

**PROTEOMIC STUDIES AND ITS APPLICATION
TO BIOLOGICAL SAMPLES USING MASS
SPECTROMETRY**

**A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of**

DOCTOR OF PHILOSOPHY

in Chemistry

**by
Melda Zeynep GÜRAY**

**May 2017
İZMİR**

We approve the thesis of **Melda Zeynep GÜRAY**

Examining Committee Members:

Prof. Dr. Ahmet KOÇ

Medical Biology and Genetics Department, İnönü University

Prof. Dr. Ahmet ÖZBİLGİN

Department of Parasitology, Celal Bayar University

Prof. Dr. Figen ZİHNİOĞLU

Department of Biochemistry, Ege University

Prof. Dr. Talat YALÇIN

Department of Chemistry, İzmir Institute of Technology

Assoc. Prof. Dr. Gülşah ŞANLI

Department of Chemistry, İzmir Institute of Technology

30 May 2017

Prof. Dr. Talat YALÇIN

Supervisor, Department of Chemistry
İzmir Institute of Technology

Assoc. Prof. Dr. Çağlar KARAKAYA

Co-Supervisor, Department of Molecular
Biology and Genetics
İzmir Institute of Technology

Prof. Dr. Ahmet Emin EROĞLU

Head of the Department of Chemistry

Prof. Dr. Aysun SOFUOĞLU

Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my supervisor Prof. Dr. Talat YALÇIN, for his professional guidance, valuable help, encouragement, understanding and endless patience not only throughout this study but also for other situations. I am so proud to receive my PhD degree under his supervision.

I would like to thank to Prof. Dr. Ahmet ÖZBİLGİN from Celal Bayar University for his collaboration in *Leishmania* study and Prof. Dr. Alan DOUCETTE from Dalhousie University for giving me opportunity to carry out acetone modification study in his laboratory in Canada, also for their beneficial suggestions and valuable comments. I am also thankful to my thesis progress committee members Prof. Dr. Ahmet KOÇ, Assoc. Prof. Dr. Gülşah ŞANLI and my co-supervisor Assoc. Prof. Dr. Çağlar KARAKAYA for their guidance and helps.

I am pleased to acknowledge the Scientific and Technological Research Council of Turkey, TÜBİTAK, with a project no 214S239.

I wish to express my thankfulness to my friends, Dr. Ahmet Emin ATİK, Dr. İrem ULUIŞIK, Begüm YETİŞER, Gözde YÜCEL, Seda ARAŞLIK, Ayça TUTAK, Dr. Hande KARAOSMANOĞLU, Dr. Hatice YİĞİT, Dr. Melis DİNÇ KANT, Dr. Sıla KARACA and my lab mates Dr. Filiz YEŞİLIRMAK, Melike DİNÇ and Dr. Çağdaş TAŞOĞLU for their friendship, kind helps and supports.

Finally, I want to express my deepest gratitude and love to my mother Hümeysra GÜRAY, my father İrfan GÜRAY, my brother Berk GÜRAY for their understanding and encouragement that they have showed in every stages of my life, including this thesis study. I also want to thank to my nephews Kuzey GÜRAY and Uzay GÜRAY for providing motivation with their love and cuteness. Last but not least, I am grateful to Mert TAŞKINARDA for being in my life, for his endless love and understanding. I would not be able to finish my study without these people.

ABSTRACT

PROTEOMIC STUDIES AND ITS APPLICATION TO BIOLOGICAL SAMPLES USING MASS SPECTROMETRY

Mass spectrometry (MS) is a powerful analytical tool with its application in the field of biological sciences for identification of proteins, defining post-translational modifications, studying protein expression and protein interactions. This thesis presents MS analyses of proteins for defining modifications observed during sample preparation and identification of proteins isolated from clinical samples and microorganisms.

The first part of the thesis includes proteomic analysis of antimony resistant *L. tropica*. The results clearly indicated that the generation of antimony resistance by parasites, either in host organism or *in vitro*, causes alteration of protein expression levels, and the mechanism of antimony resistance in host organism and *in vitro* conditions follow different strategies. In the second part of the study, proteomic analysis of Bence Jones proteins isolated from urine of multiple myeloma patients was performed. Gel electrophoresis and MS analysis revealed that the proteins from different patients with different nephrotoxicity have different tendencies to form multimeric structures and contained different type of light chain. In the third part, it was shown that precipitation of proteins in acetone causes +98 u mass artifacts on proteins when analyzed by MS. The parameters affecting the formation of modification was studied and it was revealed that this modification is dependent on solution pH, incubation time and temperature. In the last part, aspartic acid and glutamic acid containing synthetic peptides were shown to be methylated upon incubation in acidified methanol solution. MS analysis revealed that the reaction is dependent on temperature and time and is affected by the type of acid included in methanol solution.

All in all, this thesis provides a comprehensive study of proteins by mass spectrometry for identification of proteins from different sources, as well as defining protein modifications observed as artifacts during sample handling in proteomic workflows.

ÖZET

PROTEOMİK ÇALIŞMALARI VE KÜTLE SPEKTROMETRESİ KULLANILARAK BİYOLOJİK ÖRNEKLERE UYGULANMASI

Kütle spektrometrisi (MS), biyolojik bilimler alanında uygulamaları ile protein tanımlanması, post-translasyonel modifikasyonların belirlenmesi, protein ifadenmesi ve protein etkileşim çalışmaları için güçlü bir analitik yöntemdir. Bu tez, örnek hazırlama sırasında oluşan modifikasyonların belirlenmesi ve klinik örneklerden ve mikroorganizmalardan elde edilen proteinlerin tanımlanması için proteinlerin MS analizlerini sunmaktadır.

Çalışmanın ilk bölümü, antimon dirençli *L. tropica*'nın proteomik analizini içermektedir. Sonuçlar açıkça göstermiştir ki; parazit tarafından konakçı organizmada ya da *in vitro* koşullarda geliştirilen ilaç dirençliliği, parazitlerde protein ifadenme düzeylerinde değişikliklere sebep olmuştur ve konakçı hücre ve *in vitro* koşullarda oluşan ilaç dirençliliği mekanizmaları farklı yollar izlemektedir. Çalışmanın ikinci bölümünde, multipl miyelom hastalarının idrarlarından izole edilen Bence Jones proteinlerinin proteomik analizleri gerçekleştirilmiştir. Jel elektroforezi ve MS analizleri, farklı hastalardan elde edilen farklı nefrotoksik özelliğe sahip proteinlerin, farklı multimer yapılar oluşturduğunu ve farklı tür hafif zincire sahip olduğunu ortaya çıkarmıştır. Üçüncü bölümde, proteinlerin aseton ile çöktürülmesi sırasında asetonun proteinler üzerinde +98 u kalıntılar oluşturduğu MS analizleri ile gösterilmiştir. Bu modifikasyonun oluşumunu etkileyen parametreler çalışılmış ve modifikasyonun çözelti pH'ına, inkübasyon zamanına ve sıcaklığına bağımlı olduğu ortaya çıkarılmıştır. Son olarak, aspartik asit ve glutamik asit içeren sentetik peptitlerin, asidik metanol çözeltisi içinde inkübe edilmesi sebebiyle metillendiği gösterilmiştir. MS analizleri, reaksiyonun sıcaklık ve zamana bağımlı olduğunu ve çözeltinin içeriğinde bulunan asit türünden etkilendiğini ortaya koymuştur.

Sonuç olarak; bu tez, proteinlerin farklı kaynaklardan tanımlanması ve aynı zamanda proteomik çalışmalarında örnek hazırlama sırasında oluşan protein modifikasyonlarının belirlenmesinde, proteinlerin MS ile kapsamlı bir şekilde çalışılmasını sağlamaktadır.

TABLE OF CONTENTS

LIST OF FIGURES	ix
LIST OF TABLES	xii
LIST OF ABBREVIATIONS	xiii
CHAPTER 1. MASS SPECTROMETRY AND PROTEOMICS	1
1.1. Mass Spectrometry	1
1.2. Proteomics	3
1.3. Mass Spectrometry-Based Proteomics	5
1.3.1. Separation Techniques.....	5
1.3.2. Instrumentation.....	7
1.3.3. Protein Identification by Mass Spectrometry	11
1.3.4. Applications of Mass Spectrometry-Based Proteomics in Health Sciences	17
CHAPTER 2. INVESTIGATION OF DRUG RESISTANCE IN <i>IN VITRO</i> AND CLINICAL ISOLATES OF ANTIMONY-RESISTANT <i>L. tropica</i> BY PROTEOMIC ANALYSIS	19
2.1. Introduction	19
2.2. Experimental Methods.....	28
2.2.1. <i>Leishmania</i> Samples and Parasite Growth	28
2.2.2. Total Protein Extraction	28
2.2.3. Two-Dimensional Gel Electrophoresis	29
2.2.4. In-Gel Digestion and Peptide Extraction	30
2.2.5. MALDI-MS/MS Analysis and Database Search	31
2.3. Results and Discussion	31
2.3.1. Cytoskeletal Proteins	35
2.3.2. Protein Biosynthesis	48
2.3.3. Antioxidant and Detoxification	49
2.3.4. Protein Folding/Chaperons and Stress Proteins	50

2.3.5. RNA/DNA Processing.....	52
2.3.6. Metabolic Enzymes	52
2.3.7. Hypothetical Proteins	55
2.4. Conclusion.....	56
CHAPTER 3. PROTEOMIC ANALYSIS OF BENICE JONES PROTEINS ISOLATED FROM URINE OF PATIENTS WITH MULTIPLE MYELOMA	59
3.1. Introduction	59
3.2. Experimental Methods.....	64
3.2.1. Isolation and Purification of Light Chains	64
3.2.2. One- and Two-Dimensional Gel Electrophoresis.....	64
3.2.3. In-Gel Digestion and Peptide Extraction	65
3.2.4. MALDI-MS/MS Analysis and Database Search	66
3.2.5. LC-ESI-MS and MALDI-MS Analysis of Proteins.....	67
3.3. Results and Discussion	68
3.4. Conclusion.....	82
CHAPTER 4. MASS SPECTROMETRY OF INTACT PROTEINS REVEALS +98 U CHEMICAL ARTIFACTS FOLLOWING PRECIPITATION IN ACETONE	84
4.1. Introduction	84
4.2. Experimental Methods.....	87
4.2.1. Protein Precipitation	87
4.2.2. <i>E.coli</i> Proteome Extraction and Precipitation	87
4.2.3. Preparation of Apo-Cytochrome c	88
4.2.4. Protein Digestion.....	88
4.2.5. LC-MS Analysis of Intact Proteins or Peptides	89
4.2.6. LC-MS/MS Analysis of Tryptic Peptides and Database Search...	90
4.3. Results and Discussion	91
4.4. Conclusion.....	104

CHAPTER 5. OBSERVATION OF THE SIDE CHAIN METHYLATION OF GLUTAMIC ACID AND ASPARTIC ACID CONTAINING PEPTIDES BY MASS SPECTROMETRY	106
5.1. Introduction	106
5.2. Experimental Methods.....	109
5.2.1. Preparation of Peptide Samples.....	109
5.2.2. Mass Spectrometry Analysis of Peptides	109
5.2.3. HPLC Analysis.....	110
5.3. Results and Discussion	110
5.4. Conclusion.....	123
 CHAPTER 6. CONCLUSION	 125
 REFERENCES	 127
 APPENDICES	
APPENDIX A. BUFFERS AND GEL PREPARATION FOR SDS-PAGE.....	142
APPENDIX B. HIGH RESOLUTION MS SPECTRA OF CYTOCHROME C AND LIST OF ACETONE MODIFIED <i>E. COLI</i> PEPTIDES	145
APPENDIX C. ESI-MS SPECTRA OF METHYLATED PEPTIDES.....	147

