2007). As mentioned by Zurauskas & McBride (2001), patients who develop ALA are usually presented within 5 months of exposure to the disease, with clinical symptoms such as fever, and right upper abdominal quadrant pain. The majority of ALA patients do not present amoebic colitis symptoms and also *E. histolytica* cysts and trophozoites are rarely found in their stools (Fotedar et al., 2007).
Figure 2.2  A compilation of studies by Norhayati et al. (2003) on the prevalence of amoebiasis among the orang asli communities in West Malaysia.
2.1.4(b) Diagnosis

The earliest diagnosis method of amoebiasis is the microscopic examination of stool samples whereby *E. histolytica* trophozoites can be seen containing red blood cells. However, this method is prone to cause misdiagnosis as other morphological similar strains, such as *E. dispar* and *E. moshkovskii* are indistinguishable from *E. histolytica* under the microscope (Liang et al., 2009, Haque and Petri, 2006). Although microscopic method is routinely being used to diagnose amoebic colitis, it is not suitable to be performed for diagnosis of ALA cases. Despite low sensitivity of the microscopy, it is still being practised in many hospital laboratories.

Amoebic colitis patients can also be diagnosed by detecting small ulcers on colonic lesions obtained during the colonoscopic biopsy (Ohnishi et al., 2004). Furthermore, colonoscopy and subsequent sampling by means of culture swap are useful in patients with acute colitis and in cases when *E. histolytica* infection is suspected but failed to be detected in stool samples. However, these methods are time-consuming and sensitivity of the diagnosis is only 50% (Clark and Diamond, 2002). Antigen detection methods, eg. Entamoeba CELISA Path kit (Cellabs, Sydney, NSW) and the *E. histolytica* II kit (TechLab Inc, Blacksburg, Va, USA), are specific and can distinguish *E. histolytica* from *E. dispar*. The sensitivities and specificities of these various antigen detection kit ranges from 80% to 99% and from 86% to 98%, respectively (Haque et al., 1995, Gonin and Trudel, 2003, Furrows et al., 2004, Solaymani-Mohammadi et al., 2006). These tests are rapid and their interpretations are more definitive compared to the microscopic examination.

For the diagnosis of extraintestinal amoebiasis such as ALA, radiology imaging is used to detect the presence of an abscess in the liver. When the abscess is
present, further analyses such as culture, DNA detection, and/or antigen detection are performed. DNA and antigen detection-based methods performed on the abscess sample were reported to be highly sensitive (Fotedar et al., 2007, Paul et al., 2007, Tanyuksel and Petri, 2003). With serological methods, serum samples were used to detect antibodies against *E. histolytica* for the diagnosis of ALA. Commercial antibody detection assays made of native *E. histolytica* trophozoite antigens are available (Lotter et al., 1992, Ning et al., 2013). However, this method is ineffective to distinguish recent infection from past infection as high background antibody titre may persist in a population of endemic areas (Pillai et al., 1999, Zengzhu et al., 1999, Zeehaida et al., 2008, Mohamed et al., 2009).

Molecular diagnostic tests using polymerase chain reaction (PCR) to amplify *E. histolytica* DNA from the extracted faecal and pus of ALA patients are shown to be highly sensitive and specific (Gonin and Trudel, 2003, Solaymani-Mohammadi et al., 2006). Furthermore, the application of real-time PCR (RT-PCR) has significantly shortened detection time by simultaneous monitoring of the amplification process (Othman et al., 2010). The advantages of RT-PCR are the ability to detect a low number of parasite and the reliability in differentiating non-pathogenic *Entamoeba* species from *E. histolytica* (van Hal et al., 2007). However, these methods require skilled personnel and the high cost of reagents and equipment.

2.1.4(c) Treatments

Treatment for amoebiasis includes the oral administration of metronidazole or diiodohydroxyquin and in conjunction with a luminal agent such as iodoquino (Hankenson et al., 2003). For patients with invasive amoebiasis, surgical drainage
may be unnecessary to treat ALA, as drug therapy alone is efficient (Akgun et al., 1999). However, aspiration of the abscess was shown to be beneficial in patients with large abscesses (Weinke et al., 2002). Meanwhile, asymptomatic carriers should be treated with a luminal agent to reduce the spread of disease and the risk of developing symptomatic infection (Stanley Jr, 2003).

Current drug therapies have been shown to cause several side effects. According to Petri Jr & Singh (1999), effective luminal agents such as diloxanide furoate and paromomycin caused frequent gastrointestinal disturbances and rare double vision, and symptoms related to ototoxicity and nephrotoxicity. Furthermore, other drugs used to treat amoebiasis such as metronidazole may cause unpleasant side effects, such as metallic taste, nausea and headache. Although it is uncommon, metronidazole can also cause neurological side effects, such as vertigo and encephalitis. Hence, treatments are discontinued whenever harmful side effects are present.

2.1.5 Pathogenesis

According to Lejeune, Rybicka, & Chadee (2009), *E. histolytica* trophozoites can maintain a commensal relationship with the host. This is generally observed in the infected individuals. When triggered, the unharmful relationship can turn destructive, beginning from the destruction of the intestinal wall, to the extent of invading surrounding soft organs such as the liver. Several works have been performed to distinguish the virulent and attenuated strains as well as to uncover the stress-inducing components from the tissue environment (Faust and Guillen, 2012). However, the conversion of the parasite from a commensal state to a destructive state
needs to be further elucidated. The current knowledge of the pathogenesis of intestinal amoebiasis is shown in Figure 2.3.

This host tissue lysing parasite has phagocytic, proteolytic, and cytolytic capabilities. Gal-lectin, cysteine proteinase and amoebapore are the three proteins known as the main culprit in the pathogenesis of amoebiasis. The invasion of the intestinal mucosa by the degradation of the mucin layer is thought to be the first strategy to disrupt the mucus gel (Moncada et al., 2005). In this strategy, the cysteine proteases secreted by \textit{E. histolytica} effectively degraded the cysteine-rich domains of the MUC2 polymer of the mucus gel. Hence, it permits the parasite to come into contact with the epithelial surface (Lidell et al., 2006).

Then, the trophozoites attach to the tissue surface through its surface protein, namely Gal/GalNAc lectin. This leads to the cytolysis of the host cell (Tavares et al., 2005). A more recent study using an ex-vivo human intestinal model to study \textit{E. histolytica} pathogenesis found that impairing the Gal/GalNAc lectin did not inhibit the parasite’s attachment ability (Bansal et al., 2009). This study suggests that other molecules may also be involved in the adherence process. Subsequently, the destruction of the villin and microvilli occur as the trophozoites continue to secrete cysteine proteases.

The prevention of \textit{E. histolytica} from invading the sub-epithelium region involves the production of nitric oxide (NO) and reactive oxygen species (ROS) in the activated macrophages. However, the parasite is able to defend itself from the attack by neutralising both NO and ROS with peroxiredoxin (Choi et al., 2005).

Various proteins are involved in progressing intestinal amoebiasis to extraintestinal amoebiasis. For example, the amoebapores play a crucial role in the
establishment of amoebic liver abscess (ALA) (Zhang et al., 2004). In addition, according to Santi-Rocca et al. (2008), the parasite upregulates the expression of lysine and glutamic acid-rich protein (KERP1) during ALA formation. The roles of these proteins are postulated in the protection of this parasite from an acute immune response during the development of ALA.
Figure 2.3 Pathogenesis of intestinal amoebiasis (Lejeune et al., 2009)
2.2 The proteome of \textit{E. histolytica}

The term proteome was coined by Marc Wilkins in 1994 and it was defined as an entire complement of proteins expressed by the genome, cell, tissue or the entire organism (Wilkins et al., 2013). The overall aim of proteome studies is to perform large-scale discovery of the proteome by analysing many proteins at the same time. Large-scale proteomics approach, also known as a bottom-up strategy, has been proven to be an indispensable tool towards understanding the parasite pathogenesis (Veras and Bezerra de Menezes, 2016, Bertin et al., 2016).

For pathogens such as \textit{E. histolytica}, its proteome is important for the pathogenicity of the disease and the cell viability. In amoebiasis, \textit{E. histolytica} evades the immune system using surface receptor capping on the uropods. In this event, the targeted host immune components on the parasite’s surface are translocated and shed at the uropod (Avila-Calderón et al., Espinosa-Cantellano et al., 1992, Markiewicz et al., 2011). The shedding of uropods from the parasite suggests that the isolated fraction contains various molecules at the plasma membrane. The uropod shedding also suggests its involvement in the excretory mechanism of the trophozoites. In a study to understand the mechanism of uropod formation, the identification of the uropod proteome showed several numbers of multiple drug resistance proteins, ATPases, GTPases, and cysteine proteases (Markiewicz et al., 2011).

The \textit{E. histolytica} proteome plays an important role in the encystation and excystation in response to the environment. Proteome analysis of the total \textit{E. histolytica} protein identified 1029 proteins from the trophozoite form, 550 proteins from the cyst-like structure (CLS), and 411 proteins from the cysts, with only 74
proteins found to be common across all the three forms (Luna-Nácar et al., 2016). This study suggests that CLS may be an intermediate survival strategy of trophozoites towards stressful condition, of which the process enables the parasite to form a chitin-like resistant cover containing Jacob protein as a shelter.

During host infection, *E. histolytica* is exposed to reactive oxygen species that are released by the host’s immune cells at the site of the infection. Shahi et al. (2016) identified 154 oxidising proteins, in which these proteins were involved in transport, catalysis, antioxidant activity, and maintaining the parasite's cytoskeleton. They also reported the involvement of arginase in the protection of the parasite against oxidative stress that was induced by the host. These results emphasise the contribution of oxidative stress by the host cells to the pathogenesis of *E. histolytica*.

Proteome analysis aimed to elucidate the migration-related proteins have identified EhPC4 (positive coactivator 4) to be responsible for the underlying mechanisms of *E. histolytica* trophozoites migration (de la Cruz et al., 2014). They have identified 16 differentially expressed proteins, of which four up-regulated proteins were involved in cytoskeleton organisation and cell migration. They observed that the overexpression of EhPC4 induced a significant increase in the trophozoite migration and the destruction of human SW480 colon cells. Hence, these proteins play an important role in the virulence of *E. histolytica*.

Many *E. histolytica* proteins play important roles in the parasite’s pathogenesis, either by direct involvement or by important intracellular process. The PI3K family of intracellular signalling enzymes play a role in the early stages of phagosome formation (Powell et al., 2006, Nakada-Tsukui et al., 2009). Further proteome analyses on the phagosome have identified many GTPase families
(Rodríguez et al., 2000, Okada et al., 2006, Hernandes-Alejandro et al., 2013). These studies indicated that many protein members of the *E. histolytica* proteome is required for amoebic trogocytosis and phagocytosis to occur. In addition, the GTPase families include the Rab proteins, such as EhRab7A, EhRabA and EhRabB are localised at the phagocytic cup and may be part of the *E. histolytica* secreted proteins (Ralston, 2015).

2.2.1 Excretory-secretory proteins

During infection, *E. histolytica* trophozoites release excretory-secretory (ES) proteins, which are also known as excretory-secretory antigens (ESA). ES proteins are involved in the invasion of trophozoites into the colonic mucosa by degrading the glycoside substrates and proteins of the host tissues (Keene et al., 1986, Scholze and Werries, 1986, Reed et al., 1993, Moncada et al., 2005). Antibodies against ES proteins have been detected in the sera of both symptomatic and asymptomatic patients who have contracted amoebiasis (Pal et al., 1996).

The use of ES proteins as potential targets for diagnosis, treatment, and vaccine development for amoebiasis has been explored in previous studies (Quach et al., 2014, Wong et al., 2011, Saidin et al., 2014, Debnath et al., 2012). In diagnostics, the *E. histolytica* Gal/Gal-NAc lectin antigen is utilised in commercial antigen detection tests, i.e., the TechLab *E. histolytica* II ELISA (TechLab Inc). Furthermore, Gal/Gal-NAc lectin also showed potential as a vaccine candidate against *E. histolytica* (Quach et al., 2014). Another study on ES proteins showed the diagnostic potential of pyruvate phosphate dikinase (PPDK), and its recombinant form was used to develop a lateral flow dipstick test (Wong et al., 2011, Saidin et al., 2014).
addition, auronofin was identified as an effective drug which targeted *E. histolytica* thioredoxin reductase (Debnath et al., 2012).

Proteome analysis on the ES proteins of *Trypanosome* sp. has uncovered a range of proteins which include unfolding and degradation classes of proteins, such as serine, cysteine proteases, and metallopeptidases (Nten et al., 2009). These proteases play a part in the physiological and pathological functions that favour the invasion of the parasite, growth in hostile host conditions, evasion of components of the host immune defence, and hydrolysis of host proteins.

*E. histolytica* secretome comprises a spectrum of proteins that may be needed for every facet of the parasite’s life cycle including cell modulation that is due to the environmental adaptation and the evasion of host’s immune responses. Hence, proteome studies of ES proteins may open paths to initiate novel strategies for the management and prevention of amoebiasis (Ahn et al., 2018).

### 2.2.2 Membrane proteins

Membrane proteins are important in many processes ranging from basic cellular process to self-defence and disease-causing processes (Santoni et al., 2000). Many of them are also potentially good drug targets, with an estimation of more than half of all drugs that have been developed targets the membrane proteins (Klabunde and Hessler, 2002). During *E. histolytica* infection, the parasite’s membrane proteins are used for tissue invasion, as well as the establishment of intra and extra-intestinal infections. Biller et al. (2014) reported that the surface proteome of *E. histolytica*
consisted of 693 proteins, whereby 87% of the identified proteins were estimated to be localised on the membrane surface.

Perdomo et al. (2015) identified more than 1500 *E. histolytica* endomembrane proteins. The top two classes of proteins were involved in trafficking machinery and GTPases proteins with 152 and 131 proteins, respectively. The analysis revealed a high abundance of proteins that were involved in the intracellular trafficking mechanism. The most abundant protein was calreticulin, which resided at the endoplasmic reticulum and functioned as a calcium-buffer and a chaperone. It was localised at the plasma membrane and it was involved in the host cell interaction and the formation of phagocytic cups (Short et al., 2005).

The membrane proteome is a landfill for the exploitation of biomarkers. In a proteome analysis by Che et al. (2011), over two thousand *Toxoplasma gondii* membrane proteins were identified. Over 40% of the identified membrane proteins were hypothetical. Furthermore, many of the membrane proteins identified were unique to *T. gondii*. Hence, the study provided a set of proteins that are suitable for further experimental investigation.

The importance of identifying and studying membrane proteins is highlighted by the fact that they account for 70–80% of all drug targets. In addition, it is estimated that the majority of future drug targets are the membrane proteins (Hopkins et al., 2006, Overington et al., 2006). Therefore, the study of membrane protein in *E. histolytica* may even precede over ES proteins in search of potential novel biomarkers for drug targets. However, the detection of membrane proteins by standard proteomic methods is challenging due to the low abundance of membrane proteins relative to the total cell lysates, and their hydrophobic characteristics.
(Santoni et al., 2000, Wallin and Heijne, 1998). Therefore, it is important to establish a protocol for isolation of membrane proteins prior to mass spectrometry analysis.

2.3 Tools for proteome discovery

2.3.1 Sample preparation

In-gel and in-solution digestions are the two common approaches in a bottom-up proteomic sample preparation (Figure 2.4). However, in the context of a complex protein sample, in-solution digestion requires a post peptide separation while in-gel digestion is already the result of pre-protein separation (Gundry et al., 2010).

In-gel digestion followed by a mass spectrometry analysis are widely used techniques to identify proteins (Lasonder et al., 2002, Nten et al., 2009, Pomastowski and Buszewski, 2014). Before protein digestion, separation of the protein is performed using sodium dodecyl sulfate in a polyacrylamide gel (SDS-PAGE). Then, individual protein band or spot can be cored out to proceed with in-gel digestion. Also, depending on the complexity of the sample, several strategies are required for an efficient mass spectrometry analysis. This includes the need to consider the application of either one or two-dimensional gel electrophoresis (2-DE) (Pomastowski and Buszewski, 2014).

In-solution digestion followed by mass spectrometry is one of the simplest and commonly used techniques (de Souza et al., 2006, Biller et al., 2014, Perdomo et al., 2015). This technique involves denaturing, reducing, alkylating, and digesting the protein sample in the liquid phase. The fractionation is usually performed after the digestion. Nonetheless, separation can also be performed prior to digestion using
different forms of chromatography tools, including, reverse-phase, strong and weak ion exchange, as well as size exclusion chromatography (Mostovenko et al., 2013).