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
Figure 1.1. Basic diagram of a mass spectrometer	1
Figure 1.2. A mass spectrum obtained from trypsin digestion of retinol binding protein isolated from human urine	2
Figure 1.3. Formation of analyte ions by MALDI.....	8
Figure 1.4. Formation of analyte ions by ESI.....	9
Figure 1.5. MS spectra of cytochrome c ionized by (a) MALDI, (b) ESI.....	10
Figure 1.6. An overview of different strategies employed in proteomic studies for protein identification by mass spectrometry	13
Figure 1.7. The nomenclature of peptide fragment ions.....	15
Figure 2.1. Microscope images of <i>Leishmania</i> promastigotes (a) and amastigotes (b)..	20
Figure 2.2. Comparison of 2D-GE profiles of <i>in vitro</i> antimony resistant lines (DI) with their corresponding sensitive pairs (DNI).....	33
Figure 2.3. Comparison of 2D-GE profiles of <i>in vitro</i> antimony resistant lines (DI) with resistant clinical isolate (D1)	34
Figure 2.4. Comparison of 2D-GE profiles of antimony resistant clinical isolate, D1 with sensitive isolate, DNI3	35
Figure 2.5. Relative distribution of identified proteins according to their biological roles.....	48
Figure 3.1. Schematic representation of an immunoglobulin structure	60
Figure 3.2. 1D-GE of samples performed under (a) reducing conditions, (b) non-reducing conditions	68
Figure 3.3. MALDI-MS spectra of samples	70
Figure 3.4. RP-LC profiles of samples (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, (f) 6.....	74
Figure 3.5. ESI-MS spectra of samples.	75
Figure 3.6. 2D-GE analysis of samples (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, (f) 6.....	76
Figure 3.7. The full sequence of light chains obtained by database search from; spot 3 of sample 1, spot 2 of sample 2, spot 4 of sample 3, spot 7 of sample 4, spot 4 of sample 5 and spot 7 of sample 6.....	80
Figure 3.8. Effect of (a) light chains and (b) different concentration of light chains on ³ H-thymidine incorporation by human proximal tubule cells.....	81

Figure 3.9. Effect of different light chains and HSA on IL-6 production in proximal tubule cells.	82
Figure 4.1. (a) Multiply charged and (c) deconvoluted ESI-MS spectra of cytochrome c following precipitation in 80% acetone. (b) Multiply charged and (d) deconvoluted MS spectra of non-precipitated control	92
Figure 4.2. Deconvoluted MS spectra of cytochrome c following incubation in (a) 80% acetonitrile, or (b) 80% methanol, for 1 hour at -20 °C	92
Figure 4.3. LC-MS analysis of cytochrome c or following acetone precipitation.....	94
Figure 4.4. (a) Deconvoluted MS spectra of cytochrome c following incubation in 80% acetone for defined incubation periods. (b) The degree of protein modification as a function of time	95
Figure 4.5. (a) Deconvoluted MS spectra of cytochrome c following incubation in 80% acetone for 1 hour at specified temperatures (b) The level of modification is quantified as a function of temperature	96
Figure 4.6. Deconvoluted MS spectra of cytochrome c following overnight incubation in 80% acetone and subsequent evaporation of the solvent via (a) vacuum concentrator or (b) conventional protocol.....	97
Figure 4.7. Deconvoluted MS spectra of cytochrome c prepared in solutions of varying pH and precipitated in 80% acetone	99
Figure 4.8. Deconvoluted MS spectra of acetone-precipitated proteins prepared in acidic, neutral and basic solutions. (a) Myoglobin, (b) ubiquitin, (c) apo-cytochrome c, (d) hemoglobin	101
Figure 4.9. MS/MS analysis of tryptic peptides from (a) DNA-binding protein H-NS from <i>E. coli</i> , or (b) BSA.....	102
Figure 4.10. Proposed reactions of diacetone alcohol (I) or mesityl oxide (II) with protein.....	103
Figure 4.11. Deconvoluted MS spectra of cytochrome c incubated in 80% acetonitrile containing 0.1% mesityl oxide or 0.1% diacetone alcohol	104
Figure 5.1. ESI-MS spectra of (a) EGGFL-NH ₂ and (b) DGGFL-NH ₂ following thirty days of incubation in MeOH/dH ₂ O/FA (50:50:1, v/v/v) at -20 °C, 4 °C, 22 °C, 37 °C and 50 °C	111
Figure 5.2. Comparison of MS/MS spectrum of <i>m/z</i> 535 ion corresponding to methylated product of EGGFL-NH ₂ and MS/MS spectrum of [M + H] ⁺ ion of commercial E _{OMe} GGFL-NH ₂	113

Figure 5.3. LC chromatograms of freshly prepared EGGFL-NH ₂ peptide solution and EGGFL-NH ₂ peptide solution incubated in MeOH/dH ₂ O/FA (50:50:1, v/v/v) at 37 °C for a week	114
Figure 5.4. ESI-MS spectra of (a) EGGFL-OH and (b) DGGFL-OH recorded following incubation in MeOH/dH ₂ O/FA (50:50:1, v/v/v) at 37 °C over a range of incubation times	115
Figure 5.5. MS/MS spectra of mono-methylated and di-methylated products of EGGFL-OH following incubation in MeOH/dH ₂ O/FA (50:50:1, v/v/v) at 37 °C for 30 days	116
Figure 5.6. ESI-MS spectra of (a) AAAEAAA-NH ₂ and (b) AAADAAA-NH ₂ following incubation in MeOH/dH ₂ O/FA (50:50:1, v/v/v) at 37 °C for 30 days.....	117
Figure 5.7. ESI-MS spectra of (a) AAAEAAA-OH and (b) AAADAAA-OH following incubation in MeOH/dH ₂ O/FA (50:50:1, v/v/v) at 37 °C for 30 days.....	117
Figure 5.8. ESI-MS spectra of (a) AAAAAAEE-NH ₂ , (b) AAAAAAEE-OH and (c) AAAAAADD-OH following incubation in MeOH/dH ₂ O/FA (50:50:1, v/v/v) mixture at 37 °C for 30 days	119
Figure 5.9. ESI-MS spectra of (a) EGGFL-NH ₂ and (b) DGGFL-NH ₂ following incubation in MeOH/dH ₂ O (50:50, v/v) mixture containing 0.1% or 1.0% of acetic acid or formic acid, at 37 °C for 30 days.....	121
Figure 5.10. ESI-MS spectra of (a) EGGFL-NH ₂ and (b) DGGFL-NH ₂ following incubation in EtOH/dH ₂ O/FA (50:50:1, v/v/v) mixture, (c) EGGFL-NH ₂ and (d) DGGFL-NH ₂ following incubation in <i>i</i> PrOH/dH ₂ O/FA (50:50:1, v/v/v) mixture, at 37 °C for 30 days	122

LIST OF TABLES

<u>Table</u>	<u>Page</u>
Table 2.1. Differentially expressed proteins in Sb-resistant <i>Leishmania spp.</i>	25
Table 2.2. MS/MS identification of differentially expressed proteins in DNI1 and DI1 by database search.....	36
Table 2.3. MS/MS identification of differentially expressed proteins in DNI2 and DI2 by database search.....	37
Table 2.4. MS/MS identification of differentially expressed proteins in DNI3 and DI3 by database search.....	38
Table 2.5. MS/MS identification of differentially expressed proteins in D1 and DI1 by database search.....	39
Table 2.6. MS/MS identification of differentially expressed proteins in D1 and DI2 by database search.....	42
Table 2.7. MS/MS identification of differentially expressed proteins in D1 and DI3 by database search.....	44
Table 2.8. MS/MS identification of differentially expressed proteins in D1 and DNI3 by database search.....	46
Table 3.1. MS/MS identification of protein spots by database search	77

LIST OF ABBREVIATIONS

1D-GE	One-dimensional gel electrophoresis
2D-GE	Two-dimensional gel electrophoresis
2D-HPLC	Two-dimensional high pressure liquid chromatography
AA	Acetic acid
AL	Light chain amyloidosis
BJP	Bence Jones proteins
BSA	Bovine serum albumin
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate
CHCA	α -cyano-4-hydroxycinnamic acid
CID	Collision induced dissociation
CL	Cutaneous leishmaniasis
CDR	Complementarity determining region
DC	Direct current
DAA	Diacetone alcohol
DHB	2,5-dihydroxybenzoic acid
DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
EtOH	Ethanol
FA	Formic acid
FR	Framework region
FASP	Filter-aided sample preparation method
FT-ICR	Fourier transform ion cyclotron resonance
HPLC	High pressure liquid chromatography
HILIC	Hydrophilic interaction chromatography
HSA	Human serum albumin
ICAT	Isotope-coded affinity tag
IEF	Isoelectric focusing
iPrOH	Isopropanol
IEX	Ion exchange chromatography

IPG	Immobilized pH gradient
iTRAQ	Isobaric tags for relative and absolute quantification
LCDD	Light chain deposition disease
LC-MS	Liquid chromatography-mass spectrometry
MALDI	Matrix-assisted laser desorption ionization
MCL	Mucocutaneous leishmaniasis
MeOH	Methanol
MO	Mesityl oxide
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MudPIT	Multidimensional protein identification technology
NCBI	National Centre for Biotechnology Information
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCD	Programmed cell death
PCR	Polymerase chain reaction
pE	Pyroglutamic acid
PMF	Peptide mass fingerprinting
PTM	Post-translational modification
QqQ	Triple quadrupole
Q-TOF	Quadrupole/ Time-of-flight
RF	Radio frequency
RP	Reversed phase
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SILAC	Stable labeling of amino acids in cell culture
TCA	Trichloro acetic acid
TFA	Trifluoro acetic acid
TOF	Time-of-flight
VL	Visceral leishmaniasis
WHO	World Health Organization

CHAPTER 1

MASS SPECTROMETRY AND PROTEOMICS

1.1. Mass Spectrometry

A mass spectrometer is an analytical tool that measures the “mass-to-charge ratio” (m/z) of the analyte. Mass spectrometers are composed of three main components: ion source, mass analyzer and detector (Figure 1.1).

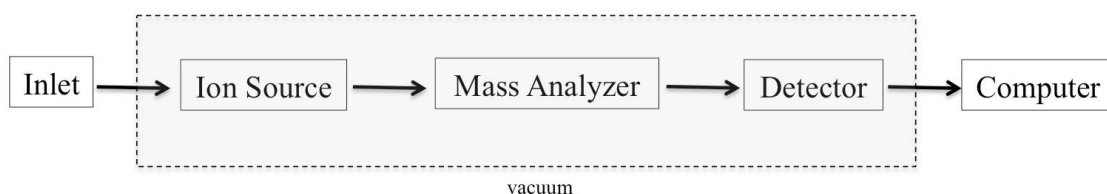


Figure 1.1. Basic diagram of a mass spectrometer

The information obtained from a mass spectrometer relies on the analysis of analyte molecules in the gas phase. The process involves some fundamental steps. First of all gas phase ions (positively or negatively charged) of the analyte molecules are formed in a suitable ion source. Next these ions are separated (electrically or magnetically) on the basis of their m/z values in a mass analyzer. Finally the ions reach to the detector where the number of ions at each m/z is detected. The main components of a mass spectrometer must be operated under high vacuum conditions (10^{-5} to 10^{-7} Torr), maintained by the vacuum systems, in order to eliminate the gas molecules other than analyte gas phase ions. At the end of analysis, the recorded data is transformed into a mass spectrum in which the abundance and m/z values of the ions are represented. An example mass spectrum is given in Figure 1.2, reflecting the abundances of peptides generated by trypsin digestion of a protein and their corresponding m/z values in Dalton (Da) unit. The data obtained from a mass spectrum not only provide information about molecular mass and abundance of the analyte molecules, but also about their purity, composition and structure.