The in-gel digestion has several advantages over the in-solution digestion. In this method, sample complexity can be reduced without using liquid chromatography (LC). On the other hand, the in-solution digestion method requires fractionation using LC after the sample digestion step. Hence, as increased sample complexity and the efficiency of protein identification require longer mass spectrometry time, cost-benefit may favour gel-based mass spectrometry compared to in-solution digestion (Rabilloud and Lelong, 2011). The in-solution digestion has its own advantages such that it is simple and straightforward to perform. Furthermore, the sample recovery of in-gel digestion is estimated to be 70 – 80% of the in-solution efficiency (Shevchenko et al., 2006, Gundry et al., 2010). In addition, the protein sample concentration and amount for in-solution digestion are fixed and hence the protein quantity can be controlled. However, for in-gel digestion, the amount of proteins digested from the gel is difficult to ascertain, though the amount of the initial protein load can be controlled (Zhou et al., 2005).
Figure 2.4  A bottom-up proteomics workflow.
2.3.2 Fractionation and peptide separation

In proteomics, several complications are present due to the unique features in a wide variety of proteins. In the cell, quantitative ratio of low abundant proteins and high abundant proteins faces a large gap of $10^6$ molecules (Veenstra, 2006, Rabilloud, 2009). Furthermore, for MS/MS analysis, the sample complexity is increased with the digestion of proteins into multiple peptides using proteolytic enzymes (Jafari et al., 2012). Although several studies have demonstrated that fractionation at the protein level can be more effective for protein profiling, fractionation at the peptide level is usually easier to perform (Thakur et al., 2011b). Fractionation reduces the complexity of a sample and improves sensitivity in MS/MS analysis. However, unavoidable sample loss in each fractionation step can be a setback in cases where samples are limited (Thakur et al., 2011b, Feist and Hummon, 2015).

Protein fractionation can be divided into two techniques: gel-based and gel-free. Examples of gel-based techniques are 1D, 2DE, native PAGE, and isoelectric focusing using immobilised pH gradient gel strips (Baggerman et al., 2005). In gel-based proteomics, protein fractionation is based on the protein’s molecular weight or isoelectric point (pI) or both. Fractionation of protein samples through gel-based techniques are less costly and easy to use (Thakur et al., 2011a, Shevchenko et al., 2006). However, challenges from the application of gel-based techniques may include poor recovery of proteins for MS/MS analysis, poor reproducibility in gel to gel replicates, and poor separation for complex protein samples of extreme MW, pI, and hydrophobicity (Granvogl et al., 2007, Albright et al., 2009).

Gel-free fractionation utilises different liquid phase chromatography methods such as reversed-phase, ion exchange, size-exclusion, and affinity chromatography.
The fractionation of protein/peptide in gel-free approaches involves different physiochemical principles of separation, such as hydrophobicity, molecular size and charge. In contrast to gel-based techniques, gel-free fractionation methods allow better reproducibility, sample recovery, and compatibility with automation as well as on-line attachment to a mass spectrometer. However, liquid chromatography techniques often require specialised equipment and skilled personnel (Simpson, 2004, Motoyama and Yates III, 2008, Chen and Pramanik, 2009).

Liquid chromatography approaches are traditionally used for fractionation of peptide samples in bottom-up proteomics. The most common liquid chromatography stationary phase used is the C-18 reverse-phase column (Feist and Hummon, 2015). The reverse-phase column takes advantage of the hydrophobicity interaction between the peptides and an adsorbent such as the C-18 column. For fractionation, a gradient of low to a high organic solvent such as acetonitrile is used for the separation. Acidified organic solvents are commonly used as organic-phase, also known as “buffer B”, due to their miscibility with aqueous solutions. Acidified water is used as the “weak” solvent, also known as “buffer A”. Both buffers are acidified with either formic acid (FA) or trifluoroacetic acid (TFA) (Li and James, 2013).

2.3.3 Protein identification by mass spectrometry

In analyses involving tandem mass spectrometry or MS/MS, the peptides are first ionised using an ionisation system. The resulting parent peptide ions are then subjected to fragmentation using collision with an inert gas. This results in raw fragmentation spectra that provide information of the parent ion.
Mass spectrometry is an important analytical technique that allows large-scale identification of proteins. In the early days, the technology lacked efficacy in the analysis of larger biomolecules such as peptides and proteins. The development of two ionisation techniques, electrospray ionisation (ESI) and Matrix-assisted laser desorption ionisation (MALDI) were a breakthrough in proteomics (Karas and Hillenkamp, 1988, Fenn et al., 1989). These two ion sources, linked to mass analysers, such as Orbitrap and Time-of-Flight (TOF), have become important tools in proteome research. This has led to the development of analytical technique, Peptide Mass Fingerprinting (PMF) followed by MS/MS which in turn allowed bottom-up proteomics to be performed (Henzel et al., 2003).

2.3.3(a) Matrix-assisted laser desorption ionisation and time-of-flight analyser (MALDI-TOF)

MALDI was introduced by Michael Karas, Franz Hillenkamp and Koichi Tanaka (Karas and Hillenkamp, 1988, Tanaka et al., 1988). The researchers found that organic samples could be ionised more easily by laser if it was mixed with other organic molecules that act as a matrix. Only a few microlitres of the sample-matrix mixture is placed on the target plate and allowed to dry. The drying process forms a crystal lattice incorporated with the sample peptides. The matrix absorbs most of the energy from the laser and transfers the charge to the sample, thus ionising the analyte. The resulting singly charged ions are caused by the protonation and deprotonation in positive and negative ion modes, respectively (Hoffman and Stroobant, 2007).
MALDI is often coupled with TOF mass/charge analyser that fits the pulsed nature of MALDI (Wolff and Stephens, 1953). The analyte accepts a single proton and this results in singly charged ions with large m/z values (Veenstra, 2006). Therefore, mass analysers with large m/z detection range such as the TOF are interfaced with MALDI. The ions generated by laser pulses in the ion source are accelerated in an electric field, then the ions enter a flight tube which has a detector at its end. The detection of mass and ion charge are determined from the time of flight and charge of the electron. In tandem MS analysis, selected precursor ions are passed into the collision cell. Here, the fragmented ions are decelerated and reaccelerated for the measurement of m/z in the second TOF analyser (Dass, 2007).

2.3.3(b) Electrospray ionisation (ESI)

The principle of ESI was first conceived by Malcolm Dole (Dole et al., 1968). The first researcher to extend the basis of this technology to study biomolecules was John Fenn (Fenn et al., 1989). Since then, ESI has augmented the ability to perform protein characterisation in MS analysis.

The principle of which ESI works is relatively simple. The sample in the solution form flows through a capillary which leads to the ionisation source region in the mass spectrometer. The ionisation of the sample begins by conducting high voltage current on the capillary where the sample flow. The sample exits the spray tip as aerosol droplets that contain both the ionised solute and analyte. With the evaporation of the solute, desorption of the analyte ion occurs. Finally, the analytes are drawn into the mass analyser region based on the differential pressure between the ionisation region and the mass analyser region (Veenstra, 2006).
2.3.3(c) The advantages and disadvantages of MALDI and ESI

The mass spectrometry analyses of macromolecules such as peptides are due to the development of soft ionization techniques such as MALDI and ESI that can transform these molecules into ions. Several studies conducted to investigate the application of MALDI and ESI on peptide detection reported that the ESI was able to continuously generate ions so that it can easily combined with liquid chromatography systems and various mass analysers. In comparison to MALDI, nanospray ESI sources has improved sensitivity in the detection of peptide in complex samples, hence resulting in high sensitivity and superior quality in MS/MS spectra (Aebersold and Mann, 2003, Nadler et al., 2017). On the other hand, MALDI couples to a MS/MS analyser has also been used for peptide identification, and is able to generate MS/MS spectra that are both analogous and complementary to the spectra produced by ESI-MS/MS (Bienvenut et al., 2002, Bodnar et al., 2003).

The versatility of ESI to be interfaced with a peptide separation technique as an on-line system, has allowed the technique to be dominant in large-scale proteomic analysis. However, for the off-line system i.e. MALDI, the chromatography separation is not coupled to the ionisation system. This has allowed for the sample reanalysis. While this situation may be advantageous in the case of reanalysis of a sample, the technique has resulted in lower introduction of precursor ions into the MS/MS system. As a consequence, data sets such as peptide and protein numbers obtained through MALDI is lesser than ESI-MS/MS (Ujang et al., 2016, Bodnar et al., 2003).

The on-line peptide fractionation technique coupled to the ionising source of ESI favours MS/MS detection chances for the hydrophobic analyte (Nadler et al.,
This occurs through the accumulation of hydrophobic analytes at the surface of a droplet in the Taylor cone emission process during the ESI process (Wilm, 2011). On the other hand, MALDI is predominantly controlled by the gas-phase basicity, whereby it is reported that ion yields in peptides identified through ESI-MS/MS contained a lower average number of basic amino acids than peptides identified in MALDI-MS/MS (Nishikaze and Takayama, 2006).

MALDI predominantly produces ions that are singly charged, hence minimizing spectral complexity. However, ESI typically produces multiply charged ions (Li and Cole, 2010). Nadler et al (2017) reported that higher charge state analyte with a mass below 1200 Da cannot be detected by ESI. This causes a disadvantage for ESI system in detection of peptides with a low molecular weight that contains additional basic amino acids.

Although MALDI-MS/MS and ESI-MS/MS are two mechanistically different ionization approaches, the complementary character of these two techniques suggest that larger proteome coverage can be achieved through the application of both techniques.

2.3.3(d) Mass spectrometry - Peptide fragmentation

Mass spectrometry techniques for peptide fragmentation include collision-induced dissociation (CID), electron capture dissociation (ECD), and electron-transfer dissociation (ETD) (Hart-Smith, 2014, Wells and McLuckey, 2005). Primarily, individual peptide ions are isolated, fragmented and the masses recorded to obtain partial or complete information. From the generated tandem MS spectrum of a
particular peptide, the partial primary sequence can be determined by the comparison of major peaks in the generated spectrum with the theoretical molecular masses of the amino acid monomers within the peptide.

The CID fragmentation requires the establishment of high vacuum prior to excitation and detection (Bogdanov and Smith, 2005). Upon that, the principle of this method depends on the collision of the analyte within a gas phase collision region such as a collision cell, post-source decay in MALDI-TOF or an ESI ion source. Subsequently, this triggers the cleavage of peptide bonds (or C-N bonds), resulting in a series of \( b \) and \( y \) fragment ions (Steen and Mann, 2004, Liao et al., 2009, Zhang et al., 2013). The generation of ion excitation through CID is vibrational in nature (Mayer and Poon, 2009). Such occurrence redistributes internal energy across the target ion and therefore cause dissociation via cleavage of the weakest bond. This results in faster and higher efficiency of ion dissociation CID.

The principles of both ETD and ECD share a common similarity by inducing fragmentation in multiply-charged cations by transferring electrons to the target ion (Elviri, 2012). As low energy electron is transferred or captured, the odd-electron species undergo rearrangement. This results in the cleavage of N−C backbone. Unlike CID, dissociation involving such low electron energy occurs without internal energy redistribution, thus avoiding the cleavage of the weakest bonds (Zubarev et al., 2002). The use of ECD was found to be capable of inducing peptide fragmentation, and the generation of c and z fragment ions while allowing post-translational modifications localisation to remain intact (Shi et al., 2001, Chalmers et al., 2004). Spectra information of CID and ECD/ETD are complementary. Hence, instrumentation with the option of alternating between the modes increases confidence in the results (Kim et al., 2011).
2.3.3(e) Database search

The identification of proteins involves the comparison of observed tandem MS spectra generated from peptide fragmentation with the theoretical MS spectra generated from the *in silico* peptide maps of a protein database (Patterson and Aebersold, 1995, Steen and Mann, 2004). In order to achieve this, various client-server applications such as ProteinPilot™ Paragon™, Proteome discoverer, PeaksStudio were developed to process and report the mass spectrometry data. Such software utilises database search engines such as SEQUEST, MASCOT, and ESIprot, which are available for interpretation from raw spectra information generated in an experiment to peptide/protein identification. (Deutsch et al., 2008). These algorithms depend on a complete sequence genome/protein database. With the advancement of DNA sequencing, the availability of complete genome database and new bioinformatic tools, large-scale protein identification has been accomplished (Zhang et al., 2013).

2.3.4 Protein topology prediction

2.3.4(a) Classically and non-classically secreted proteins

In the classical secretion system, proteins are often processed and packaged through the endoplasmic reticulum and Golgi apparatus. Then, secretory vesicles are formed and subsequently fused to the cell membrane before being released to the extracellular environment. The criteria for proteins that are classically secreted include the presence of signal peptide at the N-terminal which are 16 – 30 amino
acids long (Tsirigos et al., 2015). This segment is typically hydrophobic and can form only a single α-helix, which ends with a signal peptidase cleavage site (Petersen et al., 2011).

Proteins that are secreted without the N-terminal signal peptide are known as the leaderless secretion or the non-classical secretory pathway (Kuchler et al., 2013). The secretion of non-classically secreted proteins are mostly caused by stresses such as the lack of nutrient, endoplasmic reticulum (ER) stress, or mechanical stress that are inflicted by the host immune system (Rabouille, 2017). Like classically secreted proteins that contains the N-terminal leader sequence, these leaderless proteins can also be translocated across the plasma membrane through pores or via membrane-bound organelles. The non-classical secretory system works independently from the ER–Golgi network. These secreted proteins do not undergo glycosylation even with the presence of glycosylation motifs. However, a study by Bendsten, et al. (2005) has found that non-classically secreted proteins are proteins that contain no signal peptide or motifs but its secondary structure forms disordered regions which are more structurally disordered from their cytoplasmic protein counterparts.

2.3.4(b) Transmembrane topology

A typical membrane protein consists of at least one transmembrane segment that is predominantly made up of hydrophobic residues that are long enough to span the lipid bilayer as a helix bundle. Such characteristics of a membrane protein can be predicted based on the amino acid sequence. Hence, bioinformatic tools can then be used to analyse information such as hydrophobicity scales as a basis for
transmembrane span prediction, protein topology based on the transfer of free energy, as well as α-helix and β-sheet prediction (Koehler Leman et al., 2015)

One of the first membrane protein prediction principles was based on measurements of amino acid sequence hydrophobicity to predict transmembrane segments (Kyte and Doolittle, 1982). Simple principle derived from amino acid sequences like hydrophobicity and topology information such as simple helix bundles formation remain legit. However, with more studies, many other characteristics of a membrane protein was discovered. Such characteristics include tryptophan and tyrosine clusters near the end of a transmembrane region, “positive-inside rule” involving positively-charged amino acids in the formation of re-entry loop (Elazar et al., 2016).

Advanced algorithm has allowed the analyses on somewhat random appearance of sequence motifs, such as the GxxxG-motif in transmembrane proteins as well as other periodic patterns (Teese and Langosch, 2015). Progression such as this has been enabled by the availability of experimental information as a training data, availability of information in protein structure characteristics as well as the development of machine-learning algorithms. Machine-learning methods like the Hidden Markov’s Model (HMM) and Artificial Neural Network (ANN) are used to extract features of a protein sequence that may pertain to the characteristics of a membrane protein. Furthermore, such method uses evolutionary information like multiple sequence alignment to create consensus prediction for a submitted protein sequence (Tsirigos et al., 2015).

As more characteristics were discovered, automated analysis for protein topology predictions between various characterisation were also needed. As a result,
methods PRO-TMHMM, PRODIV-TMHMM and also OCTOPUS were created (Viklund and Elofsson, 2008). In addition, as signal peptides are often confused as TM segments, better methods that can distinguish signal peptide from transmembrane segments were developed: Phobius, Philius, and SPOCTOPUS (Käll et al., 2005, Reynolds et al., 2008). With various topology prediction methods, approaches like TOPCONS, MetaTM and CCTOP, are able to provide promising results by based on consensus decisions between the predictions (Dobson et al., 2015, Bernsel et al., 2009).
CHAPTER 3 – MATERIALS AND METHODS

3.1 Materials

3.1.1 Preparation of Culture medium, Buffers and Solutions

3.1.1(a) Trypticase-Yeast Extract-Iron and Serum Medium, TYI-S-33

Solution A was prepared by dissolving 10 g dextrose and 30 g Biosate peptone in 500 mL deionized water. Solution B was prepared by mixing 1 g potassium phosphate dibasic, 0.6 g potassium phosphate monobasic, 1 g L-cysteine, 2 g sodium chloride and 0.2 g ascorbic acid in 370 mL deionised water. Incomplete TYI-S-33 was prepared by combining solution A and solution B. After mixing solution A and B, 22.5 mg ammonium iron (III) citrate was added into the mixture. Subsequently, the pH was adjusted to 6.8 with 1 M NaOH. The solution was then filtered with a filter paper, and autoclaved at 121 °C, 15 psi for 10 min. The medium was then stored at -20 °C and used within 2 months.

Complete TYI-S-33 was prepared by adding 125 mL of heat-inactivated bovine serum and 25 mL of Diamond vitamin Tween 80 (40X) in a sterile condition. The complete medium was aliquoted into 50 mL falcon tubes and stored in 4 °C (< 1 week) or – 20 °C (< 2 months) until further use.