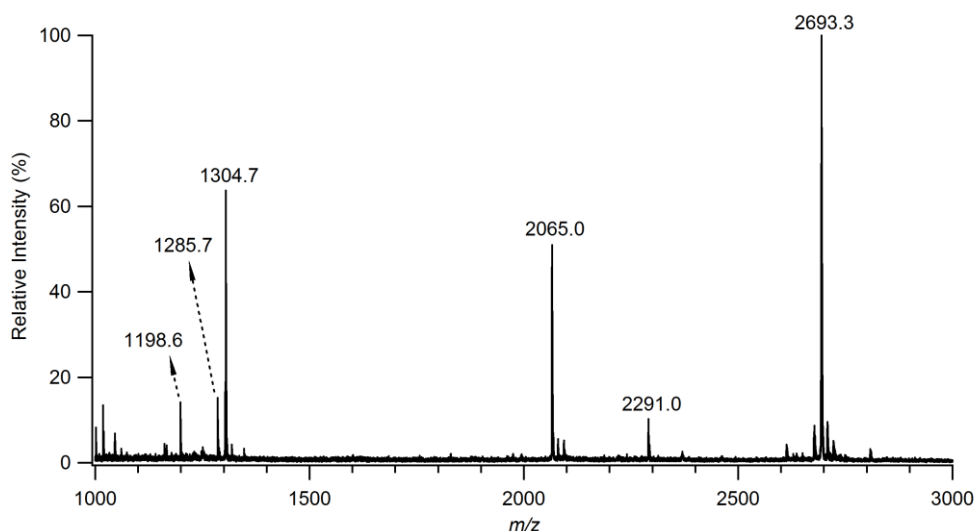


Figure 1.2. A mass spectrum obtained from trypsin digestion of retinol binding protein isolated from human urine

Mass spectrometry (MS) has a long history, which dates back to 1800s. The birth of mass spectrometry is commonly attributed to J. J. Thomson, with his discovery of electrons. Thomson measured the charge-to-mass ratio (e/m) and charge (e) of electrons using an electric field inside a cathode ray tube. Afterwards, Thomson and his students developed an instrument, which later recognized as the first mass spectrometer and it was commonly used for the needs in chemistry and physics. At the beginning of the 20th century, mass spectrometers were mainly used for the determination of the atomic masses and the discovery of the stable isotopes. By 1940s, during the World War II, with the demand for higher quality fuel and improvement of fighter aircraft performance, mass spectrometers were used in the petroleum industry for the analysis of small hydrocarbons. In the late 1950s, with the efforts of Klaus Biemann, mass spectrometers were used for characterization of small organic molecules through determination of their molecular weight. For this analysis, an ionization method, called as electron impact ionization, was used in which the analyte molecules are exposed to a beam of energetic electrons and resulted in the fragmentation of the molecule. Later on, efforts focused on the analysis of biomolecules like proteins, carbohydrates and nucleic acids. However, the current ionization methods were not suitable for the transfer of these thermally labile, large, nonvolatile and polar molecules into the gas phase without decomposition (1, 2). In 1980s, the problems associated with the transfer of biological molecules into the gas phase were resolved by the development of ‘soft ionization’ techniques, matrix-assisted laser desorption ionization (MALDI) (3, 4) and electrospray

ionization (ESI) (5), which allow biological macromolecules to be transferred into the gaseous state without dissociation. With these investments, John Fenn, for his work on ESI, and Koichi Tanaka, for his work on laser desorption ionization of proteins, shared the 2002 Nobel Prize in Chemistry. However, the credit must also be given to Michael Karas and Franz Hillenkamp's work on laser desorption ionization of biomolecules by the help of organic molecules, since the current MALDI systems operate mainly with the ideas developed by Karas and Hillenkamp.

1.2. Proteomics

The attempt to understand the linkage between genes and proteins started with the Human Genome Project and produced many fully or partially sequenced genomes of organisms. The number of complete genomes of organisms increases rapidly and they are available in the Entrez database of National Centre for Biotechnology Information (NCBI). Although genomic studies provided information about the genes and their possible protein products in organisms, the shift of the paradigm “one gene encodes one protein” to “one gene encodes more than one protein”, have made the works complicated. The human genome is assumed to have ~19,000 protein-coding genes, and if the post-genomic processes including splicing and post-translational modifications at the protein level are taken into account, there may be a million of human proteins. In contrast to genome that is static and does not change from cell to cell in a single organism, the protein content of an organism is dynamic that depends on the cell type and environmental conditions, and can change by showing responses to internal or external stimuli. In another words, although the complete genome or transcriptome (mRNA complement) of an organism form an infrastructure, they may not be sufficient to define all of the protein forms that actually exist in cells. At this point, the term ‘proteome’ was derived in 1994 to describe the entire protein complement of cells (6) and a new kind of “omics” science, referred to as proteomics, was generated. Proteomics deals with the identification, localization, interactions and modifications of all proteins encoded by the genome of organisms. Due to the importance of biological functions, modifications or interactions of proteins, proteomics has become one of the central topics in the fields of biological sciences.

The main objective of proteomics is to identify proteins expressed in cells as well as to identify the abundance, modifications and interactions. For that purpose, proteomic studies are conducted on the basis of following workflow, (i) separation and purification of proteins from complex mixtures, (ii) identifying or sequencing the protein of interest, (iii) access to DNA and protein databases and (iv) computer algorithms for data management and analysis. (7). A main tool in proteome research for separation and purification of proteins is polyacrylamide gel electrophoresis (PAGE). The technique can be applied either as one-dimensional (1D-GE) or two-dimensional (2D-GE) and helps to identify the molecular weight and/or isoelectric point of a protein. In a well-conducted 2D-gel electrophoresis experiment, more than 1000 protein bands can be separated and visualized from a whole cell extract. This technique has been widely used in proteomic studies for a long time since it can help to detect changes in protein expression in response to toxicity, drug treatment, disease conditions, differentiation, and different environmental conditions. The traditional techniques for the identification and sequencing of proteins are Western blot and Edman sequencing, respectively. In Western blotting, the proteins are separated on polyacrylamide gel and transferred onto a membrane (nitrocellulose or polyvinylidene difluoride). The identification is achieved on the basis of specific interactions between the protein of interest and an antibody, which recognizes that protein. Edman sequencing allows the determination of amino acid sequence of proteins and peptides through sequential removal of N-terminal amino acid in each cycle of Edman degradation and identification of each amino acid by liquid chromatography. Although these techniques were developed for a long time ago and are still in use, they suffer from some problems that make these techniques unsuited for proteomic work. The main problem associated with electrophoresis techniques is the low reproducibility of 2D-GE analysis. On the other hand, Western blot analysis requires the presumption of a suitable antibody and suffers from non-specific binding of antibody in some cases. The long analysis times (about 45 min of a cycle), inability for the detection of modified and N-terminally blocked proteins complicate Edman sequencing. With the introduction of the soft ionization techniques MALDI and ESI, mass spectrometry has become an unparalleled tool for the identification and sequencing of proteins. In proteomic studies, mass spectrometry provides advantages over the traditional techniques by its high sensitivity, rapid speed of analysis, ability to generate greater amount of information and determine post-translational modifications (8).

1.3. Mass Spectrometry-Based Proteomics

Mass spectrometry has evolved rapidly during the last 30 years with its applications in the field of biological sciences for identification of proteins, defining post-translational modifications, studying protein expression and protein interactions. The rapid increase in interest on mass spectrometry in proteomics studies is owing to the unmatched performance of the instrument and the continuous development in experimental methods, instrumentation and data analysis.

In a general workflow in MS-based proteomics study, first of all the protein mixture needs to be simplified into less complex components. This is achieved mainly through high-resolution 2D-GE or chromatographic techniques. Following separation, intact proteins or peptides derived from proteolytic digestion of proteins are analyzed by mass spectrometer to gather data related to their molecular mass and amino acid sequence. Lastly, the information obtained from mass spectrometer is searched against a protein database or processed by algorithms for the identification of proteins and modifications. Herein, the widely used mass spectrometer instrumentation, protein separation and identification approaches in proteomic studies will be discussed in details.

1.3.1. Separation Techniques

A key to success in MS-based proteomic studies is the correct sample preparation prior to MS analysis. To be analyzed in mass spectrometers, the biomolecules should be free of contaminants (in some cases a little amount is tolerable) and simplified in composition. Proteins or peptides obtained from complex mixtures of cell lysates include many contaminating molecules such as lipids, carbohydrates, nucleic acids, salts, detergents, which may hamper the quality of MS signals and proper identification of proteins. Therefore, appropriate separation and purification techniques should be employed prior to MS analysis. There are two main approaches used in proteomic workflows and classified as gel-based, including 2D-GE and gel-free, including liquid chromatography.

2D-GE has been used for a long time in proteomic studies. The method is based on a two-dimensional separation of proteins on polyacrylamide gels. The first

dimension is based on the separation according to isoelectric point (pI) of proteins by isoelectric focusing (IEF) method. The second dimension separates proteins by molecular weight using SDS-PAGE technique. 2D-GE allows the identification of pI and molecular weight of proteins, as well as charge and size variants of a protein in a single experiment. Moreover, it provides an estimation of protein amount by the color intensity of stained protein spots. This feature of 2D-GE forms the basis of comparative proteomics in which the difference in protein expression levels, upon i.e. growth of cells under different environmental conditions, are determined by comparing the color intensity of protein spots from different gels. The relevant protein spots from gels are excised, subjected to digestion by specific endo- or exoproteases and analyzed by several proteomic approaches for identification of proteins. Despite the advantages of 2D-GE, it suffers from several drawbacks. 2D-GE is a time consuming experiment and analysis of a high number of protein spots from gels increases this time dramatically. Moreover, proteins with high or low pI and molecular weight, or proteins with low solubility, such as membrane proteins, may not appear on gels. As stated before, poor reproducibility of gels is another issue. All in all, this technique is still commonly used in many laboratories with several improvements in instrumentation and experimental conditions.

High-pressure liquid chromatography (HPLC) takes an important place in proteomic studies. Several HPLC techniques are used in MS-based proteomic studies, and the principle of separation is dependent on the different interactions of molecules of distinct properties with a stationary phase, followed by elution by a mobile phase, which gradually disrupts the interactions between molecules and stationary phase. Chromatographic techniques are differentiated according to the used chromatographic material and some of these include; reversed phase (RP), ion exchange (IEX), hydrophilic interaction (HILIC) and size-exclusion (SEC). Among these, RP HPLC is commonly employed in separation and purification of proteins and peptides prior to MS analysis. The chromatographic material is made up of hydrophobic particles with covalently bonded alkyl chains and separation is based on the hydrophobic interactions between the stationary phase and biomolecule. The elution is achieved by a gradual increase in the hydrophobicity of mobile phase through increasing the concentration of an organic, non-polar solvent (e.g. acetonitrile) in aqueous solution. The main advantage that RP HPLC provides over other techniques is the compatibility of the mobile phase with ESI. Therefore, it can directly be coupled to ESI source and allows

the on-line analysis of biomolecules by a mass spectrometer. Another approach that is widely used in the separation of proteins and peptides employs the theory of multidimensional separation. This approach combines several separation techniques and also called as 2D-HPLC when two different chromatographic techniques are combined. An example of 2D-HPLC is the one in which the separation is initially performed by cation-exchange chromatography and a subsequent separation by RP chromatography. In cation exchange chromatography, the stationary phase is composed of negatively charged particles and separation is based on electrostatic interactions between charged groups of biomolecule and stationary phase. A salt gradient is applied for separation of biomolecules and eluted molecules are then directed to the RP column for the second dimension of separation. Multidimensional separation is a powerful technique in terms of resolution, particularly during the analysis of peptides originated from proteolytic digestion of a protein mixture, where the complexity of the sample increases as each protein results in the formation of multiple peptide products. This approach when combined with MS analysis and protein identification by database search is called as Multidimensional Protein Identification Technology (MudPIT) and automates the separation and identification of proteins in complex mixtures (9, 10).

1.3.2. Instrumentation

In analyzing biomolecules like proteins and peptides by mass spectrometry, ESI and MALDI are commonly used as ionization techniques. Both techniques allow the ionization and transfer of biomolecules into the gas phase without fragmentation, yet with different principles. In MALDI, the analyte is mixed with a solution of organic compound called matrix, which has the ability to absorb energy provided by laser (usually 337 nm). The analyte and matrix mixture is spotted on to a metal target plate, allowed for air-drying and the analyte is co-crystallized with the matrix. Formation of gas phase analyte ions occurs when the matrix molecules absorb laser energy and subsequently release that energy, accompanied with desorption of matrix together with analyte molecules into the gas phase (Figure 1.3). Positively or negatively charged ions of biomolecules are assumed to be produced by protonation or deprotonation between analyte and matrix molecules. The nature of biomolecule to be analyzed and the ionization mode of mass spectrometry are important parameters in determining which

matrix molecule to be used. The most commonly used matrix molecules for protein and peptide analysis are α -cyano-4-hydroxycinnamic acid (CHCA) for peptides and 3,5-dimethoxy-4-hydroxy-cinnamic acid (sinapinic acid) for proteins. In addition, a mixture of CHCA with another matrix, 2,5-dihydroxybenzoic acid (DHB) is also used for the analysis of proteins (9, 11).

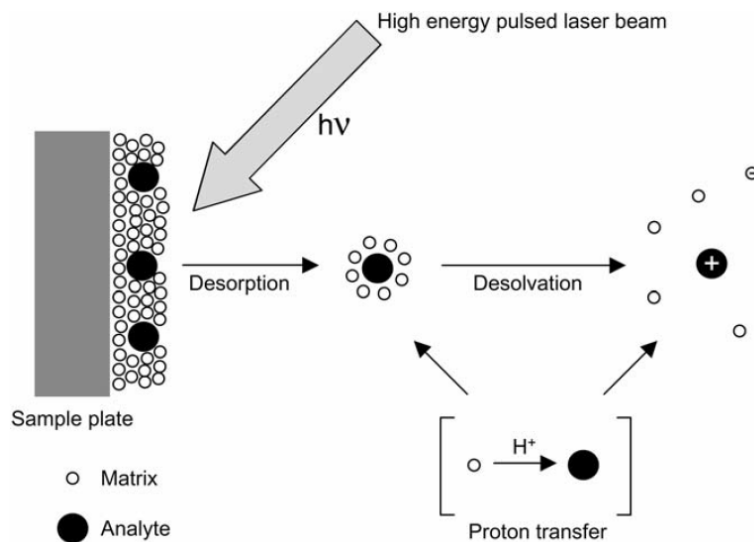


Figure 1.3. Formation of analyte ions by MALDI (11)

The other ionization technique, ESI has a different principle of producing molecular ions. In this technique, analyte solution is pumped into a tiny capillary via a syringe pump or a chromatographic system. As the analyte solution flows through the capillary, a high positive or negative voltage (+/-2-5 kV) is applied between the capillary tube and the inlet of the mass spectrometer, which then results in spraying of the analyte solution and formation of tiny, charged droplets. The droplets are driven to the inlet of the mass spectrometer by a potential difference and at the same time they are subjected to a curtain of heated nitrogen gas (N_2 , nebulizing gas) to evaporate the solvent. As the solvent evaporates, the charge remains constant and the droplets gradually shrink. At a certain point, repulsive forces, also called as coulombic forces, exceed the surface tension of the solvent and the droplets break up into smaller particles. The charged and desolvated analyte ions are then transferred into the mass spectrometer (Figure 1.4) (11, 13).

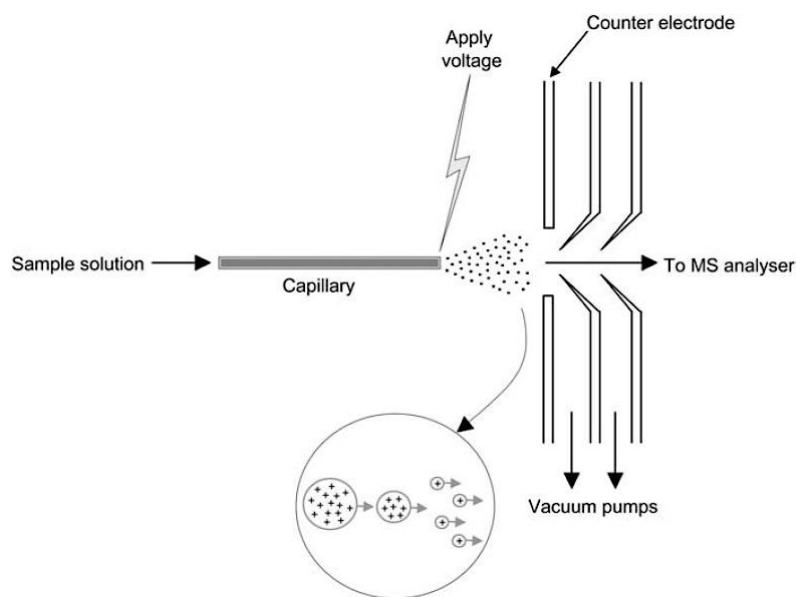


Figure 1.4. Formation of analyte ions by ESI (12)

Although both techniques perform MS analysis of biomolecules successively, they differ in the way that they create gas phase ions. The biomolecules are transferred into gas phase from solid phase in MALDI and from liquid state in ESI. This ability of ESI allows the direct coupling of liquid chromatography with ESI source of a mass spectrometer. Ionization of biomolecules by MALDI usually produces singly charged ions and allows the analysis of proteins with high or low molecular mass within a given mass range. In contrast to MALDI, ESI results in the generation of multiply charged ions of biomolecules. The mass spectrum obtained from ESI of a protein shows a distribution of different charge states. As the molecular mass is divided into different charges, peaks are observed at different m/z values in the low mass region. By this way, even biomolecules with high molecular weights can be analyzed without a limit of mass range and appear as multiply-charged species in the low mass region. As an example, mass spectra of cytochrome c obtained by MALDI and ESI techniques are given in Figure 1.5 and reflect the difference in charge state of molecular ions. Besides, MALDI offers advantages in analysis of biomolecules with its ability to tolerate the presence of contaminants (at moderate levels) like salts or detergents in sample. However, ESI cannot tolerate even the small amounts of such contaminants, since these molecules can be retained in capillary tubes of ESI and hamper the quality of MS signals.

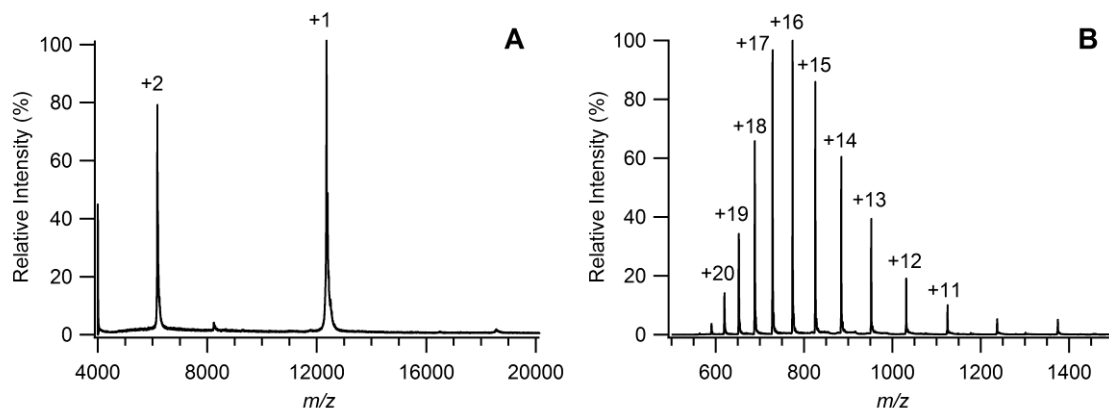


Figure 1.5. MS spectra of cytochrome c ionized by (a) MALDI, (b) ESI

As the analyte ions are formed in an ion source, they are directed into a mass analyzer in which they are separated according to their m/z values. The mass analyzers that are commonly used in proteomic studies are, time-of-flight (TOF), quadrupole, ion-trap, orbitrap and Fourier transform ion cyclotron resonance (FT-ICR). Different analyzers vary in mass accuracy, mass range, resolution, sensitivity, speed and principle of separating ions. In a TOF analyzer, the m/z values of the ions are determined by measuring the time that each ion takes to travel through a field-free flight tube. Initially, the ions are accelerated with the same potential and thus, ions with different molecular mass but with same charge and energy differ in the time they reach to the detector. Quadrupole mass analyzers consist of four parallel rods equally spaced around a central axis. Direct current (DC) and radio frequency (RF) voltages are applied to opposing pair of rods and one pair is set negative and the other positive. When the ions are introduced between rods along the central axis, a combination of RF and DC voltages is applied, which selectively stabilizes the trajectory of ions of particular m/z and allows these ions to reach to the detector. Ion trap analyzers work in a similar principle as quadrupole analyzers. It consists of a set of quadrupole electrodes; a ring electrode and two end cap electrodes. In ion traps, the ions are held inside a trap by electric field created by RF voltage and ions take an oscillating frequency related to their m/z . The field can be manipulated to selectively trap ions of a particular m/z and eject others. This ability of ion traps allows performing MS^n analysis. In a MS^n analysis, a specific molecular ion is selected and collided with an inert collision gas to induce fragmentation of the ion inside the trap and then a particular fragment ion is selected and further fragmented in the same trap. For MS^n analysis, this cycle is repeated sequentially (12). In an orbitrap mass

analyzer, ions are trapped and orbit around a central electrode and oscillate harmonically along its axis with a frequency characteristic to their m/z value. An image current is induced in the electrodes, which is then Fourier transformed to produce mass spectrum (14). Apart from these analyzers that were mentioned up to here, FT-ICR systems (also referred to as FT-MS) use a static magnetic field rather than an electric field to trap ions inside a cubic cell. In this cell, the ions are constrained by a strong magnetic field and they move in cyclotron motion. The motion of an ion is related to cyclotron frequency and its m/z value. Then by applying RF potentials, the cyclotron motion of ions with particular m/z is affected, due to an increase in the energy of these ions. The change in cyclotron motion and frequency induces an image current and is converted into a mass spectrum by applying Fourier transform (12).

The mass analyzers can be used singly in a mass spectrometer, or combined to create hybrid mass spectrometers to increase the performance of the instrument. For example, two separate TOF analyzers are combined (TOF/TOF) to perform tandem MS analysis. Moreover, TOF analyzers are combined with quadrupole analyzers. The hybrid instrument is referred to as Q-TOF systems, which can perform tandem MS analysis and provide higher mass resolution compared to only-quadrupole or only-TOF systems. Currently, FT-MS ranks in first place on the basis of its unmatched mass resolution (about 800,000). Orbitrap instruments also provide a high mass resolution (around 100,000), not as high as FT-MS systems but still higher than that of TOF, quadrupole, ion-trap or hybrid mass analyzers.

1.3.3. Protein Identification by Mass Spectrometry

Protein identification by MS is carried out either in the form of intact protein analysis (top-down) or analysis of peptides obtained from proteolytic digestion of proteins (bottom-up). An overview of these approaches is depicted in Figure 1.6. In the top-down approach, the whole protein is ionized and fragmented inside a high-resolution mass spectrometer (mainly FT-MS) in order to create fragment ions, which give information about protein sequence. In this approach, there is no need to perform protein digestion before MS analysis; the intact protein and its whole sequence are under examination. Therefore, it is easy to determine protein isoforms and post-translational modifications. Taking the advantage of high resolving power of FT-MS

instruments, simple protein mixtures can be introduced into mass spectrometer and fragmented separately, without a need for extensive protein separation or fractionation. In bottom-up approach, peptides derived from proteolytic or chemical digestion of proteins are analyzed in order to identify the corresponding proteins (10). In a classical bottom-up workflow, first of all the sample complexity is reduced by protein separation using either electrophoretic or chromatographic techniques. The proteins are then subjected to a digestion procedure with a protease. Trypsin is a widely used protease in proteomic studies, which cleaves proteins exclusively at C-terminal of arginine or lysine residues. Still other proteases exist with ability to cleave at different sites of a protein, creating peptide products of different length. Following digestion, the peptides can be analyzed directly by mass spectrometer for determination of m/z value of each peptide or namely peptide mass fingerprint. The recorded data is then submitted to a protein database in order to compare with the fingerprints of each entry in the database and the most probable proteins are retrieved. This method is known as peptide mass fingerprinting (PMF). Another way of analyzing peptides obtained from digestion is to separate these peptides by chromatographic techniques, mainly RP HPLC, and then introducing into mass spectrometer for tandem MS analysis. Briefly, tandem MS analysis enables the fragmentation of peptides in mass spectrometer to obtain partial amino acid sequences. The fragmentation pattern obtained from peptide by tandem MS is used to deduce amino acid sequence of the peptide via database search (9). Following technical developments in chromatography and mass spectrometry, a different kind of bottom-up approach has emerged and entitled as shotgun proteomics. In this approach, a complex protein mixture, generally an entire proteome isolated from cells or tissues, is subjected to proteolytic digestion without pre-fractionation. The peptides are then separated by one or multi dimensional separation techniques and analyzed by tandem MS. Shotgun approach enables the analysis of entire proteome in a simple manner and provides a high level of certainty by sequencing information rich peptides. However, it has some drawbacks. Because the digestion of whole proteome increases the complexity of the peptide mixture, some of the information may be lost and not all of the peptides can be identified by MS especially those with important modifications (14). All in all, developments in separation science, MS instrumentation and search algorithms allow efficient use of different proteomic approaches.