3.1.1(b) Heat-inactivated bovine serum

Bovine serum was completely thawed at room temperature prior to inactivation. Serum inactivation was performed at 56 °C for 30 min with a gentle shake at every
five-min interval. The serum was then aliquoted in sterile falcon tubes and stored at -20 °C until used.

3.1.1(c) Potassium phosphate monobasic solution, 0.15 M

This solution was prepared by dissolving 20.41 g of potassium phosphate monobasic in 1 L ddH₂O.

3.1.1(d) Potassium phosphate dibasic solution, 0.15 M

This solution was prepared by dissolving 26.12 g of potassium phosphate dibasic in 1 L ddH₂O.

3.1.1(e) Phosphate buffer, 0.15 M

A 0.15 M phosphate buffer was prepared by combining 400 mL of 0.15 M potassium phosphate monobasic solution into 700 mL of 0.15 M potassium phosphate dibasic solution and the pH was adjusted to 7.0. The buffer was then stored at 4 °C.

3.1.1(f) Phosphate Buffer Saline for Amoeba, PBS(A)

PBS(A) was prepared by dissolving 10.78 g of sodium chloride into 100 mL of 0.15 M phosphate buffer and then topped up with ddH₂O to a final volume of 1 L. The buffer was autoclaved at 121 °C, 15 psi for 10 min and stored at 4 °C until used.
3.1.2 Materials for isolation of excretory-secretory proteins

3.1.2(a) RPMI-C-A

A packet of 10.5 g RPMI-1640 was dissolved in 600 mL of ddH₂O. Next, 2 g of sodium bicarbonate, 1 g of L-cysteine and 0.2 g of ascorbic acid were added and stirred until dissolved. The pH of the solution was then adjusted to 7.0 and the volume was top up to 1 L with ddH₂O. Finally, the medium was sterilised using a filter in a biosafety cabinet and kept at 4 °C until used.

3.1.2(b) Ammonium bicarbonate, 50 mM

To prepare 50 mM ammonium bicarbonate, 0.19765 g ammonium bicarbonate was dissolved in 50 mL of ddH₂O. The solution was then stored in 4 °C and used within two days.

3.1.2(c) Protease inhibitor (Roche), 7×

A tablet of Roche Mini Protease inhibitor was dissolved in 1.5 mL of ddH₂O. The mixture was vortexed until the tablet was fully dissolved. The solution was then kept at 4 °C and used within 2 weeks.

3.1.3 Solutions for the conventional membrane protein extraction

3.1.3(a) Sodium phosphate monobasic, 10 mM

This solution was prepared by dissolving 0.11998 g of sodium phosphate monobasic in 100 mL deionized water.
3.1.3(b) Sodium phosphate dibasic, 10 mM

This solution was prepared by dissolving 0.14196 g of sodium phosphate dibasic in 100 mL deionized water.

3.1.3(c) 10 mM Sodium phosphate buffer, pH 8.0

This solution was prepared by slowly adding 10 mM sodium phosphate monobasic solution into 10 mM sodium phosphate dibasic until the pH reached 8.0. The solution was filtered and stored in 4 °C and used within a month.

3.1.4 Solutions for ProteoPrep® Membrane Extraction Kit (Sigma, USA)

3.1.4(a) Soluble Cytoplasmic and Loosely-bound Membrane Protein Extraction Reagent

The reagent inside the bottle was reconstituted with 125 mL of ddH$_2$O. The solution was then aliquoted and stored at –20 °C.

3.1.4(b) Protein Extraction Reagent Type 4

The reagent inside the bottle was reconstituted with 15 mL of ddH$_2$O. The solution was then stored as 1 mL aliquots at –20 °C until further use.

3.1.5 Solutions for ProteoExtract® Native Membrane Protein Extraction Kit

Extraction Buffers I and II were stored as 1 mL aliquots at -20 °C until further use.
3.1.6 Materials and buffers for SDS-PAGE

3.1.6(a) 10% SDS

Five grams of sodium dodecyl sulfate (SDS) was dissolved in 25 mL of ddH₂O and topped up to a final volume of 50 mL.

3.1.6(b) Resolving buffer

An amount of 22.7 g tris base was dissolved in 80 mL of ddH₂O. The solution was then adjusted to pH 8.8 by adding 2 M HCl. Finally, the solution was top up with ddH₂O to a final volume of 100 mL.

3.1.6(c) Stacking buffer

An amount of 7.3 g tris base was dissolved in 50 mL of ddH₂O. The solution was then adjusted to pH 6.8 by adding 2 M HCl. Finally, the solution was top up with ddH₂O to a final volume of 100 mL.

3.1.6(d) Ammonium persulfate (APS), 10%

Ten milligrams of ammonium persulfate was dissolved in 100 µL of ddH₂O. The solution was freshly prepared prior to use.

3.1.6(e) Loading buffer, 5×

A mixture comprised of 0.76 g of tris base and 1 g SDS were dissolved in 25 mL of ddH₂O. The pH of the solution was then adjusted to 6.8 with 2 M HCl. Then, 10 mL of glycerol and 0.025 g of bromophenol blue were added into the solution. Subsequently, the solution was top up with ddH₂O to a final volume of 50 mL. The loading buffer without β-mercaptoethanol was aliquoted in 900 µL and stored at -20
°C. One hundred microlitre of β-mercaptoethanol was mixed with 900 µL of incomplete loading buffer to make a complete buffer.

3.1.6(f) SDS-PAGE running buffer

To prepare 10× running buffer, 30 g of tris base, 10 g of SDS, and 144 g of glycine were dissolved in 700 mL of ddH₂O. Then, the solution was top up with ddH₂O to a final volume of 1 L. Finally, the solution was kept in 4 °C until used. Then, 100 mL of 10× running buffer was diluted with 900 mL of ddH₂O to make 1× buffer. The solution was reused for up to 4 times and was kept at 4 °C.

3.1.6(g) RAMA stain

3.1.6(g)(i) Coomassie Brilliant Blue, CBB R250, 0.05%

One gram of CBB R250 was completely dissolved in 300 mL of methanol. Then, the solution was top up to 500 mL with ddH₂O. The solution was kept at room temperature until further use.

3.1.6(g)(ii) Ammonium sulfate, 30%

Thirty grams of ammonium sulfate was completely dissolved in 100 mL of ddH₂O. The solution was kept at room temperature until used.

3.1.6(g)(iii) RAMA stain

To prepare 100 mL of working RAMA stain, 25 mL of 0.05% CBB, 10 mL of Acetic acid, 25 mL of methanol and 10 mL of 30% ammonium sulfate were mixed together. Then, the mixture was top up to 100 mL with ddH₂O. The solution was kept at room temperature until further use.
3.1.6(h) Separating gel and stacking gel

SDS-PAGE gel was prepared according to table 3.1.

Table 3.1  Preparation of SDS-PAGE gel for one small gel - 1.0 mm Mini-PROTEAN

<table>
<thead>
<tr>
<th>Solutions</th>
<th>Separating gel (10%)</th>
<th>Stacking Gel (3%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolving buffer (4×)</td>
<td>1.5 mL</td>
<td>-</td>
</tr>
<tr>
<td>Stacking buffer (4×)</td>
<td>-</td>
<td>0.5 mL</td>
</tr>
<tr>
<td>Bis-Acrylamide (30%)</td>
<td>2 mL</td>
<td>0.2 mL</td>
</tr>
<tr>
<td>SDS (10%)</td>
<td>60 µL</td>
<td>20 µL</td>
</tr>
<tr>
<td>Deionised water</td>
<td>2.374 mL</td>
<td>1.258 mL</td>
</tr>
<tr>
<td>APS (10%)</td>
<td>60 µL</td>
<td>20 µL</td>
</tr>
<tr>
<td>TEMED</td>
<td>6 µL</td>
<td>2 µL</td>
</tr>
</tbody>
</table>
3.1.7 Materials for sample digestion

3.1.7(a) RapiGest

One milligram of RapiGest lyophilized powder was reconstituted with 1 mL of freshly prepared 50 mM ammonium bicarbonate. The solution was then stored in 100 µL aliquots at -80 °C until further use. Working concentration of 0.05% RapiGest was then prepared by diluting 100 µL of 0.1% RapiGest with 100 µL of freshly prepared 50 mM ammonium bicarbonate.

3.1.7(b) 100 mM Dithiothreitol (DTT)

An amount of 7.7 mg of DTT was dissolved in 500 µL freshly prepared 50 mM ammonium bicarbonate.

3.1.7(c) 200 mM Iodoacetamide (IAA)

An amount of 18.5 mg IAA was dissolved in 500 µL freshly prepared 50 mM ammonium bicarbonate.

3.1.7(d) 1 µg/µL Trypsin

A vial containing 100 µg of lyophilized trypsin powder was reconstituted with 100 µL of the supplied trypsin resuspension buffer containing 50 mM acetic acid. The mixture was then stored as 5 µL aliquots at -80 °C until further use.

3.1.8 Materials for nanoLC-MALDI-TOF/TOF

3.1.8(a) Buffer A (2% ACN, 0.1% TFA, 97.9% ddH₂O)
To make 200 mL of buffer A, 4 mL of 100% acetonitrile (ACN), 195.8 mL of ddH₂O and 0.2 mL of trifluoroacetic acid (TFA) were mixed together. Then, the buffer was degassed using a water bath sonicator for 10 min.

3.1.8(b) Buffer B (98% ACN, 0.1% TFA, 1.9% ddH₂O)

To make 200 mL of buffer B, 196 mL of 100% ACN, 3.8 mL of ddH₂O and 0.2 mL of TFA were mixed together. Then, the buffer was degassed using a water bath sonicator for 10 min.

3.1.8(c) Matrix diluent (70% ACN, 0.1% TFA, 29.9% ddH₂O)

To make 2 mL of matrix diluent, 1.4 mL of 100% ACN, 0.598 mL of ddH₂O and 2 µL of TFA were mixed together.

3.1.8(d) Matrix α-cyanohydroxycinnamic acid (5 mg/mL)

Ten milligrams of α-cyanohydroxycinnamic acid (CHCA) powder was dissolved in 2 mL of matrix diluent. The mixture was vortexed until CHCA was fully dissolved. Then, it was briefly centrifuged to pellet down the undissolved particles.

3.1.8(e) Calibration mix

Calibration mix was prepared by adding 6 µL of CHCA matrix, 6 µL of matrix diluent, and 0.5 µL of calibration peptides. Finally, 0.3 µL of the mixture was spotted on the MALDI plate.
3.1.9 Buffers for LC – ESI (LTQ-Orbitrap Velos Pro system and Dionex UltiMate 3000 system)

Two buffers used to elute the peptides in this system consisted of Buffer A, 99.9% ddH₂O with 0.1% formic acid (FA) and Buffer B, 99.9% acetonitrile with 0.1% FA. To make 100 mL of buffer A, 99.9 mL of 100% ACN, and 0.1 mL of FA were mixed together. To make 100 mL of buffer B, 99.9 mL of ACN and 0.1 mL FA were mixed together. Both buffers were degassed using a water bath sonicator for 10 min prior to use.

3.2 Methods

3.2.1 Axenic culture of *E. histolytica* trophozoites

Before performing the maintenance of the culture, the growth of the trophozoites and the presence of contamination were observed using an inverted microscope. Then, a culture tube was gently tapped throughout and inverted a few times until sediment cells were unsettled from the bottom of the tube. Under a sterile environment, the old medium was discarded. Then, 9 mL of fresh TYI-S-33 medium (section 3.1.1 (a)) was refilled into the tube. The tube was incubated at 36 °C for 48-72 hours before subsequent maintenance. Culture waste was discarded into a container containing 10% sodium hypochlorite (Chlorox, Malaysia).
3.2.1(a) Expansion of the culture

For sub-culturing, a culture tube was flicked and gently inverted until unattached cells were mixed well. The cell suspension was then transferred into two new culture tubes. The new tubes were then closed and incubated horizontally with slant end facing downwards at room temperature for 5 min for the attachment of trophozoites. Subsequently, the old culture medium was discarded and replaced with the fresh culture medium. The culture tubes were incubated at 36 °C for 48-72 hours.

3.2.1(b) Mass Culture

Three culture tubes were flicked vigorously and gently inverted until unattached cells were mixed well. The cell suspensions were then transferred into a Nunc 75 cm² culture flask. The flask was then closed and incubated horizontally at room temperature for 10 min for attachment of the trophozoites. Next, the supernatant was discarded and the cell pellet was resuspended with fresh TYI-S-33 medium. The cell density was estimated via Trypan blue exclusion method. Subsequently, 2-3 × 10⁶ of viable cells were loaded into a 60 cm³ culture flask and filled with fresh TYI-S-33 medium to a final volume of 56 mL. The culture flask was then incubated at 36 °C for 48-72 hours.

3.2.2 Protein isolation and extraction

3.2.2(a) Isolation of excretory-secretory proteins

The RPMI-C-A medium used was prepared as described in section 3.1.2 (a). Trophozoites were rinsed three times with RPMI supplemented with 0.1% L-cysteine
and 0.02% ascorbic acid (RPMI-C-A) by centrifugation at 220 × g for 2 min. Trophozoites were then seeded into a culture tube filled with 7 mL RPMI-C-A medium at a cell density of 0.5 × 10^6 cells/ml, then incubated at 36 °C for 6 hours. The culture was centrifuged at 22 × g for 2 min and the supernatant was collected and mixed with 1 mM iodoacetamide. Then, centrifugation at 10,000 × g for 5 min was performed. Subsequently, the supernatant was collected and filtered through Minisart 0.2 µm pore membrane (Sartorius, Germany). The resultant excretory-secretory product was concentrated 1000 times using Vivaspin (GE Healthcare, UK), with molecular weight cut-off (MWCO) of 5 kDa by centrifugation at 3000 × g, 4 °C. To prevent protein degradation, protease inhibitors (Roche, Germany) was added. The ES proteins extracted was stored at -80 °C until further use. The protein concentration was estimated using RCDC assay.

3.2.2(b) Extraction of membrane proteins

3.2.2(b)(i) ProteoExtract® Native Membrane Protein Extraction Kit (Calbiochem, Germany)

Prior to the membrane protein extraction, Wash Buffer, Extraction Buffers I and II were thawed and mixed well by vortexing. Then, buffers were kept on ice throughout the extraction steps. Meanwhile, the Protease Inhibitor Cocktail was left at room temperature. Approximately 5 × 10^6 E. histolytica trophozoites cell pellet were washed with 2 mL ice-cold Wash Buffer. The cell pellet was released by resuspending with a pipette. The cells were then pelleted down at 440 × g, 4 °C for 2 min. Subsequently, the supernatant was discarded without disturbing the cell pellet. The washing step was then repeated once. After the final washing step, the Wash Buffer was completely removed. Next, 10 µL Protease Inhibitor Cocktail and 2 mL
ice-cold Extraction Buffer I was added to the cell pellet. The cell pellet was then completely resuspended using a pipette. The suspension was then incubated for 10 min at 4 °C under gentle agitation. After incubation, the suspension was then pelleted down at 16,000 × g, 4 °C for 15 min. The supernatant was then discarded. Subsequently, 5 µL Protease Inhibitor Cocktail and 1 mL ice-cold Extraction Buffer II were added to the pellet. The cell pellet was again gently resuspended using a pipette and was incubated for 30 min at 4 °C under gentle agitation. The insoluble material was then pelleted down at 16,000 × g, 4 °C for 15 min. Finally, the supernatant containing the membrane fraction was collected and stored at -20 °C.

3.2.2(b)(ii) ProteoPrep® Membrane Extraction Kit (Sigma, USA)

Soluble Cytoplasmic and Loosely-bound Membrane Protein Extraction Reagent and Protein Extraction Reagent Type 4 were completely thawed at room temperature. Then, the Soluble Cytoplasmic and Loosely-bound Membrane Protein Extraction Reagent were kept on the ice and the Protein Extraction Reagent Type 4 was kept at room temperature for further extraction steps. To begin, 5 × 10⁶ E. histolytica trophozoites cell pellet were suspended in 10 mL Soluble Cytoplasmic and Loosely-bound Membrane Protein Extraction Reagent. The suspension was then sonicated on ice with an ultrasonic probe for 1 min with power level 2 and 0.5 sec pulse on and off. Next, 30 mL Soluble Cytoplasmic and Loosely-bound Membrane Protein Extraction Reagent was added. The suspension was then incubated at 4 °C for 1 hour under gentle agitation. Subsequently, the suspension was ultracentrifuged at 115,000 × g and 4 °C for 1 hour. Next, the supernatant was discarded. The resulting pellet was washed twice with ddH₂O with centrifugation at 20,000 × g and 4 °C for 20 min. Then, the cell pellet was resuspended in 2 mL of Protein Extraction Reagent Type 4 and sonicated on ice at power level 2 with 0.5 sec pulse on and off for 1 min.
The suspension was then centrifuged at 14,000 × g, 15 °C for 45 min. Finally, the supernatant containing the membrane fraction was collected and stored at -20 °C.