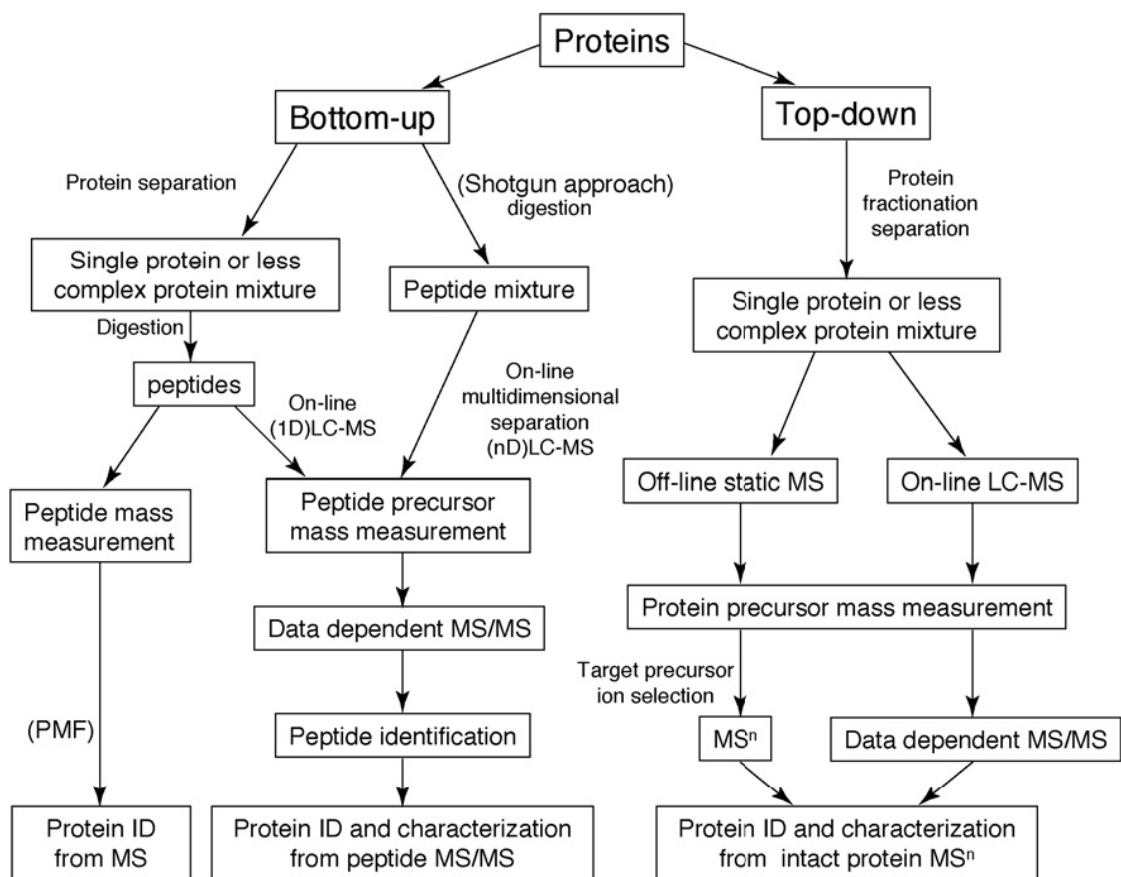


Figure 1.6. An overview of different strategies employed in proteomic studies for protein identification by mass spectrometry (14)

Tandem MS (MS/MS) is a key technique for sequencing of proteins and peptides by mass spectrometry. In order to perform MS/MS experiments, the mass spectrometers usually contain multiple mass analyzers connected in a series. This combination can either include same analyzers, such as TOF/TOF or triple quadrupole (QqQ), or include different type of analyzers, such as Q-TOF. First of all, an ion with a specific m/z (precursor ion) is selected in the first mass analyzer. Then, the precursor ion is focused into a collision cell where the ion is collided with an inert gas and collisions between the ion and gas result in fragmentation of the ion. During collisions, a portion of the kinetic energy of precursor ion is converted into vibrational energy, which induces fragmentation of specific bonds in the molecule. Finally, the m/z values of the resulting ions, called as fragment ions or product ions, are measured in the second mass analyzer. This process is called collision-induced dissociation (CID) and is the most widely used MS/MS technique in proteomic studies (11). In addition, this process can be repeated several times for further fragmentation and obtaining detailed structural

information, leading to MSⁿ experiments (n is the number of repeating MS experiments). Only several mass analyzers, such as ion trap and FT-ICR, can perform MSⁿ experiments without a need for an additional mass analyzer (15).

CID is routinely employed in proteomic studies where the aim is sequencing of proteins and peptides. The collision of peptide ions with gas molecules causes formation of fragment ions due the cleavage of certain chemical bonds in peptide. CID can be performed at low-energy conditions (10-100 eV) or high-energy conditions (above 1 keV) and the energy of the collisions determine which bonds in a peptide to be cleaved (11). Fragment ions are generated by cleavage of C α -C, C-N or N-C α peptide bonds and take different names depending on the cleaved bond and the location of charge following fragmentation. If the charge is retained on N-terminal side of the peptide, the fragment ions are called as *a*-, *b*- or *c*-ions. On the other hand, the fragment ions are referred to as *x*-, *y*- and *z*-ions if the charge stays at C-terminal (15). The nomenclature of peptide fragment ions (16, 17) is shown in Figure 1.7. The *a*-*x*, *b*-*y* and *c*-*z* are complementary ions and the subscripts indicate the number of amino acid residues. Under low-energy CID conditions, *a*-, *b*- and *y*-ions are formed by cleavage of amide bonds (C-N) (18). The high-energy CID process allows the cleavage of other backbone bonds and the formation of other types of ions. Moreover, it causes the dissociation of side chain bonds (19). Although high-energy CID provides more information when compared with low-energy CID, it results in the generation of complex CID spectra, which can be difficult to interpret. Fragmentation patterns of peptides and any type of resulting fragment ions contain sequence specific information and are very useful in determination of peptide sequence. In principle, the mass difference between consecutive ions in the same ion series is used to identify corresponding amino acids, as well as determination of modifications formed on amino acids either by post-translational processes or as artifacts during sample preparation. In proteomic studies, usually low-energy CID is employed and *b*-/*y*-ion series are used in determination the sequence of peptides and proteins. However, in some cases high-energy CID is required to obtain detailed information of chemical bonds. An example of such a case is the identification of leucine and isoleucine amino acids in a peptide. The molecular masses of these amino acids are exactly the same, so it is not possible to distinguish them on the basis of mass difference. Because these amino acids differ in connectivity of atoms in their side chains, only high-energy CID conditions can be used to distinguish these amino acids.

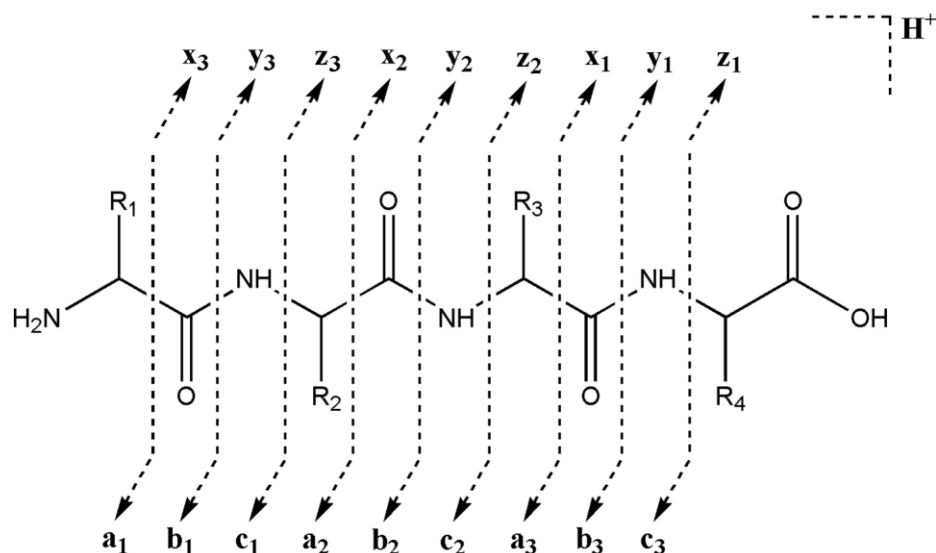


Figure 1.7. The nomenclature of peptide fragment ions (20)

The last step in proteomics workflow is the identification of peptides and proteins. To that end, MS or MS/MS data, acquired for each peptide in a single bottom-up experiment are searched against a protein database using a search engine. The search engine uses each MS or MS/MS spectrum to determine which peptide sequence in the protein database gives the best match. Several search engines have been developed for this purpose and Mascot, Sequest, MS-Tag, OMSSA, X!Tandem are some them. These search engines identify the best matching peptide on the basis of different algorithms. For example, Mascot uses probability based matching by comparing the calculated masses of peptides or MS/MS fragments, which are obtained by applying enzyme cleavage rules to each entry in protein database, to that of experimentally obtained peptides. As a result of this comparison, a score is calculated that is related with statistical significance of the match between experimental and theoretical masses. Sequest employs a cross-correlation algorithm in which peptide amino acid sequences from protein database are used to construct theoretical mass spectra and the best match is determined by overlap, or cross correlation of theoretical and experimental spectra (12). During database searching, additional information are required by search engines under the title of search parameters. These parameters include database, taxonomy, mass tolerance, enzyme and modifications. With the help of these parameters, users are allowed to perform search by selecting a suitable protein database, such as NCBI and Swissprot, and can restrict the search to entries for a particular organism. If the taxonomy information is available, specifying the search for a particular organism is

recommended because it speeds up the search and simplifies the results by eliminating homologous proteins from other species. The enzyme used in proteolytic digestion should be indicated in search parameters because the same digestion rules are applied to proteins in database for presumption of theoretical peptides. In addition, it is widely accepted that digestion mixtures contain not only expected peptide products of a perfect digestion but also some peptides with missed cleavage sites. Therefore, generally a setting of 1 or 2 missed cleavage sites for a specified enzyme is applied in order to include the presence of these partially digested peptides during database search. Besides these parameters, setting an appropriate mass tolerance value for both precursor and fragment ions allows to specify an error window on experimental masses of peptides and proteins. Another important parameter in search is to assign modifications on proteins. Proteins and peptides can be modified during post-translational processes or sample preparation. The modifications are classified into two types in search parameters, namely fixed and variable. Fixed modifications are applied to every specified residue or terminus. For example, the use iodoacetamide in alkylation of cysteines causes a mass shift of +57 Da and should be assigned as a fixed modification to be applied to all cysteine residues. Variable modifications are those who may present or not present. An example would be the oxidation of methionine residues. This modification can be generated in organism or can also be observed as artifacts in sample preparation. Therefore, for a peptide containing 3 methionine residues, assigning methionine oxidation as variable modification, the search engine will search the oxidation of 0, 1, 2, and 3 methionine residues and include all possible combinations of oxidized peptide (21). Once the MS or MS/MS data is submitted and appropriate search parameters are assigned, identified peptides are given as a protein hit-list. From this list, true positive results are retrieved among best matches by a careful examination of the results on the basis of score or significance threshold values and the number of peptides identified for a single protein. Usually one protein matching just a single peptide sequence is treated as suspect and it is recommended to consider proteins matching more than one peptide. In some cases, database searching fails in determination the sequence of peptide. This usually originates from the lack of sequence in the database. The limitation of database search can be partially solved by de novo sequencing in which the sequence of peptide is derived from peptide fragmentation data without using database.

1.3.4. Applications of Mass Spectrometry-Based Proteomics in Health Sciences

The advent of MS-based proteomics gives insight into the biology of organisms including human. Since the proteome of an organism is affected by the conditions of that organism, studying the entire proteome of given tissues or cells may give insights into the pathologic states of diseases and drug mechanisms by examining protein expression levels, protein post-translational modifications, subcellular protein localizations, protein-protein and protein-drug interactions. To have a better understanding in disease development and to be used in the drug development studies, proteomic approaches needed to be used for determining disease-related changes on protein structure and in cellular function. In an effort to speed up early diagnosis of diseases, elucidation of disease mechanisms and drug development, the researchers are continuously working on the discovery of biomarkers for neurological, nephrological, immunological, endocrinological, cardiovascular diseases and various types of cancer.

Biomarkers are molecules, such as proteins, peptides, small molecules or metabolites in organisms, which are considered as indicators of biological processes (22). Proteomics studies enable discovery of disease-related biomarkers by examining expression levels of proteins, protein modifications, functions of proteins, and protein localizations. Expression proteomics analyzes protein expression levels in a given tissue, cell or body fluid and aims to identify biomarkers or drug targets. This is mainly achieved through comparing the proteomes of samples obtained from patients and healthy individuals in order to find changes in relative abundances of proteins. To that end, 2D-GE and multidimensional separation techniques are employed for protein separation and fractionation, followed by mass spectrometry analysis for identification of proteins. Quantification of the level of change in protein expression can be simply achieved through 2D-GE and comparison of color intensity of protein spots in gels of samples obtained from different conditions. Another way is to employ quantitative proteomics technique in which the proteins are labeled with stable isotopes (^2H , ^{13}C , ^{15}N , ^{18}O). The labeling can be performed during protein synthesis by including i.e. ^{15}N in cell culture and the technique is referred to as stable labeling of amino acids in cell culture (SILAC) and categorized as metabolic labeling. Another option is the enzymatic labeling where ^{18}O is incorporated into proteins or peptides, during or after enzymatic

digestion. Chemical labeling targets reactive groups on side chains of amino acids, such as free cysteines in isotope-coded affinity tag (ICAT) approach or free amines in isobaric tags for relative and absolute quantification (iTRAQ) approach (10, 14). Specific softwares are then used to determine peptide ratios quantified for a specific protein. Structural proteomics aims to determine functions of proteins localized in a specific organelle and makes a correlation between protein function, structure and localization. Functional proteomics studies the functions of proteins and regulation of their expression in biological systems (23). The change of protein function upon protein modification is a hot topic in this area. After translation and synthesis, proteins can undergo post-translational modifications, which can alter their properties and structure hence alter their function. There are a great number of post-translational modifications observed in living systems; phosphorylation, glycosylation, acetylation, methylation, ubiquitination and oxidation are only some of them. Several post-translational modifications, such as phosphorylation, are reversible, physiological and required in cellular regulation. However, some of the modifications, such as oxidation, are irreversible and may be associated with different conditions of organisms such as diseases. Therefore, the study of disease related post-translational modifications could help to elucidate the cellular mechanism of diseases. Mass spectrometry enables the identification and determination of post-translational modifications as well as the position of modification. Unfortunately, in some cases, the modified proteins may be underrepresented due to their low initial amounts in a given cell or tissue and this would in turn cause difficulties in identification of these proteins. To overcome this problem, appropriate enrichment techniques are employed prior to mass spectrometry analysis. Coupling affinity chromatography techniques with mass spectrometry is a common strategy for the enrichment and subsequent identification of modified proteins. By this way, mass spectrometry becomes a very powerful tool in the analysis of post-translational modifications. Consequently, proteomic studies enable the identification of disease related biomarkers, which could be used in disease diagnosis, staging and treatment monitoring, as well as possible targets for new drugs (23).

CHAPTER 2

INVESTIGATION OF DRUG RESISTANCE IN IN VITRO AND CLINICAL ISOLATES OF ANTIMONY- RESISTANT *L. Tropica* BY PROTEOMIC ANALYSIS

2.1. Introduction

Leishmania, protozoan parasites of the Trypanosomatidae family, are the causative agents of a complex group of diseases called leishmaniasis (24). More than 20 species of *Leishmania* are responsible for the varying forms of leishmaniasis. The disease is transmitted by the bite of a sandfly. The nature and severity of the disease depends on causative *Leishmania* species. One form of the disease is cutaneous leishmaniasis (CL), which is caused by *L. major*, *L. tropica* and *L. mexicana* and manifests as localized cutaneous lesions. A classical lesion starts as a nodule or papule at the site of sandfly bite, grows slowly, then generally self-heal and leads to subsequent immunity. Mucocutaneous leishmaniasis (MCL), in which the parasite causes metastasis to the mucosal tissues of the mouth and upper respiratory tract, is observed if the infecting species is *L. braziliensis* or *L. panamensis*. The other form, visceral leishmaniasis (VL), is caused by *L. donovani* and *L. infantum* (25, 26). In visceral leishmaniasis, the parasites disseminate from the site of sandfly bite to visceral organs resulting in the enlargement of lymph glands, liver and spleen that is associated with fever, wasting, anemia and weight loss. Visceral leishmaniasis is lethal in almost all untreated cases (27). Besides, genetic exchange between different *Leishmania spp.*, such as *L. infantum* and *L. major*, has been reported in literature (28). The genetic exchange resulted in the generation of hybrid strains, which are responsible for the different clinical manifestations and tissue tropisms.

Leishmania have dimorphic life cycle, namely promastigote and amastigote stages (Figure 2.1). Promastigote is the stage in which the parasite is found in the midgut of *Phlebotomus* or *Lutzomyia* sandfly vector as an extracellular flagellated form, whereas amastigote stage is characterized as aflagellated intracellular form in the

macrophages of the mammalian host. The life cycle of *Leishmania* is initiated when sandfly bites an infected mammal (rodents, dogs or humans). Promastigotes increase in number and undergo several developmental stages in the sandfly vector and then transmitted to other mammals by the bite of the sandfly. In the mammalian host, promastigotes are rapidly engulfed by phagocytotic cells, where the differentiation of promastigotes into amastigotes is triggered. The amastigote forms of the parasite reside within the host macrophage. Because *Leishmania* have dimorphic life cycle, they are adapted to live in different conditions of pH and temperature of the different hosts. Thus, they have developed strategies to evade digestion enzymes of the sandfly and host immune responses (24).

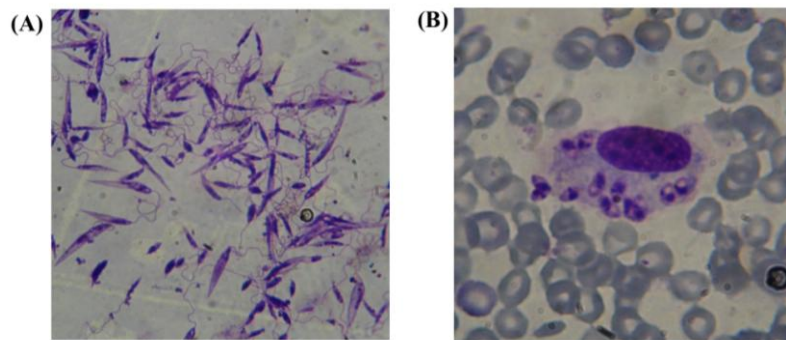


Figure 2.1. Microscope images of *Leishmania* (a) promastigotes and (b) amastigotes

Leishmaniasis represent a major public health risk and is widespread in a total of 98 countries or territories on five continents; Asia, Europe, Africa, North America and South America. According to World Health Organization (WHO), more than 350 million people are at risk with an estimation of 12 million people infected with *Leishmania*, 2 million new cases (0.5 million of visceral leishmaniasis and 1.5 million of cutaneous leishmaniasis) reported each year and 50,000 deaths. From the data available, WHO estimated that 90% of cases of visceral leishmaniasis occur in Bangladesh, Brazil, Ethiopia, India, Nepal and the Sudan; whereas up to 90% of the cases of cutaneous leishmaniasis occur in Afghanistan, Algeria, Brazil, Colombia, the Islamic Republic of Iran, Nicaragua, Peru, Saudi Arabia and the Syrian Arab Republic (26). In Turkey, leishmaniasis poses an emerging threat to public health, as well. Cutaneous leishmaniasis is mainly caused by *L. tropica* and frequently observed in southeastern Anatolia, central Anatolia, western regions, and less frequently in Aegean

and Mediterranean regions. On the other hand, visceral leishmaniasis, in which the causative species is *L. infantum*, occurs mainly in Aegean, Mediterranean and central Anatolia regions. Some cases of cutaneous leishmaniasis caused by *L. infantum* have also been reported in the eastern Mediterranean region (29). Ministry of Health reported 46,000 new cutaneous leishmaniasis cases between 1990 and 2000, 96% of which were from southeastern Anatolia region (30). The global incidence of leishmaniasis increases due to the lack of effective vaccines against disease, problems in controlling the vector and emergence of drug-resistant parasites (31).

The infecting *Leishmania* species, as well as host genetic background and immune status, are the main factors for the determination of severity and nature of leishmaniasis. Early diagnosis and appropriate treatment are key steps for the control of the disease. The diagnosis is mainly achieved by microscopic examination of tissue aspirate or lesion biopsy samples and cultured material in specific medium. Serological tests, such as ELISA (enzyme-linked immunosorbent assay) or Western blot, and molecular methods such as PCR (polymerase chain reaction), are also used for diagnosis of leishmaniasis (26, 30). Besides, a recent study reported the use of MALDI-TOF MS as a rapid and reliable tool for the identification of *Leishmania* species (30). In treatment of leishmaniasis, chemotherapy represents the only way, since there is no vaccine available against the disease. Pentavalent antimonials (SbV), such as meglumine antimoniate (Glucantime) and sodium stibogluconate (Pentostam), are the first-line drugs for treatment of all clinical cases of leishmaniasis (32). Other drugs are used as second-line treatment and include pentamidine, amphotericin B and miltefosine. However, their usage is limited due their side effects and high cost. Pentavalent antimonials have been used for more than 70 years, but the treatment is challenged due to emergence of antimony resistant *Leishmania* species (33).