3.2.2(b)(iii) Conventional Membrane Protein Extraction Method (Texeira)

The method was performed according to Teixeira & Huston (2008). Initially, 10 mM sodium phosphate buffer, pH 8 was kept on ice for the subsequent extraction steps. 10 × 10⁶ *E. histolytica* trophozoites cell pellet were suspended in 10 µL Protease Inhibitor Cocktail and 1 mL 10mM sodium phosphate buffer. The suspension was then sonicated on ice for 1 min using 0.5 sec pulse on and off with power level 2. Next, the suspension was centrifuged at 50,000 × g at 4 °C for 1 hour. Then, the supernatant was discarded and the cell pellet was resuspended in 1 mL ice-cold 10 mM Sodium Phosphate buffer and centrifuged at 100,000 × g and 4 °C for 1 hour. Subsequently, the supernatant was discarded and the resulting pellet containing membrane fraction was resuspended with 200 µL of 50 mM ammonium bicarbonate.

3.2.3 Protein concentration and buffer exchange

3.2.3(a) Protein concentration

A frozen medium containing ES proteins was thawed and concentrated using a column namely Vivaspin (GE Healthcare, UK) with 5kDa molecular weight cut-off. The tube containing the column was centrifuged at 4000 × g at 4 °C. After 100× concentration, 7× protease inhibitor (Roche, Germany) was added into the sample at a final concentration of 1× protease inhibitor. Subsequently, buffer exchange was performed three times by adding 3 mL of 50 mM ammonium bicarbonate and centrifuged at 4000 × g at 4 °C.
The membrane protein extracts were concentrated using Vivaspin (GE Healthcare, UK) with 5 kDa molecular weight cut-off. It was centrifuged at 3000 × g at 4 °C. After the 10× concentration, the samples were precipitated using 100% cold acetone.

### 3.2.3(b) Acetone precipitation

Prior to use, the 100% acetone was chilled at – 20 °C for at least 2 hours. Then, the cold acetone was added to the sample at a volume ratio of 5:1. The mixture was then incubated at -20 °C for 6 hours, followed by centrifugation at 15,000 × g for 10 min. Then, the supernatant was discarded. The precipitated protein pellet was resuspended in 50 mM ammonium bicarbonate with half the volume of the original sample volume.

### 3.2.3 Determination of protein concentration

The protein concentration was measured using Bio-Rad RC DC™ protein assay kit. This kit consisted of RC Reagent I, RC Reagent II, DC Reagent B, and DC Reagent S. Solution A’ was prepared by the addition of 5 µL of DC Reagent S into 250 µL of DC Reagent A.

Firstly, six dilutions of a protein standard consisted of 0.5 µg/µL, 1.0 µg/µL, 1.5 µg/µL, 2.0 µg/µL, 2.5 µg/µL, 3.0 µg/µL, 3.5 µg/µL, and 4.0 µg/µL were prepared from a 10 µg/µL bovine serum albumin (BSA) stock solution. Then, 12.5 µL of protein standard and samples were mixed well with 62.5 µL of RC Reagent I. The tubes were then incubated for 1 min at room temperature. Then, 62.5 µL of RC Reagent II was added to each tube and vortexed. The tubes were centrifuged at 15,000 × g for 5 min. The supernatant was carefully discarded without disturbing the protein pellet and the tubes were allowed to dry completely. After that, 63.5 µL of
solution A’ was added to each microfuge tube and vortexed. The tubes were incubated at room temperature for 5 min. Subsequently, 1 ml of DC Reagent B was added to each tube and was vortexed immediately. Then, the tubes were incubated at room temperature for 15 min. Finally, the absorbance was read at 750 nm within an hour using a spectrophotometer (NANODROP, ThermoScientific, USA). Protein concentration was determined using a standard curve based on a serial dilution of known bovine serum albumin concentrations (Figure 3.1).
Figure 3.1  A standard curve constructed from a serial dilution of known BSA concentrations.
3.2.4 Gel electrophoresis

Sixteen microlitre of protein sample containing an amount of 30 µg protein was mixed with 5× sample buffer at a ratio of 4:1. The mixture was incubated in a water bath at 100 °C for 2 min. Meanwhile, a reservoir consisting of the preparative gels was submerged with 1× running buffer. The sample was cooled in a water bath at 25 °C for a few seconds. Then, the sample was pipetted into the well. One well was designated for pre-stained protein marker as the protein size reference. Proteins were separated according to molecular weights with 10% resolving gel using Mini-PROTEAN Tetra cell as described in section 3.1.6 (h). Electrophoresis was run at a constant voltage of 100 V until the dye reached approximately 1 cm from the gel end. After electrophoresis, the gel was stained with RAMA stain for 30 min and de-stained in ddH2O until the background was cleared. RAMA stain was prepared as described in section 3.1.6 (g).

3.2.5 Protein digestion

One hundred micrograms of protein was top-up to 100 µL of 50 mM ammonium bicarbonate to achieve a concentration of 1 µg/µL. Next, 100 µL of 0.05% RapiGest was added to the mixture and mixed well by vortexing. The sample was then incubated at 80 °C for 15 min. Subsequently, the sample was allowed to cool at room temperature for approximately 5 min before 5 µL of 100 mM DTT was added. The sample was mixed and incubated at 60 °C for 15 min. Next, the sample was cooled to room temperature again for 5 min and 5 µL of 200 mM IAA was added. The sample was mixed well and then incubated at room temperature in the dark for 30 min. Then, 2 µL of 1 µg/µL trypsin was added. The mixture was incubated at 37 °C for 16 hours. Subsequently, 2 µL TFA was added and the mixture was incubated at 37 °C
for 20 min to stop the digestion and remove the RapiGest. The aggregated RapiGest was then pelleted down at 14,462 × g for 15 min. The supernatant containing the peptide mixtures was then collected and filtered with 0.45 µm minisart syringe filter. The peptide samples were then stored at -80 °C until further analysis.

3.2.6 Mass spectrometry analysis

3.2.6(a) LC–MALDI–TOF/TOF

Fractionation of the peptides was performed using Eksigent nanoLC ultra 1D plus linked to an automated MALDI spotter (Eksigent, Netherlands). To achieve spatial discrimination of the peptide mixtures, 2 – 5 µl of peptide samples was auto-loaded and packed into a C18 column. The gradient pump was set to elute the peptides with 20 to 80% acetonitrile for a duration of 165 min and at a flow rate of 0.3 µL/min. Mobile phase buffer A consisted of 0.1 % TFA in 2 % ACN and 97.9 % water while mobile phase buffer B consisted of 0.1 % TFA in 98 % ACN and 1.9 % water. The system was linked on-line to an automated MALDI spotter in which eluted peptides were spotted between the 30th and 160th min of gradient phase, with CHCA matrix flow of 1.8 µL/min for a duration of 25 sec for each spot.

Mass spectrometry analyses were performed in an automated (LC mode) on the AB Sciex TOF/TOFTM 5800 system. Data were obtained in the MALDI reflector mode using at least 6 spots of the internal calibration standard (TOF/TOF calibration mixture). Mass spectra from each spot were obtained in the m/z range from 800 to 4000, whereby up to 500 laser shots were accumulated per spectrum. The signal-to-noise (S/N) ratio was set to a minimum of 10, and the spots with the highest intensity of precursor ion were subjected to MS/MS analysis. A maximum of ten precursors were allowed for the MS/MS analysis; for each spectrum, up to 2000
laser shots were accumulated per spectrum, and the S/N were set to a minimum ratio of 15 S/N. The mass spectrometry data were analysed using ProteinPilot™ Software 4.5 and searched using Paragon against a combined AmoebaDB_4.1 and cRAP (‘protein contaminants database’) which was set to search with the following parameters: false discovery rate of <1%, detected protein threshold of >0.47 (66%), and competitor error margin of 2.00. The cRAP includes the possible contaminant proteins in this study such as BSA and keratin.

3.2.6(b) LC–ESI–MS/MS (LTQ-Orbitrap Velos Pro)

The ES protein peptide samples were sent to Analytical Biochemistry Research Centre (ABrC) at Universiti Sains Malaysia for mass spectrometry analysis. Peptide fractionation and mass spectrometry analysis were performed by using the liquid chromatography system, Easy-nLC II (Thermo Scientific, San Jose, CA, USA) coupled with a mass spectrometer, LTQ-Orbitrap Velos Pro (Thermo Scientific, San Jose, CA, USA). Fractionation of peptides was performed using Easy-Column C18-A2 (100 × 0.75 mm i.d., 3 µm; Thermo Scientific, San Jose, CA, USA) coupled with pre-column (Easy-Column, 20 × 0.1 mm i.d., 5 µm; Thermo Scientific, San Jose, CA, USA) at a flow rate of 0.3 µL/min and sample injection volume of 10 µL. The pre-column was equilibrated for 15 µL at a flow rate of 3 µL/min whereas the analytical column was equilibrated for 4 µL at a flow rate of 0.3 µL/min. The running buffers used were (A) deionised distilled water with 0.1 % formic acid and (B) ACN with 0.1 % formic acid. The samples were eluted using a gradient of B from 5 % to 100 % in 100 min. The fragmentation technique used was Collision Induced Dissociation (CID). De Novo sequencing and database matching (against AmoebaDB_4.1) was performed using Peaks Studio Version 7 (Bioinformatics Solution, USA). The sequenced peptides were also matched against the cRAP
database (www.thegpm.org/crap/) that contained a list of commonly found contaminating proteins. The parameters used in De Novo sequencing were precursor mass tolerance at 0.1 Da and fragment mass error tolerance at 0.8 Da. Subsequently, for the database matching, the parameters were carbamidomethylation and methionine oxidation as fixed modifications, 2 maximum missed cleavages, false detection rate (FDR) <0.1% and parent mass and 0.1 Da precursor mass tolerance. Besides, significant score (-10lgP) for protein acceptance were set at >20, whereas minimum unique peptide was set at 1.

3.2.6(c) LC–ESI–MS/MS (Orbitrap Fusion)

The peptides from the membrane and cytosolic fractions were sent to Proteomics Core Facility, Malaysia Genome Institute, National Institutes of Biotechnology Malaysia (NIBM) for mass spectrometry analysis. Briefly, the peptides separation was performed using Dionex UltiMate 3000 RSLCnano (Thermo Fisher Scientific, USA) nano-liquid chromatography system while mass spectrometry analysis was performed using Orbitrap Fusion (Thermo Fisher Scientific) mass spectrometer. For spatial discrimination of the peptides, they were packed into and eluted with EASY-Spray Column PepMap® RSLC, C18, 2 µm particle size, 50 µm id x 150 mm coupled with pre-column (µ-Precolumn PepMap 100, C18. 3µm particle size, 300µm id x 5mm) at a flow rate of 0.3 µL/min with mobile phase buffer A (deionised distilled water with 0.1 % formic acid) and B (ACN with 0.1 % formic acid). The sample was eluted for 101 min with a gradient of mobile phase B, 5-95% for 93 min, 95% for 2 min, and back to 5% in 2 min. For mass spectrometry analysis, the instrument was operated in the data-dependent mode. The parameters for the full scan spectra were as follow: scan range 310-1800 m/z, resolving power of 120000, AGC target of 4.0 e5 (400 000), and maximum injection time of 50 ms. The method
consisted of 3 sec Top Speed Mode where precursors were selected for a maximum of 3 second cycle. Only precursors with an assigned monoisotopic $m/z$ and a charge state of 2 – 7 were further analysed for MS/MS. All precursors were filtered using a 20 second dynamic exclusion window and intensity threshold of 5000. The MS/MS spectra were analysed using the following parameters: rapid scan rate with a resolving power of 60000, AGC target of 1.0e2 (100), 1.6 $m/z$ isolation window, and a maximum injection time of 250 ms. Precursors were fragmented by CID and HCD at normalised collision energy of 30% and 28%, respectively. Data analysis and database matching against AmoebaDB_4.1 and cRAP was performed using Proteome DiscovererTM Software Version 2.1 (Thermo Scientific). The parameters used in the analysis were fixed modifications: carbamidomethylation (C), variable modification: Oxidation (M), deamidation of asparagine (N) and glutamine (Q), maximum missed cleavage set at 2, false detection rate (FDR) <0.1% and parent mass and precursor mass tolerance at 10 ppm and 0.6 Da, respectively. Significant score (-10lgP) for protein acceptance were set at >20, whereas minimum unique peptide was set at 1.

3.2.7 Data analysis

Match peptide of more than 1 was considered for further analysis. Common and uniquely identified proteins by LC–MALDI–TOF/TOF and LC–ESI–MS/MS were illustrated with the aid of a Venn diagram generator at http://bioinfogp.cnb.csic.es/tools/venny/.

3.2.7(a) Prediction of excretory-secretory proteins

The hydrophobic N-terminal Signal Peptide that is required for classically secreted proteins were predicted by TOPCONS 2.0 server at http://www.topcons.net/pred/.
Meanwhile, the non-classically secreted proteins were predicted using SecretomeP 2.0 server at http://www.cbs.dtu.dk/services/SecretomeP/. Proteins that exceeded NN score of 0.6 and did not have N-terminal Signal Peptide that were predicted by TOPCONS 2.0 were considered as non-classical secretory proteins.

3.2.7(b) Prediction of membrane protein

Membrane protein prediction was analysed by TOPCONS 2.0 server at http://www.topcons.net/pred/. In this server, amino acid sequences of the identified proteins were analysed for protein topology using five sub-methods (OCTOPUS, Philius, PolyPhobius, SCAMPI and SPOCTOPUS). The final results were decided from consensus prediction of all the sub-methods based on the presence of signal peptides and the number of transmembrane regions of $\geq 1$.

3.2.7(c) Functional annotation

Identified proteins were classified using PANTHER (Protein Analysis Through Evolutionary Relationships) Classification System version 11.1 (released 2016-10-24) at http://www.pantherdb.org/. Functional annotation based on biological processes and molecular functions were searched specific to *E. histolytica* using AmoebaDB based accession number.
CHAPTER 4 – RESULTS

4.1 ES proteins

4.1.1 Protein profile of ES proteins

The isolated ES protein band profiles between three biological replicates were observed in the SDS-PAGE. As shown in Figure 4.1, similar protein band profiles on lane 2, 3 and 4 indicated good reproducibility. The RPMI-C-A as a blank medium in Lane 5 showed that the medium was indeed protein free and did not contribute any artefact in the ES protein band profiles. In addition, visual observation showed distinct protein bands with no protein smearing. This suggested that the ES proteins were suitable for downstream mass spectrometry analysis.

4.1.2 ES proteins identification by LC-MALDI-TOF/TOF and LC-ESI-MS/MS

Protein identification was performed using two mass spectrometry systems, namely LC-MALDI-TOF/TOF and LC-ESI-MS/MS. Two replicates were analysed with LC-ESI-MS/MS, and three replicates with LC-MALDI-TOF/TOF. Examples of the mass spectrometry results were summarised in Table 4.1 and 4.2. The overall results can be found in supplementary material 1 and 2.

A total of 209 *E. histolytica* ES proteins were identified. As shown in Figure 4.2, 97 and 191 proteins were identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS, respectively. Both systems identified 79 proteins. Meanwhile, 18 and 112 proteins were exclusively identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS, respectively.
Figure 4.1 Protein profiles of *E. histolytica* ES proteins from three biological replicates.

An amount of 30 µg protein samples were loaded to each well. The samples were denatured in 1× loading buffer and incubated in 100 °C water bath. The SDS-PAGE was ran under the following conditions: 3% stacking gel, 10% resolving gel under 100 V for 90 min. Lane 1: Precision Plus Proteins Kaleidoscope Standards (Bio-rad, USA); Lane 2 – 4: three independent batches of ES proteins; Lane 5: Blank RPMI-C-A medium.
Table 4.1  Examples of *E. histolytica* ES proteins identified by LC-ESI-MS/MS

<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>-10lgP</th>
<th>Matched Peptides</th>
<th>Coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Rep 1</td>
<td>Rep 2</td>
<td>Rep 1</td>
</tr>
<tr>
<td>EHI_150490</td>
<td>aldehyde-alcohol dehydrogenase 2 putative</td>
<td>149.5</td>
<td>180.5</td>
<td>56</td>
</tr>
<tr>
<td>EHI_051060</td>
<td>pyruvate:ferredoxin oxidoreductase</td>
<td>148.6</td>
<td>174.2</td>
<td>50</td>
</tr>
<tr>
<td>EHI_009530</td>
<td>pyruvate phosphate dikinase</td>
<td>147</td>
<td>172.1</td>
<td>47</td>
</tr>
<tr>
<td>EHI_030750</td>
<td>hypothetical protein</td>
<td>133.3</td>
<td>159.7</td>
<td>39</td>
</tr>
<tr>
<td>EHI_166920</td>
<td>hypothetical protein</td>
<td>130.7</td>
<td>156.6</td>
<td>35</td>
</tr>
<tr>
<td>EHI_044970</td>
<td>malic enzyme putative</td>
<td>132.6</td>
<td>156.4</td>
<td>33</td>
</tr>
<tr>
<td>EHI_008200</td>
<td>glyceraldehyde-3-phosphate dehydrogenase putative</td>
<td>117.6</td>
<td>151.6</td>
<td>26</td>
</tr>
<tr>
<td>EHI_077500</td>
<td>galactose-specific adhesin 170kD subunit</td>
<td>134</td>
<td>154.3</td>
<td>36</td>
</tr>
<tr>
<td>EHI_167320</td>
<td>glyceraldehyde-3-phosphate dehydrogenase putative</td>
<td>117.6</td>
<td>151.6</td>
<td>26</td>
</tr>
<tr>
<td>EHI_096710</td>
<td>type A flavoprotein putative</td>
<td>126</td>
<td>150.2</td>
<td>30</td>
</tr>
</tbody>
</table>
Table 4.2 Examples of *E. histolytica* ES proteins identified by LC-MALDI-TOF/TOF

<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>Score</th>
<th>Matched peptides</th>
<th>Coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHI_130700</td>
<td>enolase, putative</td>
<td>23.15</td>
<td>13</td>
<td>59.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td>55.1</td>
</tr>
<tr>
<td>EHI_150490</td>
<td>aldehyde-alcohol dehydrogenase 2</td>
<td>18.96</td>
<td>12</td>
<td>27.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>25.4</td>
</tr>
<tr>
<td>EHI_009530</td>
<td>pyruvate phosphate dikinase</td>
<td>23.65</td>
<td>10</td>
<td>38.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>36.7</td>
</tr>
<tr>
<td>EHI_198760</td>
<td>alcohol dehydrogenase 3, putative</td>
<td>13.9</td>
<td>8</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>38</td>
</tr>
<tr>
<td>EHI_008200</td>
<td>glyceraldehyde-3-phosphate dehydrogenase, putative</td>
<td>14.15</td>
<td>9</td>
<td>45.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>40.4</td>
</tr>
<tr>
<td>EHI_043640</td>
<td>actin, putative</td>
<td>8.27</td>
<td>7</td>
<td>21.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>20.7</td>
</tr>
<tr>
<td>EHI_044970</td>
<td>malic enzyme, putative</td>
<td>16.92</td>
<td>8</td>
<td>42.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>48.1</td>
</tr>
<tr>
<td>EHI_051060</td>
<td>pyruvate: ferredoxin oxidoreductase</td>
<td>11.17</td>
<td>6</td>
<td>18.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>21.1</td>
</tr>
<tr>
<td>EHI_025360</td>
<td>14-3-3 protein 1</td>
<td>4.29</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>EHI_096710</td>
<td>type A flavoprotein, putative</td>
<td>9.05</td>
<td>5</td>
<td>24.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>27.8</td>
</tr>
</tbody>
</table>
Figure 4.2 Venn diagram represents the *E. histolytica* ES proteins identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS.

Three independent biological replicates were performed whereby three and two MS/MS analysis were performed using LC-MALDI-TOF/TOF and LC-ESI-MS/MS, respectively. The data analysis shown above only involved proteins that were identified in at least two of the biological replicates.
4.1.3 Prediction of classical and non-classical secretion

The *E. histolytica* protein sequence database consists of 8306 proteins. As shown in Figure 4.3, the prediction analysis on the *E. histolytica* protein database identified 340 classical and 2328 non-classical secreted proteins. In this study, out of 209 identified ES proteins, 8 and 31 of the proteins were predicted as classically and non-classically secreted proteins, respectively. The secretome prediction results can be found in supplementary material 3.

4.1.4 Functional annotation of ES proteins

The ES proteins identified using the mass spectrometry analyses were classified using PANTHER classification system (Figure 4.4). Out of 209 proteins, the largest number of proteins were attributed to the oxidoreductase class, which comprised of 30 proteins. This was followed by 23 hydrolase proteins, 16 cytoskeletal proteins, 13 transferases, 10 nucleic acid binding proteins, 10 enzyme modulators, 9 lyases, 5 isomerases, 4 ligases, 3 calcium-binding proteins, 2 chaperone proteins, 2 membrane traffic proteins, 1 signaling molecule protein, and 1 transcription factor protein. Meanwhile, out of the 209 *E. histolytica* ES proteins found in this study, 21 and 55 were hypothetical and putative proteins, respectively. The protein functional annotation results can be found in supplementary material 4.
Figure 4.3 Venn diagram represents the prediction analysis of classically and non-classically secreted proteins between the *E. histolytica* protein database and the ES proteins identified by both mass spectrometry systems.
### Figure 4.4

The protein classes of *E. histolytica* excretory-secretory proteins.

<table>
<thead>
<tr>
<th>Protein Class</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signaling molecule (PC00207)</td>
<td>1</td>
</tr>
<tr>
<td>Cytoskeletal protein (PC00085)</td>
<td>16</td>
</tr>
<tr>
<td>Isomerase (PC00135)</td>
<td>5</td>
</tr>
<tr>
<td>Calcium-binding protein (PC00060)</td>
<td>3</td>
</tr>
<tr>
<td>Ligase (PC00142)</td>
<td>4</td>
</tr>
<tr>
<td>Nucleic acid binding (PC00171)</td>
<td>10</td>
</tr>
<tr>
<td>Transcription factor (PC00218)</td>
<td>1</td>
</tr>
<tr>
<td>Transferase (PC00220)</td>
<td>13</td>
</tr>
<tr>
<td>Lyase (PC00144)</td>
<td>9</td>
</tr>
<tr>
<td>Enzyme modulator (PC00095)</td>
<td>10</td>
</tr>
<tr>
<td>Oxidoreductase (PC00176)</td>
<td>30</td>
</tr>
<tr>
<td>Hydrolase (PC00121)</td>
<td>23</td>
</tr>
<tr>
<td>Chaperone (PC00072)</td>
<td>2</td>
</tr>
<tr>
<td>Membrane traffic protein (PC00150)</td>
<td>2</td>
</tr>
</tbody>
</table>
4.2 Membrane proteins

4.2.1 Comparison of protein yield between ProteoExtract® kit, ProteoPrep® kit and conventional method

The ProteoExtract® kit and the conventional extraction method yielded approximately similar amounts of proteins i.e. 416.2 µg and 427.16 µg, respectively. Meanwhile, the protein yield of ProteoPrep® was 230.8 µg, which was lesser by two-fold when compared to the other two methods (Figure 4.5). The results showed significant differences between the protein yields of ProteoExtract® and ProteoPrep® or the conventional method. However, there was no significant difference between the protein yields of ProteoExtract® kit and the conventional method.

4.2.2 SDS-PAGE protein profiles of the three extraction methods

Thirty-five µg of membrane fraction proteins from each extract were separated using SDS-PAGE. As depicted in Figure 4.6, the membrane protein fraction of E. histolytica protein band profiles extracted using the ProteoExtract® kit (lanes 2 and 3) displayed a similar band pattern to the conventional method (lanes 4 and 5). A higher number of protein bands were observed in the samples extracted using the ProteoExtract® kit and the conventional method compared to the ProteoPrep® kit (lanes 6 & 7). In addition, the 20 kDa protein band was more clearly seen in the extracts of ProteoExtract® and conventional method compared to ProteoPrep® kit. Overall, the reproducibility of each extraction methods was reliable as indicated by the consistency of protein band profiles among the replicates.
Figure 4.5  Protein yields from the three extraction methods.

Data expressed as the mean ± SD of three independent batches. The means were 416.2 µg, 230.08 µg and 427.16 µg, while the standard deviations were 85.77 µg, 12 µg, 50.83µg for the ProteoExtract® kit, ProteoPrep® kit and conventional method, respectively. *p<0.05 was considered statistically significant.
The protein profiles of *E. histolytica* membrane fractions extracted using the ProteoExtract® kit, conventional method and ProteoPrep® kit. An amount of 30 µg protein samples were loaded to each well. The samples were denatured in 1× loading buffer and incubated in 100 °C water bath. The SDS-PAGE was ran under the following conditions: 3% stacking gel, 10% resolving gel under 100 V for 90 min. Lane 1: Precision Plus Proteins Kaleidoscope Standards (Bio-rad, USA); Lane 2 and 3: ProteoExtract®; Lane 4 and 5: Conventional method; Lane 6 and 7: ProteoPrep®.
4.2.3 *E. histolytica* protein identification of the membrane fractions extracted using the ProteoExtract® kit, ProteoPrep® kit and conventional method

The representatives of the identified *E. histolytica* membrane protein fractions are listed in Table 4.3 – 4.8. As shown in Figure 4.7A, LC-MALDI-TOF/TOF identified a total of 81 membrane fraction proteins. Meanwhile, LC-ESI-MS/MS identified a total of 456 membrane fraction proteins extracted using the ProteoExtract® kit. A total of 47 proteins were identified by both the mass spectrometry systems. In addition, 34 and 409 proteins were uniquely identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS, respectively.

From the membrane fractions extracted using ProteoPrep® kit (Figure 4.7B), LC-MALDI-TOF/TOF identified 42 proteins whereas LC-ESI-MS/MS identified a total of 482 proteins. Thirty-two proteins were identified by both the mass spectrometry systems. On the other hand, 10 and 450 proteins were uniquely identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS, respectively.

In the membrane fractions extracted using the conventional method (Figure 4.7C), LC-MALDI-TOF/TOF identified 151 proteins and LC-ESI-MS/MS identified 551 proteins from the membrane fractions. Meanwhile, 115 proteins were identified by both the mass spectrometry systems. Furthermore, 36 and 436 proteins were uniquely identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS, respectively.

From the combined data of all three extraction methods, a total of 876 proteins were identified. LC-ESI-MS/MS and LC-MALDI-TOF/TOF identified a total of 828 and 200 proteins, respectively. Meanwhile, LC-ESI-MS/MS and LC-MALDI-TOF/TOF exclusively identified a total of 676 and 48 proteins, respectively. Both the systems identified 152 similar proteins (Figure 4.7D). Overall, LC-ESI-
MS/MS was shown to identify a larger number of proteins than LC-MALDI-TOF/TOF.

Membrane protein prediction analyses using TOPCONS were performed for all the identified membrane fraction proteins. The ProteoPrep® kit was able to extract the highest number of predicted membrane proteins, this was followed by the conventional method and ProteoExtract® kit (Figure 4.8).

This study found a total of 222 proteins that were commonly extracted between the three extraction methods. Totals of 150, 125 and 130 proteins were exclusively extracted using the ProteoExtract® kit, ProteoPrep® kit and conventional method, respectively (Figure 4.9). Membrane protein prediction analyses showed 98 predicted membrane proteins were found from the membrane fractions extracted using the three methods (Figure 4.10). Furthermore, 2, 63 and 11 predicted membrane proteins were exclusively extracted by ProteoExtract®, ProteoPrep® kits and conventional method, respectively. In addition, the membrane protein prediction showed that LC-MALDI-TOF/TOF has only exclusively identified 6 predicted membrane proteins from the list of identified proteins (Table 4.9). In this study, 265 proteins identified exclusively using LC-ESI-MS/MS were predicted as membrane proteins. Details of the overall results can be found in Supplementary material 1 and 5.
Table 4.3  
Examples of proteins extracted using the ProteoExtract® kit and identified by LC-MALDI-TOF/TOF

<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>Replicate 1</th>
<th></th>
<th>Replicate 2</th>
<th></th>
<th>Replicate 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>% Cov (95%)</td>
<td>Peptides (95%)</td>
<td>Total</td>
<td>% Cov (95%)</td>
<td>Peptides (95%)</td>
</tr>
<tr>
<td>EHL_150490</td>
<td>aldehyde-alcohol dehydrogenase 2, putative</td>
<td>12.44</td>
<td>11.49</td>
<td>9</td>
<td>4.7</td>
<td>4.02</td>
<td>2</td>
</tr>
<tr>
<td>EHL_136160</td>
<td>calreticulin, putative</td>
<td>13.72</td>
<td>27.25</td>
<td>7</td>
<td>4.35</td>
<td>11.05</td>
<td>2</td>
</tr>
<tr>
<td>EHL_107290</td>
<td>actin, putative</td>
<td>8</td>
<td>16.76</td>
<td>5</td>
<td>6</td>
<td>13.83</td>
<td>3</td>
</tr>
<tr>
<td>EHL_182900</td>
<td>Actin</td>
<td>8</td>
<td>16.76</td>
<td>5</td>
<td>6</td>
<td>13.83</td>
<td>3</td>
</tr>
<tr>
<td>EHL_130700</td>
<td>enolase, putative</td>
<td>7.85</td>
<td>10.55</td>
<td>4</td>
<td>6</td>
<td>6.19</td>
<td>3</td>
</tr>
<tr>
<td>EHL_178470</td>
<td>hypothetical protein</td>
<td>8</td>
<td>8.94</td>
<td>4</td>
<td>3.58</td>
<td>5.92</td>
<td>2</td>
</tr>
<tr>
<td>EHL_023110</td>
<td>NADP-dependent alcohol dehydrogenase</td>
<td>4</td>
<td>11.20</td>
<td>3</td>
<td>4.02</td>
<td>11.20</td>
<td>2</td>
</tr>
<tr>
<td>EHL_197480</td>
<td>actophorin, putative</td>
<td>6</td>
<td>31.88</td>
<td>3</td>
<td>4</td>
<td>23.19</td>
<td>2</td>
</tr>
<tr>
<td>EHL_201250</td>
<td>Peroxiredoxin</td>
<td>4</td>
<td>13.68</td>
<td>3</td>
<td>1.44</td>
<td>10.68</td>
<td>1</td>
</tr>
<tr>
<td>EHL_004490</td>
<td>thioredoxin, putative</td>
<td>3.36</td>
<td>30.10</td>
<td>2</td>
<td>3.72</td>
<td>30.10</td>
<td>2</td>
</tr>
</tbody>
</table>

"#N/A – data not available
Table 4.4  Examples of proteins extracted using the ProteoExtract® kit and identified by LC-ESI-MS/MS

<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>Replicate 1</th>
<th>Replicate 2</th>
<th>Replicate 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sum PEP Score</td>
<td>% Cov Peptide</td>
<td>Sum PEP Score</td>
</tr>
<tr>
<td>EHI_110180</td>
<td>myosin heavy chain</td>
<td>188.54</td>
<td>25.76</td>
<td>325.10</td>
</tr>
<tr>
<td>EHI_009530</td>
<td>pyruvate phosphate dikinase</td>
<td>200.84</td>
<td>54.24</td>
<td>224.45</td>
</tr>
<tr>
<td>EHI_201510</td>
<td>clathrin heavy chain, putative</td>
<td>120.03</td>
<td>29.68</td>
<td>223.88</td>
</tr>
<tr>
<td>EHI_006980</td>
<td>Gal/GalNAc lectin subunit Ig1</td>
<td>201.43</td>
<td>43.42</td>
<td>751.60</td>
</tr>
<tr>
<td>EHI_044970</td>
<td>malic enzyme, putative</td>
<td>163.36</td>
<td>68.99</td>
<td>162.00</td>
</tr>
<tr>
<td>EHI_030750</td>
<td>hypothetical protein</td>
<td>125.38</td>
<td>28.97</td>
<td>66.02</td>
</tr>
<tr>
<td>EHI_198620</td>
<td>hypothetical protein</td>
<td>117.65</td>
<td>27.93</td>
<td>62.44</td>
</tr>
<tr>
<td>EHI_014030</td>
<td>NAD(P) transhydrogenase subunit alpha,</td>
<td>88.57</td>
<td>25.21</td>
<td>486.78</td>
</tr>
<tr>
<td></td>
<td>putative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EHI_052860</td>
<td>heat shock protein 70, putative</td>
<td>131.04</td>
<td>36.89</td>
<td>158.00</td>
</tr>
<tr>
<td>EHI_101230</td>
<td>P-glyco protein 6, putative</td>
<td>113.91</td>
<td>25.66</td>
<td>283.66</td>
</tr>
</tbody>
</table>

"#N/A – data not available
Table 4.5  Examples of proteins extracted using ProteoPrep® kit and identified by LC-MALDI-TOF/TOF