Antimony resistance in *Leishmania* has been extensively studied so as to unravel the molecular mechanisms of resistance and improve effective treatment strategies. Resistance involves numerous mechanisms including uptake, metabolism, efflux and/or sequestration of active drug molecules (34). Primarily, it is accepted that SbV is a prodrug and needs to be reduced to active form, trivalent antimony (SbIII), in order to be effective on both stages of the parasite (35, 36). This reduction is assumed to proceed via two pathways within the parasite: non-enzymatic and enzymatic reduction. Non-enzymatic reduction involves the use of parasite thiols such as spermidine, cysteine, glutathione (GSH) and trypanothione (TSH) (34). For enzymatic reduction, two parasite

specific enzymes, which require GSH as reducing agent, have been identified: thiol dependent reductase 1 (TDR1) (37) and arsenate reductase (LmACR2) (38). The reduction of SbV to SbIII is achieved not only in parasite, but also in the macrophage. GSH is the primary thiol that takes part in the reduction of SbV, but this reduction is pH dependent, having higher conversion rates at acidic pH, and takes part in acidic compartments of macrophage such as endosome, lysosome and phagolysosome (39). It has been shown that, exposure to SbIII inhibits trypanothione reductase (TR) (40) and glutathione synthetase (41). In addition, it results in the efflux of TSH and GSH from promastigotes and amastigotes (42). In order to be active, antimony has to enter macrophage and act on parasite. The route of entry of antimonials is still unclear and suggested to be not the same for SbV and SbIII (33). It was reported that the entry of pentavalent arsenite (AsV), a metal related to SbV, is achieved via a phosphate transporter (43). On the other hand, aquaglyceroporin (AQP1), which takes part in transport of trivalent metalloids in prokaryotes and eukaryotes, has been identified in *Leishmania* and associated with SbIII uptake (44). Later on, it was shown that expression of the gene coding AQP1 was down regulated in antimony-resistant clinical isolates (45). The efflux and/or sequestration of drug molecules through membranous transport systems are both common strategies for drug-resistance observed in several microorganisms. Proteins that belong to ATP Binding Cassette (ABC) family have important roles in efflux and sequestration of drug molecules (34). Multidrug Resistance-related Protein of *Leishmania* (MRPA) is an ABC transporter but unlike other ABC type proteins it does not act by efflux of the drug; rather it acts by sequestration of the SbIII-TSH complex into intracellular vesicles (46). Overexpression of the gene encoding MRPA has been linked to a decrease in the influx of antimony, rather than increase in the efflux, therefore accounting for drug resistance (47). Another type of ABC transporter protein that is associated with antimony resistance is PRP1. Although its exact function is unknown, it was reported that the overexpression of this gene leads to resistance to antimonials (48). Maintenance of intracellular reducing environments via thiols has a major role in drug resistance, where the parasite protects itself against damage caused by oxidative stress created inside the macrophage by oxidants and xenobiotics (49). Antimony resistant *Leishmania* species were found to have increased levels of intracellular thiols (50). TSH is a trypanosomatide specific thiol and is formed by conjugation of spermidine and GSH (51). Overexpression of genes encoding ornithine decarboxylase (in the biosynthesis of spermidine) (52) and

gamma glutamylcysteine synthetase (in the biosynthesis of GSH) (53) has been reported in arsenite resistant *Leishmania*. Moreover, it was shown that inhibition of these genes by specific inhibitors reversed the arsenite or antimony resistance of *Leishmania* and implies the importance of thiol levels for drug resistance (52-55). In addition to all these, there are some other mechanisms of antimony resistance in *Leishmania*. It is known that antimonials kill parasites by inhibiting the enzymes of metabolic pathways, by effects on thiol redox potential, DNA fragmentation and by inducing synthesis of proteins involved in programmed cell death (PCD) (33). It has been shown that, differential expression of some proteins, such as 14-3-3 protein, small kinetoplastid calpain-related protein 1.14 (SKCRP1.14) and heat shock proteins (HSPs), were related with programmed cell death. 14-3-3 protein, which has the ability to bind phosphorylated proteins in apoptosis, was over expressed and SKCRP1.14, which is often associated to several key effector proteins of PCD, was down regulated in drug resistant *L. donovani* (56). These results imply that the process of cell death in drug resistant strains is impaired. Although HSPs have already been reported to be involved in cell death by modulating some steps of apoptosis pathway, the increase in the production of a variety of HSPs (HSP70, HSP83 and HSP65) in drug resistant strains is mainly associated with a protective mechanism against the toxic effects of drugs (34).

In addition to the genomic studies, proteomic studies involving the comparison of proteome of antimony resistant and sensitive parasites were conducted. Up to date, comparative proteomic analyses of antimony resistant and sensitive *L. braziliensis* (57, 58), *L. infantum* (32, 58, 59), *L. donovani* (56, 60, 61), *L. tropica* (62), and *L. panamensis* (63) were performed using different proteomic approaches. These studies enabled the identification of a great number of differentially expressed (up or down regulated) proteins in antimony resistant species, as summarized in Table 2.1. The identified proteins are distributed into seven biological process categories: cytoskeletal, metabolic, transport, antioxidant/detoxification, protein folding/stress response, protein biosynthesis and RNA/DNA processing. In this table, the change in expression is abbreviated as UR for up regulated and DR for down regulated proteins. Detailed examination of this table highlights some proteins that showed a common trend in their expression status from different antimony resistant *Leishmania* species. Of the metabolic enzymes, glyceraldehyde 3-phosphate dehydrogenase, fructose 1,6-bisphosphate aldolase, aldehyde dehydrogenase, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase, enolase, and isocitrate dehydrogenase together with other

important metabolic enzymes, were up regulated. Up regulation of glycolysis and TCA cycle enzymes shows that the energy metabolism is increased in antimony resistant strains, which provides energy for their proliferation. The abundance of trypanothione peroxidase, pterin reductase, HSPs (HSP 70, HSP 83, HSP 60), MRPA and ATPases implies that parasites are trying to cope with the toxic effects of drug either by efflux or sequestration of drug molecules via transporters, creating a reducing environment and protection via antioxidants and heat shock proteins. Moreover, the increase in the abundance of ribosomal proteins, proteins involved in translation machinery such as elongation factor 2, and DNA processing proteins such as proliferative cell nuclear antigen, which takes part in replication and DNA repair, suggests an increase in replication, DNA repair and protein biosynthesis in antimony resistant *Leishmania* species. As a result, it is obvious that proteomic studies together with genomic studies are powerful approaches to study drug resistance and made important contributions to elucidate the mechanism of drug resistance in *Leishmania*.

The emergence of antimony resistant *Leishmania spp.* is a main obstacle in the treatment of leishmaniasis. A literature survey on the proteomic analysis of antimony resistance showed that the studies were mainly performed by investigating differentially expressed proteins of *in vitro* antimony resistant lines and clinical resistant isolates, in comparison to corresponding sensitive pairs. According to the results of these studies, it has been suggested that antimony resistance in *Leishmania*, generated *in vitro* and in host organism, may follow similar mechanisms, albeit some differences could exist due to the differences in conditions that the parasite is exposed in host organism and culture.

In this study, the aim is to perform a comprehensive proteomics analysis of antimony resistance in *Leishmania*. To that end, the proteome of *in vitro* antimony resistant lines and antimony resistant clinical isolate of *L. tropica* were compared with antimony sensitive isolates. Moreover, *in vitro* resistant lines and clinical resistant isolate were compared to gain more insights into the different resistance phenotypes. The results of this study can make important contributions to understand the mechanism of resistance and can be used to improve the current treatment methods of leishmaniasis, as well as to develop more efficient therapeutics.

Table 2.1. Differentially expressed proteins in Sb-resistant *Leishmania spp.*

Protein Function	Protein Identity	Expression
Antioxidant/ detoxification	Tryparedoxin peroxidase	UR (57, 58)
	Pterin reductase 1	UR (57, 58)
	Trypanothione reductase	DR (58)
	3-hydroxacyl-ACP dehydratase, MaOC	UR (58)
	Prostaglandin f2-alpha synthase	DR (62), UR (63)
	Kinetoplastid membrane protein-11	DR (57, 59)
	Peroxidoxin	UR (57)
Cytoskeletal	Alpha tubulin	DR (57, 62) UR (58, 61, 63)
	Beta tubulin	DR (57, 59), UR (58, 60, 61)
	Actin-like protein	DR (58)
	Paraflagellar rod protein	DR (57, 58)
	Kinesin	UR (32)
Intracellular survival/ processing	Nascent polypeptide associated complex subunit-like protein	DR (58), UR (61)
	Metallo-peptidase, Clan MF	UR (58, 61)
	Metallo-peptidase, Clan MA(E), Family 32	UR (57)
	Proteasome regulatory ATPase subunit	UR (61)
	Mannose-1-phosphate guanyltransferase	UR (61)
	IgE-dependent histamine releasing factor	DR (61)
	Protein phosphatase	DR (58, 61)
	GP63, leishmanolysin, metallo-peptidase, Clan MA(M), Family M8	DR (32)
	calpain-like cysteine peptidase	DR (32)
	Proteasome beta 5 subunit	UR (32)
	Proteasome alpha 5 subunit	UR (60)
	Proteasome alpha 1 subunit	DR (57)
	Protein disulfide isomerase	DR (62)
	GPI protein transamidase	UR (60)
	Carboxypeptidase	UR (60)
Oligopeptidase b	UR (63)	
Small kinetoplastid Calpain-related protein	DR (56)	
Protein folding/ chaperones and stress proteins	Calmodulin	UR (61)
	Heat shock protein hsp70	UR (32, 56, 57, 58, 60, 61)
	T-complex protein 1	UR (61)
	Chaperonin Hsp60	UR (32, 58), DR (61)
	Heat shock protein 83-1	UR (56, 57, 58, 60)
	Glucose-regulated protein	DR (58)
	Cyclophilin a	DR (58)
	Stress induced protein sti1	DR (58), UR (63)
	Heat shock protein DNAJ	DR (32)
	Calreticulin	DR (57)
TPR domain protein	DR (57)	

(cont. on next page)

Table 2.1 (cont).

Protein Function	Protein Identity	Expression
Protein biosynthesis	60S ribosomal protein	DR (32, 61), UR (60, 61)
	40S ribosomal protein	UR (32, 57, 61), DR (61)
	Elongation factor 1-alpha	UR (32, 61), DR (58)
	Ribosomal protein s29, L1a, I3	UR (61)
	Elongation factor 2	DR (57), UR (58, 61, 63)
	Translation elongation factor-1-beta	UR (61)
	Eukaryotic translation initiation factor 2	DR (61)
	Eukaryotic translation initiation factor 1	UR (57)
	Seryl-tRNA synthase	UR (58)
	Small ubiquitin protein	UR (58)
	Argininosuccinate synthase	UR (57, 59)
	Eukaryotic initiation factor 4a	DR (57)
	Lysyl-tRNA synthetase	DR (57)
Metabolic	Triose phosphate isomerase	UR (61)
	Glyceraldehyde 3-phosphate dehydrogenase	UR (32, 61)
	Pyruvate kinase	DR (57, 59), UR (61)
	Fructose 1,6-bisphosphate aldolase	UR (60, 61)
	Glycosomal malate dehydrogenase	UR (61)
	Fumarate hydratase	UR (61)
	Glucose-6-phosphate 1-dehydrogenase	UR (61)
	Aldose 1-epimerase	UR (61)
	Short chain dehydrogenase	UR (61)
	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	UR (32, 61)
	Aldehyde dehydrogenase	UR (32, 61)
	Glycosomal phosphoenolpyruvate carboxykinase	DR (61)
	Glycosamine-6-phosphate isomerase	DR (61)
	Nucleoside diphosphate kinase b	UR (57), DR (61)
	Enolase	DR (56), UR (58, 60)
	Rieske iron-sulfur protein precursor	UR (58)
	Aminoacylase	DR (58)
	Nitrilase	DR (58)
	Glutamine synthase	DR (58)
	Cathepsin L-like protease	DR (58)
	Pyruvate dehydrogenase E1 beta subunit	DR (58)
	Inorganic pyrophosphatase	UR (58)
	Isocitrate dehydrogenase	UR (32, 58)
	S-adenosyl methionine synthase	UR (58, 63)
	S-adenosylhomocysteine hydrolase	DR (57), UR (61, 63)
	2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase	UR (58)
	Biotin/lipoate protein ligase-like protein	UR (58)
	3-hydroxyisobutyryl-coenzyme A hydrolase-like protein	UR (58)
	Long chain fatty acid CoA ligase	UR (32)
	protoporphyrinogen oxidase-like protein	DR (32)
	Pyruvate phosphate dikinase	UR (32)
	Sterol 14-alpha-demethylase	UR (32)

(cont. on next page)

Table 2.1 (cont).