<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>Replicate 1</th>
<th></th>
<th></th>
<th>Replicate 2</th>
<th></th>
<th></th>
<th>Replicate 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>% Cov (95%)</td>
<td>Peptides (95%)</td>
<td>Total</td>
<td>% Cov (95%)</td>
<td>Peptides (95%)</td>
<td>Total</td>
<td>% Cov (95%)</td>
</tr>
<tr>
<td>EHI_055400</td>
<td>NAD(P)( ) transhydrogenase (AB-specific), alpha subunit, putative</td>
<td>26.7</td>
<td>18.31</td>
<td>15</td>
<td>6</td>
<td>5.01</td>
<td>4</td>
<td>4</td>
<td>2.73</td>
</tr>
<tr>
<td>EHI_012270</td>
<td>Gal/GalNAc lectin heavy subunit</td>
<td>25.82</td>
<td>15.09</td>
<td>14</td>
<td>4</td>
<td>1.55</td>
<td>2</td>
<td>8.32</td>
<td>5.60</td>
</tr>
<tr>
<td>EHI_133900</td>
<td>galactose-inhibitable lectin 170 kDa subunit, putative</td>
<td>26.31</td>
<td>14.46</td>
<td>14</td>
<td>4</td>
<td>1.55</td>
<td>2</td>
<td>8.32</td>
<td>5.26</td>
</tr>
<tr>
<td>EHI_049690</td>
<td>galactose-specific adhesin light subunit, putative</td>
<td>8</td>
<td>20.83</td>
<td>9</td>
<td>2</td>
<td>4.86</td>
<td>4</td>
<td>4.03</td>
<td>10.42</td>
</tr>
<tr>
<td>EHI_006980</td>
<td>Gal/GalNAc lectin subunit Ig1</td>
<td>8.59</td>
<td>8.08</td>
<td>8</td>
<td>4</td>
<td>2.27</td>
<td>2</td>
<td>6.14</td>
<td>4.36</td>
</tr>
<tr>
<td>EHI_035690</td>
<td>galactose-inhibitable lectin 35 kDa subunit precursor</td>
<td>6.16</td>
<td>19.10</td>
<td>8</td>
<td>2</td>
<td>4.86</td>
<td>4</td>
<td>#N/A</td>
<td>#N/A</td>
</tr>
<tr>
<td>EHI_178470</td>
<td>hypothetical protein</td>
<td>15.74</td>
<td>14.08</td>
<td>8</td>
<td>4</td>
<td>4.80</td>
<td>2</td>
<td>#N/A</td>
<td>#N/A</td>
</tr>
<tr>
<td>EHI_158240</td>
<td>fatty acid elongase, putative</td>
<td>15.01</td>
<td>17.98</td>
<td>7</td>
<td>2</td>
<td>3.43</td>
<td>1</td>
<td>2</td>
<td>3.43</td>
</tr>
<tr>
<td>EHI_047800</td>
<td>hypothetical protein</td>
<td>10</td>
<td>22.37</td>
<td>5</td>
<td>4</td>
<td>10.53</td>
<td>2</td>
<td>4.14</td>
<td>8.22</td>
</tr>
<tr>
<td>EHI_192810</td>
<td>small GTPase Rab7A</td>
<td>5.34</td>
<td>20.39</td>
<td>3</td>
<td>4</td>
<td>19.90</td>
<td>2</td>
<td>#N/A</td>
<td>#N/A</td>
</tr>
</tbody>
</table>

aN/A – data not available
<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>Replicate 1</th>
<th></th>
<th>Replicate 2</th>
<th></th>
<th>Replicate 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EHI_101230</td>
<td>P-glyco protein 6, putative</td>
<td>194.61</td>
<td>32.84</td>
<td>32</td>
<td></td>
<td>314.57</td>
<td>42.59</td>
</tr>
<tr>
<td>EHI_054830</td>
<td>Plasma membrane calcium-transporting ATPase, putative</td>
<td>141.72</td>
<td>34.68</td>
<td>31</td>
<td></td>
<td>205.85</td>
<td>39.10</td>
</tr>
<tr>
<td>EHI_100320</td>
<td>multidrug resistance protein, putative</td>
<td>147.19</td>
<td>24.94</td>
<td>30</td>
<td></td>
<td>192.68</td>
<td>28.13</td>
</tr>
<tr>
<td>EHI_015380</td>
<td>immuno-dominant variable surface antigen</td>
<td>152.98</td>
<td>40.61</td>
<td>29</td>
<td></td>
<td>193.07</td>
<td>43.49</td>
</tr>
<tr>
<td>EHI_030830</td>
<td>Plasma membrane calcium-transporting ATPase, putative</td>
<td>124.21</td>
<td>34.11</td>
<td>23</td>
<td></td>
<td>198.62</td>
<td>40.16</td>
</tr>
<tr>
<td>EHI_199590</td>
<td>70 kDa heat shock protein, putative</td>
<td>159.79</td>
<td>43.12</td>
<td>22</td>
<td></td>
<td>245.26</td>
<td>54.92</td>
</tr>
<tr>
<td>EHI_147010</td>
<td>long-chain-fatty-acid--CoA ligase, putative</td>
<td>110.60</td>
<td>43.26</td>
<td>22</td>
<td></td>
<td>121.80</td>
<td>43.11</td>
</tr>
<tr>
<td>EHI_183510</td>
<td>nuclear pore protein, putative</td>
<td>85.67</td>
<td>18.17</td>
<td>22</td>
<td></td>
<td>77.80</td>
<td>13.64</td>
</tr>
<tr>
<td>EHI_035730</td>
<td>hypothetical protein</td>
<td>112.90</td>
<td>38.44</td>
<td>21</td>
<td></td>
<td>138.92</td>
<td>42.42</td>
</tr>
<tr>
<td>EHI_111990</td>
<td>CXXC-rich protein</td>
<td>102.55</td>
<td>26.12</td>
<td>21</td>
<td></td>
<td>167.03</td>
<td>31.04</td>
</tr>
</tbody>
</table>

“#N/A – data not available
Table 4.7  Examples of proteins extracted using the conventional method and identified by LC-MALDI-TOF/TOF

<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>Replicate 1</th>
<th>Replicate 2</th>
<th>Replicate 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>% Cov  (95%)</td>
<td>Peptides (95%)</td>
</tr>
<tr>
<td>EHI_043640</td>
<td>actin, putative</td>
<td>12</td>
<td>24.70</td>
<td>7</td>
</tr>
<tr>
<td>EHI_065330</td>
<td>Gal/GalNAc lectin subunit Igl2</td>
<td>10</td>
<td>7.60</td>
<td>7</td>
</tr>
<tr>
<td>EHI_096830</td>
<td>glycogen phosphorylase, putative</td>
<td>11.3</td>
<td>12.67</td>
<td>7</td>
</tr>
<tr>
<td>EHI_107290</td>
<td>actin, putative</td>
<td>12</td>
<td>21.81</td>
<td>7</td>
</tr>
<tr>
<td>EHI_136160</td>
<td>calreticulin, putative</td>
<td>9.74</td>
<td>18.77</td>
<td>7</td>
</tr>
<tr>
<td>EHI_049690</td>
<td>galactose-specific adhesin light subunit, putative</td>
<td>6.08</td>
<td>15.63</td>
<td>6</td>
</tr>
<tr>
<td>EHI_164800</td>
<td>hypothetical protein</td>
<td>10.67</td>
<td>11.95</td>
<td>6</td>
</tr>
<tr>
<td>EHI_033710</td>
<td>cysteine proteinase 2</td>
<td>6</td>
<td>21.27</td>
<td>5</td>
</tr>
<tr>
<td>EHI_060740</td>
<td>EF-hand calcium-binding domain containing protein</td>
<td>8</td>
<td>25.11</td>
<td>5</td>
</tr>
</tbody>
</table>

"#N/A – data not available"
Table 4.8  Examples of proteins extracted using the conventional method and identified by LC ESI-MS/MS

<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>Replicate 1</th>
<th>Replicate 2</th>
<th>Replicate 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sum PEP</td>
<td>% Cov</td>
<td># Peptides</td>
</tr>
<tr>
<td>EHI_150490</td>
<td>aldehyde-alcohol dehydrogenase 2, putative</td>
<td>418.872</td>
<td>71.149</td>
<td>47</td>
</tr>
<tr>
<td>EHI_101230</td>
<td>P-glyco protein 6, putative</td>
<td>260.911</td>
<td>37.051</td>
<td>39</td>
</tr>
<tr>
<td>EHI_014030</td>
<td>NAD(P) transhydrogenase subunit alpha, putative</td>
<td>238.774</td>
<td>40.905</td>
<td>38</td>
</tr>
<tr>
<td>EHI_148560</td>
<td>4-alpha-glucanotransferase, putative NAD(P)( ) transhydrogenase (AB-</td>
<td>208.846</td>
<td>53.341</td>
<td>38</td>
</tr>
<tr>
<td>EHI_055400</td>
<td>specific), alpha subunit, putative</td>
<td>206.228</td>
<td>37.432</td>
<td>34</td>
</tr>
<tr>
<td>EHI_015380</td>
<td>immuno-dominant variable surface antigen</td>
<td>176.453</td>
<td>45.553</td>
<td>32</td>
</tr>
<tr>
<td>EHI_004640</td>
<td>vesicle-fusing ATPase, putative</td>
<td>130.532</td>
<td>57.201</td>
<td>31</td>
</tr>
<tr>
<td>EHI_199590</td>
<td>70 kDa heat shock protein, putative</td>
<td>209.479</td>
<td>60.968</td>
<td>30</td>
</tr>
<tr>
<td>EHI_051060</td>
<td>pyruvate:ferredoxin oxidoreductase</td>
<td>128.906</td>
<td>34.079</td>
<td>29</td>
</tr>
<tr>
<td>EHI_030830</td>
<td>Plasma membrane calcium-transporting ATPase, putative</td>
<td>133.295</td>
<td>38.694</td>
<td>27</td>
</tr>
</tbody>
</table>
Figure 4.7  The Venn diagrams represent the number of proteins identified by LC-MALDI-TOF/TOF and LC-ESI-MS/MS from the membrane fractions

(A) ProteoExtract®, (B) ProteoPrep® and (C) conventional method.
(D) The Venn diagram depicts the combined number of proteins identified in all three extraction methods by LC-MALDI-TOF/TOF and LC-ESI-MS/MS.
Figure 4.8 The number of predicted membrane proteins versus non-membrane proteins identified from the membrane fractions using the three extraction methods.
Figure 4.9 A combination of the identified *E. histolytica* membrane fraction proteins extracted by the three extraction methods.

Three independent biological replicates were performed for each of the membrane extraction methods. The data analysis shown above only involved proteins that were identified in at least two of the biological replicates.
Figure 4.10  The Venn diagram represents the predicted membrane proteins identified in at least two replicates from the membrane fractions of the three extraction methods.

As mentioned previously in Figure 4.9, the data analysis shown above only involved proteins that were identified in at least two of the biological replicates. Subsequently, the proteins identified were subjected to transmembrane protein topology prediction through a webserver at www.topcons.net.
**Table 4.9**  List of all the predicted membrane proteins exclusively identified by LC-MALDI-TOF/TOF

<table>
<thead>
<tr>
<th>No.</th>
<th>Length</th>
<th>No. TM</th>
<th>Accession no.</th>
<th>Sequence name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>96</td>
<td>3</td>
<td>EHI_040820</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2</td>
<td>160</td>
<td>4</td>
<td>EHI_042580</td>
<td>dolychil-diphosphooligosaccharide-- protein glycosyltransferase subunit, putative</td>
</tr>
<tr>
<td>3</td>
<td>409</td>
<td>5</td>
<td>EHI_060020</td>
<td>NAD(P) transhydrogenase beta subunit, putative</td>
</tr>
<tr>
<td>4</td>
<td>384</td>
<td>10</td>
<td>EHI_103330</td>
<td>mannosyltransferase, putative</td>
</tr>
<tr>
<td>5</td>
<td>265</td>
<td>5</td>
<td>EHI_153260</td>
<td>hypothetical protein, conserved</td>
</tr>
<tr>
<td>6</td>
<td>93</td>
<td>2</td>
<td>EHI_200720</td>
<td>signal peptidase complex subunit, putative</td>
</tr>
</tbody>
</table>
4.2.4 Evaluation of the membrane extraction methods selectivity by comparing the identified membrane fraction and cytosolic fraction proteins

To properly assess the selectivity and specificity of the extraction methods, the corresponding left-over fractions known as the cytosolic fractions were analysed. In this study, the LC-ESI-MS/MS identified more number of proteins compared to LC-MALDI-TOF/TOF. Hence, the respective cytosolic fractions were analysed only by LC-ESI-MS/MS. Subsequently, only the corresponding membrane fraction results from LC-ESI-MS/MS were used for the selectivity and specificity analysis. The mass spectrometry results for proteins identified in the cytosolic fractions can be found in supplementary material 6.

LC-ESI-MS/MS identified a total of 456, 482 and 551 proteins from the membrane fractions extracted using the ProteoExtract® kit, ProteoPrep® kit and conventional method, respectively. Meanwhile, LC-ESI-MS/MS identified a total of 487, 611 and 343 proteins from the cytosolic fractions extracted using the ProteoExtract® kit, ProteoPrep® kit and conventional method, respectively. (Figure 4.11). The results for the membrane and cytosolic fraction revealed that more number of proteins were detected in the cytosolic fractions compared to the membrane fractions extracted using the ProteoExtract® kit and the ProteoPrep® kit. Meanwhile, for the conventional method, more number of proteins were identified from the membrane fractions than the cytosolic fractions.

To assess the selectivity of the membrane protein extraction methods, the identified proteins from the cytosolic and membrane fractions were compared. The selectivity of the extraction methods was indicated by the overlapped regions between the membrane and the cytosolic fractions (Figure 4.12). The method with
the least overlapped region between the membrane and cytosolic fraction was extracted using the ProteoPrep® kit, followed by the conventional method and the ProteoExtract® kit. This indicated that the ProteoPrep® kit was the most efficient in separating the membrane from the cytosolic fraction compared to the other two methods. Details of the result can be found in Supplementary material 1.
Figure 4.11  The number of proteins identified from the membrane fractions and the cytosolic fractions of the ProteoExtract® kit, ProteoPrep® kit and conventional method.
Figure 4.12  The Venn diagrams compare the proteins identified using LC-ESI-MS/MS from the membrane and cytosolic fractions.
4.2.5 Evaluation of the membrane extraction methods specificity by comparing the predicted membrane proteins among the identified membrane and cytosolic fraction proteins

As mentioned in chapter 4.2.4, membrane protein extraction specificity analysis was based on the proteins identified by LC-ESI-MS/MS. As such, Figure 4.13 depicts a summary of the number of membrane and cytosolic fraction proteins identified using the LC-ESI-MS/MS only. In addition, Figure 4.13 also shows that the cytosolic fractions extracted using the ProteoPrep® kit contained the least percentage of predicted membrane proteins. On the other hand, the conventional method extracted the least number of predicted membrane proteins from the cytosolic fraction. The results for membrane protein prediction can be found in supplementary material 7.

Subsequently, a Venn diagram was constructed to analyse the differences between the predicted membrane proteins found in the membrane fractions (MF) and cytosolic fractions (CF). As shown in Figure 4.14, the overlapped region between the MF and CF of ProteoExtract® kit, ProteoPrep® kit, and conventional method were 14%, 3.3% and 1.6%, respectively. The predicted membrane proteins were exclusively identified from the cytosolic fraction with only 6%, 1.7% and 2.8% by the ProteoExtract® kit, ProteoPrep® kit and conventional method, respectively. This is important because a low percentage of predicted membrane proteins found in the CF indicates high specificity of the membrane protein extraction method. Overall, the ProteoPrep® kit showed the highest specificity in the extraction of *E. histolytica* membrane proteins.
Figure 4.13  The number of predicted membrane proteins versus non-membrane proteins identified from the membrane fraction (MF) and cytosolic fraction (CF) of the three extraction methods.
Figure 4.14  The Venn diagrams compare the predicted membrane proteins from the membrane and cytosolic fractions.
4.2.6 Functional annotation of membrane proteins

All the proteins identified in the membrane fractions extracted using all three membrane protein extraction methods were classified according to their protein classes (Figure 4.15). The most represented protein class consisting of 84 proteins were attributed to nucleic acid binding protein. This was followed by 69 hydrolases, 44 enzyme modulators, 43 membrane traffic proteins, 38 transferases, 37 oxidoreductases, 28 cytoskeletal proteins, 28 transporter proteins, 17 ligases, 9 receptor proteins, 8 calcium-binding proteins, 8 transfer/carrier proteins, 7 lyases, 5 extracellular matrix proteins, 4 transcription factor proteins, 3 signalling molecule proteins, 3 chaperone proteins, 3 isomerase, 2 cell adhesion molecule proteins, 1 transmembrane receptor regulatory/adaptor protein, and 1 protein involved in the cell junction. Whereas for cellular compartment annotation, a total of 306 proteins were a part of cell part (Figure 4.16). This was followed by 252, 146, 129, and 6 proteins associated to the organelle, macromolecular complex, membrane, and extracellular region, respectively. Meanwhile, 10 out of the 876 proteins remained unclassified. Furthermore, out of the 8306 E. histolytica proteins that were deposited into the AmoebaDB protein database, the TOPCONS predicted 1352 (16%) membrane proteins. This study has only identified 258 (19%) predicted membrane proteins from the predicted membrane proteins from the database (supporting material 1). The functional annotation results can be found in supplementary material 8.
Figure 4.15  The proteins classes of *E. histolytica* membrane fraction proteins.
Figure 4.16  The *E. histolytica* membrane fraction proteins that are classified according to the cellular compartment.
CHAPTER 5 – DISCUSSION

5.1 The excretory-secretory (ES) proteins of *E. histolytica*

5.1.1 Isolation of the ES proteins

Challenges in analysing ES proteomes of *E. histolytica* trophozoites include overcoming contamination of proteins from components of the culture medium, sustaining cell viability throughout the incubation period to avoid contamination of non-excreted-secreted proteins, overcoming the high proteolytic activity in *E. histolytica* cells and overcoming the low abundance of excreted-secreted proteins (Flores et al., 2005, Chevallet et al., 2007). Several strategies have been applied to address these challenges in order to maximise the proteome coverage.

The isolation of *E. histolytica* ES proteins without contamination of non-*E. histolytica* proteins were crucial. Therefore, in this study, protein-free medium such as RPMI-C-A was used instead of culture medium containing serum proteins. Wong et al. (2011) reported that the RPMI-C-A was found to be most suitable protein-free medium with at least 95% cell viability in over 8 hours incubation. In order to minimise the carryover of serum proteins from the culture medium, washing steps were performed thoroughly with the protein-free media. Furthermore, in order to isolate pure ES proteins, seeding process was only carried out using viable trophozoites that showed healthy characteristics such as the absence of contaminants, good amoeboid movements and attachment to the tube surface. In addition, only culture flasks that met at least 95% cell viability was used for the protein isolation procedure.
One of the crucial ways to capture the low abundance ES proteins was to extend the collection time before cell lysis occurs. Hence, a chemically-defined media was used, whereby the addition of 0.1 % of L-cysteine and 0.02 % ascorbic acid in the medium have been reported to extend the cell viability duration of the *E. histolytica* trophozoites (Gillin and Diamond, 1980). The addition of these supplements was reported to act as a reducing agent that protected the *E. histolytica* trophozoites against oxidative stress. The minimal cell lysis using the modified RPMI medium was proven using Western blotting analysis where the crude soluble antigen (CSA) and ESA proteins were probed with patient sera showed significant differences in the protein band patterns (Sengupta et al., 2000, Wong et al., 2011). Therefore, the isolation method in this study was efficient in isolating true ES proteins. On the other hand, *E. histolytica* trophozoites express high amount of proteases, including cysteine protease which can cause the digestion of the ES proteins (Que and Reed, 2000, Sengupta et al., 2000, Flores et al., 2005). In order to preserve the isolated ES proteins, protease inhibitor cocktail and iodoacetamide were added to inhibit the cysteine protease activity in this study.

There are two well-known methods to capture proteins in low concentration or diluted extracts. The methods include protein concentration step by either precipitation or ultrafiltration (Jiang et al., 2004). The former requires precipitation of proteins using organic solvents like acetone while the latter requires specialised ultrafiltration tubes such as the Vivaspin (GE Healthcare, UK). Both methods have disadvantages, the precipitation method requires a minimum of four-fold volume of solvent to sample volume, hence it is more suitable for small sample volumes. Meanwhile, samples subjected to ultrafiltration method is prone to clogging of the semi-permeable membrane (Fic et al., 2010, Eppler et al., 2011). In this study,
ultrafiltration method was applied to increase the concentration of the isolated ES proteins due to the high initial volume.

5.1.2 ES proteins identification and prediction of the classical and non-classical secretion pathways

This study revealed 291 ES proteins by the combination of LC-MALDI-TOF/TOF and LC-ESI-MS/MS analyses. From this protein list, 52 proteins (25%) were well-annotated proteins. However, 136 proteins (65%) and 21 proteins (10%) remained putative and hypothetical, respectively. This showed that a large percentage of *E. histolytica* ES proteins have not been well characterised. In a study by Ahn et al. (2018), a total of 79 individual *E. histolytica* ES proteins were separated and identified using 2D-gel electrophoresis and MALDI-MS/MS. Their study found less number of proteins compared to the present study. However, the secretome profile reported in their study and the present study generally matched well.

In the classical secretion system, proteins are often processed and packaged through the endoplasmic reticulum and Golgi apparatus. Then, secretory vesicles are formed and subsequently fused to the cell membrane before being released to the extracellular environment. These proteins contain the N-terminal signal peptide. On the other hand, proteins that are secreted without the N-terminal signal peptide is known as the leaderless secretion or the non-classical secretory pathway (Kuchler et al., 2013). The non-classical secretory system works independently from the ER–Golgi network. These secreted proteins do not undergo glycosylation even with the presence of glycosylation motifs.
In this study, only a small percentage of the identified ES proteins contained either the N-terminal signal peptide or the non-classical leaderless sequence. This suggested that a large portion of the identified proteins might not be a part of a secretion process, and may be the by-products generated during metabolic turnovers. In addition, the identified proteins may have originated from the membrane surface and were released during the membrane shedding process. This finding concurred with previous studies that demonstrated the lack of signal peptides in the *E. histolytica* secretory proteins (Biller et al., 2014, Ahn et al., 2018). Alternatively, these *E. histolytica* ES proteins that lack either signal peptides or non-leader sequence may have been excreted through the extracellular vesicles, as observed in the metazoan parasites such as *Echinostoma caproni* (Marcilla et al., 2012).

This study has identified a number of proteins that are considered as the surface membrane proteins which may be excreted during the membrane shedding process. More than half of the proteins that were identified in this study were similar to the *E. histolytica* membrane surface proteome reported by Biller et al. (2014) such as disulphide isomerase, enolase, heat shock proteins, malate dehydrogenase, triosephosphate isomerase, thioredoxin and superoxide dismutase. The identification of membrane-related proteins in the ES proteins may be due to membrane recycling event. For example, *Acanthamoeba* proteins were shown to be secreted when trapped in the shuttle vesicles that were transported from secondary lysosomes to the surface membrane (Hohman and Bowers, 1984).
5.1.3 Functional protein classification

Among the identified ES proteins, the most represented protein class was the oxidoreductase (PC00176) which consisted of the dehydrogenase (PC00092), oxidase (PC00175), peroxidase (PC00180) and reductase (PC00198). One of the well-represented protein family found under this category was the peroxiredoxin (EHI_139570, EHI_145840, EHI_121620, EHI_122310, EHI_123390, EHI_061980, EHI_001420, EHI_201250, and EHI_114010). According to the annotated function, peroxiredoxin is involved in redox metabolism, oxidoreductase and peroxidase activities. In addition, this study has found peroxiredoxin to be non-classically secreted. As previously mentioned in the literature review, the presence of these antioxidant enzymes in the extracellular environment aid in the defence mechanism of the parasite against reactive oxygen species that is imposed by the host immune system (Choi et al., 2005).

Another oxidoreductase protein is the iron-containing superoxide dismutase (EHI_159160). This protein was reported to play important roles in drug resistance as it was found highly expressed in metronidazole resistant *E. histolytica* trophozoites (Wassmann et al., 1999). Interestingly in this study, the superoxide dismutase was found in all the biological replicates, however, it was not classified under both the classically and non-classically secreted proteins. Furthermore, it contained no predicted transmembrane domain. Similar to peroxiredoxin, the superoxide dismutase mediates detoxification of superoxide radicals imposed by the host’s immune system (Tannich et al., 1991). This protein is known to be only effective against superoxide radicals in the intracellular environment of the parasite (Hiraishi et al., 1994). Hence, an important question arises on how an intracellular protein was found in the extracellular environment if not because of cell lysis? Based
on Escuyer et al. and Kang et al., the secretion of superoxide dismutase has been reported to occur in a few virulent mycobacteria, including *Mycobacterium tuberculosis* as well as the virulent form of *Nocardia asteroides* (Escuyer et al., 1996, Beaman et al., 1985, Kang et al., 1998). To date, the mechanism of extracellular secretion of the superoxide dismutase is still not fully understood. Several studies have reported that the extracellular superoxide dismutase consists of CuZn-superoxide dismutase and Mn-superoxide dismutase, however, no study has been reported on the secretion of iron-containing superoxide dismutase (Nozik-Grayck et al., 2005).

Hydrolase (PC00121) is the second most represented protein class in the ES proteins identified in this study. One of the *E. histolytica* ES proteins annotated under this category was the cysteine proteinase (EHI_033710, EHI_168240, EHI_050570 and EHI_074180). As mentioned in the literature review, the cysteine proteinase plays an important role in the pathogenesis and the virulence of the *E. histolytica* trophozoites (Lidell et al., 2006). The roles of cysteine proteinase include the degradation of fibrinogen, collagen and the basement membrane matrix (Keene et al., 1986, Scholze and Werries, 1986, Reed et al., 1993). This study showed that the cysteine proteinase 2 (EHI_033710) contained N-terminal signal peptide, hence it is a classically secreted protein. Based on the studies of Hellberg et al. (2001) and Lidell et al. (2006), when this enzyme was overexpressed by the *E. histolytica* trophozoites in the extracellular environment, it caused damage to the cysteine-rich MUC2 polymer of the host tissue.

Only one identified ES protein was classified under the signalling molecule category which is known as Rho GDP exchange inhibitor (EHI_147570). This protein is a signalling molecule as indicated by the functional annotation. However,
in this study, this protein was not predicted as classically and non-classically secreted protein. The Rho GDP exchange inhibitor (Rho-GDI) extracts Rho family GTPases from membranes and solubilises them in the cytosol and also inhibit GTP hydrolysing activities on Rho proteins (Fauré and Dagher, 2001). This protein regulates the Rho signalling pathway, which is responsible for various cellular regulations including morphology, migration, phagocytosis, vesicular transport, as well as the assembly and contraction of stress fibres (Etienne-Manneville and Hall, 2002). Furthermore, the activation of this signalling pathway in E. histolytica was reported to be up-regulated during an invasive behaviour (Franco-Barraza et al., 2006). A study on the knock-out of this gene in Dictyostelium has led to the defects in cytokinesis, pinocytosis and the contractile vacuole system (Imai et al., 2002).

The Gal/GalNAc lectin-related proteins (EHI_006980, EHI_065330 and EHI_012270) were not annotated in any functional classification by PantherDB. However, these proteins were identified in all three biological replicates. The GalNAc lectin proteins are well studied and their potential has already been explored and exploited (Wong et al., 2011, Watanabe and Petri, 2015). These proteins matched the criteria for classically and non-classically secreted proteins. Furthermore, the Gal/GalNAc lectin is an adhesin that is needed for the attachment of the parasite to the host. This was observed when the parasite failed to engage adhesion and contact-dependent cytotoxicity with cells that lacked terminal Gal/GalNAc residues (Ravdin and Guerrant, 1981, Tavares et al., 2005). They contributed to the virulence of the parasite by aiding in the adherence to the host tissue, cause contact-dependent cytolysis to target cells, involve in phagocytosis, and contribute to the resistance of cell lysis that is inflicted by the host (Petri Jr et al., 2002). GalNAc lectin protein was also reported by Wong et al. (2011) as one of the
E. histolytica ES components. The abundance of Gal/GalNAc lectin proteins on the surface of the parasite and its antigenic property have allowed it to be used for antigen detection assay namely the TechLab Entamoeba histolytica II ELISA kit (TechLAB Inc.) (Haque et al., 2000).

5.2 The membrane proteins of E. histolytica

5.2.1 Comparison of membrane protein extraction between ProteoExtract® kit, ProteoPrep® kit and conventional method

Selectivity and specificity are the most important characteristics to consider in choosing a subcellular extraction method. Nonetheless, additional factors such as protein yield, cost, time and the impact on the downstream analyses may influence the choice of a method. In particular, it is very challenging to perform large-scale analysis of membrane proteins. This is due to the fact that the membrane proteins are generally hydrophobic and insoluble in nature (Santoni et al., 2000).

Membrane proteins extracted using the ProteoExtract® kit and conventional method yielded similar amount of proteins, whereas the ProteoPrep® kit yield was lesser by two-folds (Figure 4.5 and 4.6). The ProteoExtract® kit and the conventional method minimised undesired protein loss by allowing the procedures to be performed in the same tube. On the other hand, the ProteoPrep® method required separate tubes between incubation and ultracentrifugation steps. The low protein amount yielded by the ProteoPrep® kit may be due to the inevitable sample loss through the multiple contact points such as during the sample transfer between multiple tubes for sonication, incubation, and ultracentrifugation steps. In contrast, the conventional method omitted sample transfers between multiple tubes and required only a single contact with the sonicator probe. Similarly, the high protein yield by the
ProteoExtract® kit was also probably achieved due to minimal contact points as the protocol was performed in a single tube. A summary of the protein yield, time, cost and the use of ultracentrifuge are listed in Table 5.1.

Table 5.1 Comparison of the protein yield, time, cost and the need of ultracentrifugation of each method

<table>
<thead>
<tr>
<th></th>
<th>ProteoExtract®</th>
<th>ProteoPrep®</th>
<th>Conventional method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein yield (µg)</td>
<td>358.8 - 514.8</td>
<td>217.3 - 241.1</td>
<td>374.4 - 475.8</td>
</tr>
<tr>
<td>Cost per extraction</td>
<td>RM93</td>
<td>RM260</td>
<td>RM6</td>
</tr>
<tr>
<td>Time (h)</td>
<td>1.5 - 2</td>
<td>3.5 - 4</td>
<td>2 - 2.5</td>
</tr>
<tr>
<td>Ultracentrifugation</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Using mammalian cells, a similar observation was reported by Bünger et al. (2009) whereby the ProteoExtract® kit gave the highest protein yield when compared to three other undisclosed methods. In an attempt to identify the best protocol to extract membrane proteins, they have demonstrated that the use of the ProteoExtract® extraction kit could extract 0.3-1.3 mg of membrane protein per 8-10⁷ colorectal carcinoma cell line SW620.

5.2.2 Assessment of method’s selectivity and specificity

The number of identified proteins may not be coherent with the results from the protein yield assay and the SDS-PAGE analysis, despite the standardisation of the initial protein amount prior to mass spectrometry analysis (Section 4.2). For
example, the ProteoExtract® kit and the conventional method extracted high protein yields but had a different total number of identified proteins, whereby the ProteoExtract® kit had the lowest number of proteins identified compared to the conventional method. Although the SDS-PAGE protein band profiles and the protein yield for the membrane fractions extracted using the ProteoPrep® kit resulted in a lower number of protein bands and yield, the number of identified proteins were considerably high compared to the ProteoExtract® kit.

The selectivity analysis of the membrane and cytosolic fractions extracted using the ProteoPrep® kit showed the lowest percentage of cross-contamination between the cytosolic and the membrane fraction. This was followed by the conventional method which was indicated by 17% contamination. On the other hand, the ProteoExtract® kit revealed the highest percentage of proteins crossover between the two fractions at 39.5%. The relatively high percentage of cytosolic protein contamination in the membrane fraction extracted using the ProteoExtract® kit indicated a low-quality separation between the two subcellular fractions. However, a study by Bünger et al. (2009) reported that the ProteoExtract® kit was ideal for the extraction of membrane proteins among four other undisclosed commercial membrane protein extraction kits, whereby, they have reported excellent separation with minimal cross-contamination between the membrane and the cytosolic fractions. However, their study may not be comprehensive enough as the selectivity analysis between the subcellular fractions were based on Western blotting, using only three membrane protein markers and two cytosolic protein markers.

In the present study, the ProteoPrep® extraction kit exhibited the highest number of predicted membrane proteins (254 proteins) though the number of identified proteins was less than the other two methods. This may be due to the use
of differential and sequential extraction strategies which resulted in the enrichment of membrane proteins. This was performed by partitioning the membrane proteins to allow differential solubilisation of cytosolic and membrane proteins. After the removal of the cytosolic proteins, the resultant membrane pellet was washed and partially solubilised, thus resulting in a reduced number of carryover soluble proteins compared to a single step solubilisation strategy (Molloy et al., 1998). Although the ProteoExtract® kit extracted high protein yield, the extraction buffers were meant for mild differential extraction, and membrane proteins were extracted based on their association with the cellular membranes instead of their intrinsic hydrophobicity (ProteoExtract® Native Membrane Protein Extraction Kit, 2010, Zhang et al., 2007). Hence, the number of identified proteins and predicted membrane proteins by the ProteoExtract® kit were lower than the conventional method since the majority of hydrophobic proteins were not extracted by the kit.

Meanwhile, the membrane extract from the conventional method gave the highest number of the protein identification with 35% of the proteins were predicted to be membrane proteins. However, the use of high pH phosphate buffer extracted more hydrophilic proteins than hydrophobic membrane proteins (Shevchenko et al., 2012). The buffer also allowed the disruption of the non-covalent interaction of peripheral membrane proteins but was inefficient in solubilising integral membrane proteins (Speers and Wu, 2007, Shevchenko et al., 2012). In addition, the sonication step was used to open the membrane vesicles, thus allowing the release of soluble and membrane-associated proteins (Speers and Wu, 2007).
5.2.3 Functional annotation

The topmost abundant protein class was the nucleic-acid binding proteins (PC00171) which include sub-classes such as the DNA binding proteins (PC00009), RNA binding protein (PC00031) and nuclease (PC00170) (Figure 4.15). Among the most abundant sub-class in this category was the ribosomal proteins (PC00202) and consisted of 58 proteins, for instance, the 60s acidic ribosomal proteins (EHL_010650, EHL_138770, EHL_052610, EHL_141940, EHL_103310 etc.). To date, only two studies have reported on these proteins in *E. histolytica* (del Pilar Crisóstomo-Vázquez et al., 2014, Shahi et al., 2016). These proteins are strong acidic proteins that are commonly found on the surface of the ribosome of all organisms (Tchórzewski, 2002). Eukaryotic acidic ribosomal proteins are a structural constituent of the ribosome and are involved in nucleic acid binding and translation (Zinker and Warner, 1976). Furthermore, the presence of these proteins have been reported to be related to infection caused by protozoan, for example anti-ribosomal P-antibodies were previously detected against *Trypanosoma cruzi* and *Leishmania* species (Mesri et al., 1990, Soto et al., 1996).

The second most represented class from the identified membrane proteins was the hydrolase (PC00121). Under this category, six calcium-transporting ATPases were identified, e.g. EHL_054830, EHL_016480, EHL_030830, EHL_027710, EHL_065670, EHL_136610. These proteins also fall under the transporter protein class (PC00227). The present study identified six of the eight calcium-ATPase found in the *E. histolytica* protein database and they were also reported by Martinez-Higuera et al. (2013). The identified proteins were reported to contain the E1–E2 ATPase motif that belongs to the P-type ATPases (Brini and Carafoli, 2009). This E1–E2 motif was not shared in the EHL_001150 and EHL_159040 sequences that
were not identified in this study as well as the study by Martinez-Higuera et al. (2013). In addition, the identified calcium-ATPases were predicted to contain between four and six transmembrane regions, hence, confirming that these proteins resided at the membrane of *E. histolytica*. Calcium-ATPases are involved in calcium homeostasis by catalysing the active efflux of calcium ion from the cytoplasm (Martinez-Higuera et al., 2013). Furthermore, these proteins play an important role in the cellular process signalling of *E. histolytica*, including cell development and pathogenesis (Martinez-Higuera et al., 2013).

Another interesting set of proteins identified from the membrane fractions and classified under the hydrolase protein class (PC00121) were the signal peptidases (EHI_031250, EHI_121860, EHI_193470, EHI_197020, EHI_051790 and EHI_200720). EHI_200720 was also highlighted in Table 4.9 whereby this subunit was exclusively identified by LC-MALDI-TOF/TOF. Signal peptidases are important for normal cell function as they are needed to release translocated pre-proteins from the membrane to extracellular environment (Auclair et al., 2012). These proteins cleave the N-terminal signal peptides of secretory proteins at the membrane and also rudiment precursor protein molecules that are under elongation by a ribosome that is bound to the endoplasmic reticulum translocation site (Barrett et al., 2012). The disruption of the signal peptidase activity has proved to cause significant effects. For example, the replication of Zika and Dengue viruses were inhibited with the inhibition of the signal peptidase in the human cell line (Estoppey et al., 2017).

Important proteins that were annotated under the membrane category (GO:0016020) include the multidrug resistance protein (EHI_100320) and ATP-binding cassette proteins (EHI_095820, and EHI_178580) (Figure 4.16). These
proteins are present in both trophozoites and cysts stages of *E. histolytica* (Luna-Nácar et al., 2016). Multidrug resistance protein was identified in the membrane fraction of all three extraction methods. Furthermore, membrane protein prediction analysis identified a total of 16 transmembrane regions in the sequence. Multidrug resistance protein can be found in the cytoplasm, vacuole and plasma membrane as annotated in the cellular component. Membrane protein prediction analyses found a total of 12-13 transmembrane regions in the ATP-binding cassette (EHI_095820, EHI_178050 and EHI_178580). The ATP-binding cassette is a membrane transport protein known for their involvement in regulating and rendering resistance to chemotherapeutic drugs by limiting the intracellular accumulation of certain cytotoxic agents and xenobiotics (Cole et al., 1992).

Several actin-binding proteins were identified in this study. Previously, Váazquez et al. (1995) described four actin binding proteins (vinculin, α-actinin, tropomyosin and myosin I) of amoebic adhesion plaques/plates. In the present study, actin-binding proteins (EHI_168340), cofilin/tropomyosin family (EHI_186840), F-actin capping protein beta subunit (EHI_005020), actophorin (197480) and myosin heavy chain (EHI_110180) were identified. These proteins do not only exist on the surface of *E. histolytica* as an adhesion plaques component but also aid the parasite in locomotion and invasion into the host tissue (Váazquez et al., 1995).

The pore-forming peptide amoebapore B precursor (EHI_194540) is a protein that is involved in the cytolytic activity of *E. histolytica*. In a study by Nickel et al. (1999), *E. histolytica* expressed more amoebapore and showed higher cytolytic activity as compared to *E. dispar*. This protein was identified in this study, but it was not categorised in any of the functional ontology. Interestingly, this protein contains an N-terminal signal peptide, which indicates that this protein is a classically secreted
protein. However, it was not detected in the ES protein study and yet was identified in the membrane fractions extracted using all three membrane protein extraction protocols. The absence of amoebapore among the ES proteins was consistent with the findings of Leippe et al. (1995) whereby they described that amoebapores were secreted by *E. histolytica* only if it was in contact with the target host cell.

Another group of proteins that were not categorised under any protein class are the heat shock proteins (EHI_001950, EHI_148990, and EHI_199590). In this study, these proteins were classified as secreted proteins, but they were also found consistently in all the membrane fractions of *E. histolytica*. In addition, heat shock protein (HSP) 70 of *E. histolytica* was reported to stimulate immune responses in patients with amoebiasis (Ortner et al., 1992). Furthermore, amoebic HSPs were suggested to be involved in the degradation of cytoskeletal proteins for the purpose of encystation (Field et al., 2000). These proteins are involved in response to a stimulus. However, since HSPs were observed in this study whereby the parasites were not subjected to stress, the presence of HSPs may not necessarily indicate that the *E. histolytica* trophozoites were under stress. Similarly, a previous study demonstrated that HSPs were consistently present in heated and unheated cells of *Amoeba proteus* (Podlipaeva, 2001).

### 5.3 The complementary use of LC-MALDI-TOF/TOF and LC-ESI-MS/MS system

The application of LC-MALDI-TOF/TOF and LC-ESI-MS/MS in the ES and membrane protein studies in chapter 4.1 and chapter 4.2 has highlighted the advantages of using two MS systems to increase the coverage of protein hits. Previous studies have shown that LC-MALDI-TOF/TOF and LC-ESI-MS/MS are
complementary ionisation techniques and that the simultaneous use of both systems have led to the identification of more peptides (Baldwin et al., 2001, Medzihradszky et al., 2001, Hansen et al., 2003, Heller et al., 2003, Stapels and Barofsky, 2004). In the present ES and membrane proteins study, the ESI system identified more proteins as compared to MALDI. This was consistent in Yang et al. (2007) as they have demonstrated MALDI detected a lower percentage of sequence coverage than the ESI system. Similarly, Medzihradszky et al. (2001) described the superiority of the ESI system, whereby nine peptides were identified with LC-MALDI-TOF/TOF, whereas ESI was able to detect six additional peptides. Yet, other studies have also reported that MALDI was superior to the ESI system (Stapels and Barofsky, 2004, Bereszczak and Brancia, 2009). For example, Stapels and Barofsky (2004) observed that the LC-MALDI-TOF/TOF identified two times more proteins than the LC-ESI-MS/MS. This was probably due to the nature of the abundant peptides in their samples that better suited the characteristics of ionisation by MALDI.

According to Bodnar et al. (2003), the low number of proteins identified by the MALDI system can be due to the suppression of peptide ion signals by the matrix. On the other hand, the better sensitivity of the ESI system was due to the multiply charged peptides and cleaner spectra with less internal fragments compared to the singly charged ions produced by MALDI. However, this can sometimes lead to less structural information, thus leading to lesser peptide identification (Bodnar et al., 2003). In Medzihradszky et al. (2001), peptides that were unique to the ESI system had either ended with a lysine residue or contained no basic amino acid residues. Stapels and Barofsky (2004) described that the ESI system tends to ionise hydrophobic peptides while MALDI tends to identify basic and aromatic peptides. MALDI is able to detect higher proportions of arginine (R) terminating peptides.
while ESI favours lysine (K) terminating peptides. Hence, the complementary application of the two systems can increase peptide coverage by approximately 10% (Stapels and Barofsky, 2004).
CHAPTER 6 – SUMMARY AND CONCLUSION

The proteome of *E. histolytica* plays an integral part in the pathogenesis of the disease. During *E. histolytica* infections, the interactions between the *E. histolytica* proteome and the host leads to tissue invasion, which eventually leads to the development of the disease. Elucidation of the function of *E. histolytica* proteins can aid in the discovery of new biomarkers. Hence, an in-depth proteome profiling study is needed. To achieve this goal, this study was performed to identify the ES and the membrane proteins.

Protein-free media was used to isolate the ES proteins of *E. histolytica* trophozoites in this study. The ES proteins were identified using LC-MALDI-TOF/TOF and LC-ESI-MS/MS. This study has identified a total of 209 ES proteins. Subsequently, classical and non-classical secretory prediction analyses revealed that only 19% of the ES proteins contained motifs of either classically or non-classically secreted proteins.

In this study, three membrane protein extraction methods were compared. The ProteoExtract® kit and the conventional method extracted the highest protein yield compared to the ProteoPrep® kit. The reproducibility of all three extraction methods was shown in the SDS-PAGE protein band patterns. Nonetheless, the protein identification for the membrane fraction extracted using the conventional method resulted in the highest number of proteins. However, when the selectivity analysis was performed, the result indicated that the ProteoPrep® kit extracted the membrane and the cytosolic fractions with the lowest percentage of cross-contamination. By using membrane protein prediction analysis, the ProteoPrep® kit was determined to have the highest specificity for the extraction of membrane
proteins. There were proteins exclusively identified in each of the membrane fractions extracted by the ProteoExtract® kit, ProteoPrep® kit and the conventional method. Therefore, the combined membrane fraction proteins from each method were used as a reflection of the *E. histolytica* trophozoite’s membrane proteome. In this study, 876 membrane proteins were identified, of which 254 were predicted to contain at least one transmembrane domain.

The combination of LC-MALDI-TOF/TOF and LC-ESI-MS/MS has clearly improved the proteome coverage of the *E. histolytica* ES and membrane proteins. Although the high throughput methodology employed in this study has shed some light into the proteome of the *E. histolytica* trophozoites, shortcomings were undeniably present at all levels of the experiments. To address the current limitation, future efforts are needed to improve the experiment by culturing the trophozoites in an environment that mimics the human intestinal condition and proteomic analyses should be performed for a larger number of biological replicates. For ES proteome study, the *E. histolytica* trophozoites should include investigations on the relationship between the parasite and the host’s gut microflora. Furthermore, future study should also focus on characterising and validating the subcellular localisation of these alleged *E. histolytica* membrane proteins.

In conclusion, the present study has shown enhanced proteome coverage with the complementary use of two different mass spectrometry systems. In addition, the comparisons of the membrane extraction methods were conclusive whereby ProteoPrep® kit was identified as the best method, nevertheless, the membrane protein information obtained from this kit alone was insufficient. The proteomic analyses of the ES and membrane proteins in this study have provided information
for further elucidation of the *E histolytica* proteome and its involvement in the pathogenesis.
REFERENCES


RABILLOUD, T. 2009. Membrane proteins and proteomics: love is possible, but so difficult. Electrophoresis, 30.


and PCR of Human Liver Abscess Aspirates from Patients in China for *Entamoeba histolytica*. *Journal of clinical microbiology*, 37, 3034-3036.


Appendix 1


**Proteome analysis of excretory-secretory proteins of *Entamoeba histolytica* HM1:IMSS via LC-ESI-MS/MS and LC-MALDI-TOF/TOF**

Jorim Anak Ujang¹, Kwan Soon Hong², Mohd Nazri Ismail,² Lim Boon Huat³, Rahmah Noordin,¹ Nurulhasanah Othman,¹,*

Institute for Research in Molecular Medicine (INFORMM), Universiti Sains Malaysia, 11800, Penang, Malaysia;¹ Analytical Biochemistry Research Centre, Universiti Sains Malaysia, 11800, Penang, Malaysia,² School of Health Sciences, Universiti Sains Malaysia, 16150, Kelantan, Malaysia³

**Abstract**

Background: Excretory-secretory (ES) proteins of *E. histolytica* are thought to play important roles in the host invasion, metabolism, and defence. Elucidation of the types and functions of *E. histolytica* ES proteins can further our understanding of the disease pathogenesis. Thus, the aim of this study is to use proteomics approach to better understand the complex ES proteins of the protozoa. Methods: *E. histolytica* ES proteins were prepared by culturing the trophozoites in protein-free medium. The ES proteins were identified using two mass spectrometry tools, namely, LC–ESI–MS/MS and LC–MALDI–TOF/TOF. The identified proteins were then classified according to their biological processes, molecular functions, and cellular components using the Panther classification system (PantherDB). Results: A complementary list of 219 proteins was identified; this comprised 201 proteins detected by LC–ESI–MS/MS and 107 proteins by LC–MALDI–TOF/TOF. Of the 219 proteins, 89 were identified by both mass spectrometry systems, while 112 and 18 proteins were detected exclusively by LC–ESI–MS/MS and LC–MALDI–TOF/TOF respectively. Biological protein functional analysis using PantherDB showed that 27% of the proteins were involved in metabolic processes. Using molecular functional and cellular component analyses, 35% of the proteins were found to be involved in catalytic activity, and 21% were associated with the cell parts. Conclusion: This study showed that complementary use of LC–ESI–MS/MS and LC–MALDI–TOF/TOF has improved the identification of ES proteins. The results have increased our understanding of the types of proteins excreted/secreted by the amoeba and provided further evidence of the involvement of ES proteins in intestinal colonisation and evasion of the host immune system, as well as in encystation and excystation of the parasite.

Keywords: *Entamoeba histolytica*, Excretory-secretory (ES) proteins, LC–ESI–MS/MS, LC–MALDI–TOF/TOF
Appendix 2

OP39: Membrane proteome analysis of *Entamoeba histolytica*

*Ujang Jorim, Noordin, Rahmah and Othman, Nurulhasanah*

Institute for Research in Molecular Medicine (INFORMM), Universiti Sains Malaysia, Penang, Malaysia

*jorim_ujang@hotmail.com*

Abstract

*Entamoeba histolytica* causes amoebiasis, which leads to amoebic dysentery and amoebic liver abscess, of which is fatal if untreated. The membrane proteins of *E. histolytica* are key molecules in the development of the disease and are crucial to the discovery of biomarkers for drug therapy, vaccine development and diagnostic markers. In this study, two commercial membrane proteins extraction methods, ProteoExtract Native Membrane Protein Extraction Kit (Calbiochem) and ProteoPrep Membrane Extraction Kit (Sigma), and a conventional method were compared. The amount of proteins yielded for Calbiochem extracts were significantly higher than Sigma and conventional method. A total number of 171 proteins were identified using a combination of the three extraction methods by LC-MALDI-TOF/TOF, in which 56, 30 and 127 proteins were identified from Calbiochem, Sigma and conventional method, respectively. Furthermore, 35, 7 and 35 proteins were unique to Calbiochem, Sigma and conventional method. However, only 10 proteins were found to be common among the three methods. Functional classification of protein classes using PantherDB software revealed that the proteins are involved as oxidoreductase (23.9%), cell adhesion molecule (0.8%), membrane traffic protein (4.4%), receptor (3.5%), signaling molecule (1.9%) and transfer/carrier protein (0.6%). Overall, the Calbiochem kit gave the highest yield of membrane protein extract, while, the conventional method gave the best hits of protein identification. In conclusion, the methods compared in this study were complementary to each other based on mass spectrometry analyses and its combination improved membrane proteins coverage.

*Keywords: Entamoeba histolytica; membrane proteins; LC-MALDI-TOF/TOF*