Protein Function	Protein Identity	Expression
Metabolic	NADH-cytochrome B5 reductase	UR (32)
	Hexokinase	UR (32)
	Fumarate hydratase	UR (32)
	Succinate dehydrogenase flavoprotein	UR (32)
	Formate-tetrahydrofolate ligase	UR (32)
	GCVL-2 dihydrolipoamide dehydrogenase	UR (32)
	2-oxoglutarate dehydrogenase E1 component	UR (32)
	Activated protein kinase c receptor (LACK)	DR (59), UR (62)
	Succinyl-CoA ligase beta-chain	DR (57), DR (58)
	STP-binding cassette sub-family F member 1	DR (57)
	Transitional endoplasmic reticulum ATPase	DR (57), UR (61)
	14-3-3 protein-like protein	UR (56), DR (57)
	Acidocalcisomal pyrophosphatase	UR (57)
	Proteasome activator protein	DR (59)
RNA/DNA processing	Histone H2B	UR (61)
	Histone H4	UR (32)
	High mobility group protein homolog tdp-1	UR (61)
	ATP-dependent RNA helicase	UR (61)
	RNA-binding protein	DR (58), UR (61)
	Basic transcription factor 3a	UR (61)
	Proliferative cell nuclear antigen	UR (57, 58, 60)
	Ran-binding protein	UR (57)
	UV excision repair RAD23-like protein	UR (57)
	Nucleosome assembly protein-like protein	UR (57)
RNA helicase	DR (57)	
Transport	ATPase alpha subunit	UR (61)
	ATPase beta subunit	UR (58, 63)
	Vacuolar ATP synthase subunit	DR (58)
	Vacuolar ATPase subunit like protein	UR (32)
	FT1 folate/biopterin transporter	DR (32)
	MRPA (ABCC3) ABC-thiol transporter	UR (32, 60)
	Vacuolar-type proton translocating pyrophosphatase 1	DR (32)
	ER-golgi transport protein p24	UR (32)
	SEC61-like (pretranslocation process) protein	UR (32)
	Vesicular transport protein	DR (62)
	GTP-binding protein	UR (63)
	Rab7 GTP binding protein	UR (58)

2.2. Experimental Methods

2.2.1. *Leishmania* Samples and Parasite Growth

In this study, antimony sensitive and resistant *L. tropica* were used. Dr. Ahmet Özbilgin and his group in Celal Bayar University Medical School performed parasite isolation, growth and protein extraction steps.

The parasites used in this study were categorized into three types and coded as DNI, DI and D. DNI type refers to clinical isolate of antimony sensitive *L. tropica*. This type includes; DNI1, DNI2 and DNI3 coded promastigotes, which were isolated from skin lesions of patients with cutaneous leishmaniasis in Turkey. These patients have never received treatment with any type of anti-leishmanial drugs before isolation of the parasite from patient. Following isolation, meglumine antimoniate treatment of these patients resulted in the healing of the cutaneous lesions. DI type refers to *in vitro* antimony resistant *L. tropica* and includes DI1, DI2, and DI3 coded promastigotes. Antimony resistant DI type promastigotes were generated *in vitro* from DNI type, via treatment with increasing concentrations of meglumine antimoniate. D type refers to clinical isolate of antimony resistant *L. tropica*, and included D1 coded promastigote. D1 was isolated from a skin lesion of a patient with cutaneous leishmaniasis in Turkey, who has received three courses of antimony treatment and was found to be unresponsive to the treatment. Clinical *L. tropica* isolates (DNI and D type) were first cultivated in NNN medium and then transferred into RPMI-1640 medium supplemented with 10% FCS, 200 U Penicillin/mL and 0.2 mg Streptomycin/mL. The generation of *in vitro* antimony resistant *L. tropica* (DI type) was performed by treatment of DNI type promastigotes with increasing concentrations (10-300 mg/mL) of meglumine antimoniate (Glucantime). Cell growth was performed at 25 °C and promastigotes were harvested by centrifugation at 4400 rpm, for 10 min at 4°C.

2.2.2. Total Protein Extraction

Cells were washed 3 times with PBS buffer, resuspended in 1 mL of Mammalian Cell Lysis Reagent (Fermentas Life Sciences) and incubated at room

temperature on a shaker. Following, cell lysate was centrifuged at 14,500 rpm for 15 min and supernatant was stored at -20 °C for further analysis. Protein extracts were received from Celal Bayar University in cell lysis reagent. In order to concentrate total protein from *Leishmania* and exchange cell lysis reagent with two-dimensional gel electrophoresis rehydration buffer, proteins in supernatant were subjected to acetone precipitation protocol. Briefly, 600 µL of supernatant was mixed with four volumes of ice-cold acetone and incubated at -20 °C for 2 h. Following incubation the mixture was centrifuged at 14,000 rpm for 20 min at 4 °C. The supernatant containing acetone fraction was discarded without disturbing protein pellet at the bottom and the pellet was solubilized in 200 µL of urea-thiourea solution (7 M urea, 2 M thiourea). Protein concentration was determined by Bradford method (Pierce, Rockford, IL).

2.2.3. Two-Dimensional Gel Electrophoresis

For two-dimensional gel electrophoresis experiments, protein concentration of each sample was adjusted in rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 65 mM DTT, 2.5% pH 3-10 ampholyte solution) to a final concentration of 1 mg/mL. The first dimension, isoelectric focusing, was performed with non-linear pH 3-10, 17 cm IPG strips (Bio-Rad) using a Protean IEF cell (Bio-Rad). 350 µg of total protein in 350 µL of rehydration buffer was loaded on a strip; passive rehydration was performed for 2 h at room temperature followed by an active rehydration at 50 V for 16 h. The strips were then subjected to a total of 60000 Vh of electrophoresis. IEF steps were as follows; step 1: linear increase, 200 V, 300 Vh; step 2: linear increase, 500 V, 500 Vh; step 3: linear increase, 1000 V, 1000 Vh; step 4: linear increase, 4000 V, 4000 Vh; step 5: rapid increase, 8000 V, 24000 Vh; step 6: rapid increase, 8000 V, 30000 Vh. The second dimension, SDS-PAGE, was performed using Protean II XL System (Bio-Rad). Prior to SDS-PAGE, the strips were equilibrated first in equilibration buffer I (6 M urea, 0.375 M Tris-HCl pH 8.8, 2% SDS, 20% glycerol, 2% DTT) for 15 min and then in equilibration buffer II (6 M urea, 0.375 M Tris-HCl pH 8.8, 2% SDS, 20% glycerol, 2.5% iodoacetamide) for an additional 15 min. The strips were placed onto 12% polyacrylamide gels prepared according to the Laemmli Buffer System (Appendix A) and then sealed with an agarose solution (0.5% agarose, 0.003% bromophenol blue in 1X Tris-Glycine-SDS Buffer). Electrophoresis was carried out at 16 mA for the first

hour and then at 180 V until the dye front reached the bottom of the gel. Following electrophoresis, gels were washed with distilled water and stained with colloidal Coomassie Brilliant Blue G-250.

2.2.4. In-Gel Digestion and Peptide Extraction

Protein spots were excised from gel with a clean pipette tip, placed into clean micro centrifuge tubes and subjected to in-gel digestion protocol. First, each gel was divided into smaller pieces and 200 μ L of wash solution (50% methanol, 5% acetic acid in MilliQ water) was added on to each sample. The gel pieces were rinsed with wash solution on a shaker at room temperature, by several changes of the solution until all the dye is removed from gel pieces. The wash solution was removed, 200 μ L of acetonitrile was added and gel pieces were dehydrated at room temperature for 5 min. Acetonitrile was removed and gel pieces were completely dried in a vacuum centrifuge for 2-3 min. Then, 30 μ L of 10 mM DTT solution, prepared in 100 mM ammonium bicarbonate solution, was added on each sample and proteins were reduced for 30 min at room temperature. Following incubation, DTT solution was replaced with 30 μ L of 100 mM iodoacetamide solution, prepared in 100 mM ammonium bicarbonate solution, and alkylation was performed for an additional 30 min at room temperature. Iodoacetamide solution was removed from each sample, 200 μ L acetonitrile solution was added and gel pieces were dehydrated for 5 min at room temperature. Rehydration of gel pieces was achieved by addition of 100 mM ammonium bicarbonate solution and incubation of gel pieces for 10 min at room temperature. Dehydration step was repeated for the last time and then gel pieces were completely dried in a vacuum centrifuge for 2-3 min. Trypsin solution was prepared in 50 mM ammonium bicarbonate solution at a final concentration of 20 μ g/mL and 20 μ L of this solution was added to each sample. Digestion was carried out overnight at 37 °C. Peptides, obtained from digestion were extracted from gel pieces in a three-step procedure. First, 30 μ L of 50 mM ammonium bicarbonate solution was added to each sample and incubated for 10 min with occasional gentle vortex mixing. The sample was centrifuged and the supernatant was transferred into a clean micro centrifuge tube. Then, 30 μ L of extraction solution (50% acetonitrile, 5% formic acid in MilliQ water) was added to each sample and incubated for 10 min with occasional gentle vortex mixing. The supernatant after centrifugation

was combined with the former extract. The last step was repeated once more, and the volume of peptide extract was reduced to 10 μ L in a vacuum centrifuge.

2.2.5. MALDI-MS/MS Analysis and Database Search

Before mass spectrometry analysis, each sample was desalted with C18-ZipTip (Millipore). Mass spectrometry experiments were conducted on an Autoflex III smartbeam MALDI TOF/TOF MS (Bruker Daltonics, Germany) instrument. Two-layer matrix preparation method was employed, using alpha-cyano-4-hydroxycinnamic acid (CHCA) as matrix. The first layer consisted of 0.6-1.2 mg CHCA dissolved in 100 μ L solution of 20% methanol in acetone and the second layer consisted of 1 mg CHCA dissolved in 100 μ L of 40% methanol in 0.1% TFA solution. One μ L of first layer was spotted onto MALDI target plate (MTP-384 massive target gold plated T, Bruker Daltonics), followed by spotting 1 μ L of peptide-matrix mixture, mixed at a ratio of 1:5, onto first layer and left air drying. The mass spectra of each sample were recorded in reflectron positive ion mode and from each mass spectrum; peptides were selected for fragmentation in MS/MS mode using argon as a collision gas. The data obtained from MS/MS analysis were searched against NCBIprot database (selecting other eukaryotes taxonomy) using Mascot search engine (V2.4, Matrix Science) for identification of the corresponding proteins. The search parameters were set as follows; 200 ppm mass tolerance in MS mode, 0.5 Da mass tolerance in MS/MS mode, cysteine carbamidomethylation as fixed modification, methionine oxidation as variable modification and allowing up to 1 missed cleavage.

2.3. Results and Discussion

This study conducted a comprehensive proteomic analysis of antimony resistance in *L. tropica*. For that, first of all comparative proteomic analyses between sensitive *L. tropica* isolates (DNI1, DNI2, DNI3) and their *in vitro* generated resistant lines (DI1, DI2, DI3) were performed. Secondly, proteome of antimony resistant clinical isolate of *L. tropica* (D1) was compared with sensitive isolate (DNI3). Finally, proteome of antimony resistant clinical isolate (D1) was compared with *in vitro* generated resistant lines (DI1, DI2, DI3). Previous reports on the proteomic analysis of

antimony resistance in *Leishmania spp.* are mainly based on the comparison of sensitive isolates with *in vitro* selected resistant lines or with resistant clinical isolates. However, none of them combined the comparative proteomic analysis of *in vitro* generated and clinical resistant isolates of *Leishmania spp.* Therefore, the current study was conducted in order to elucidate differences and similarities of antimony resistance in *L. tropica*, gained either in laboratory conditions or in host organism, and in the hope of finding new insights into resistance mechanism.

The 2D-GE images and pairwise comparisons of the gels are given in Figures 2.2, 2.3 and 2.4. Considering significant differences in their expression levels between compared pairs, a total of 119 protein spots were selected. Out of 119 spots, 48 were identified as distinct proteins via MS/MS analysis and database search. The differentially expressed proteins are given in Tables 2.2 - Table 2.8, and in each result table, up regulated, down regulated and newly synthesized proteins were abbreviated as UR, DR and N, respectively. The identified proteins were classified into seven categories on the basis of their biological activities as follows: antioxidant/detoxification, cytoskeletal proteins, metabolic enzymes, protein biosynthesis, protein folding/chaperons and stress proteins, RNA/DNA processing and hypothetical proteins. Moreover, the relative distribution of identified proteins according to their biological activities were given as pie diagrams in Figure 2.5. These diagrams reveal that approximately 50% of the identified proteins were ascribed to the metabolism group in all comparisons. Clinical resistant isolate showed altered levels of antioxidant/detoxification process and chaperons/stress proteins. On the other hand, *in vitro* resistance resulted in the differential expression of hypothetical proteins, chaperons/stress proteins, proteins involved in antioxidant/detoxification process and protein biosynthesis. Moreover, these diagrams manifest some differences in the relative distribution of identified proteins among *in vitro* resistant lines. This might be attributed to the possibility of the generation of hybrid strains of *L. tropica* obtained from different patients.

The identified proteins as a result of comparisons, and their possible roles in resistance mechanism on the basis of their biological functions are discussed individually for each category in the following sections.

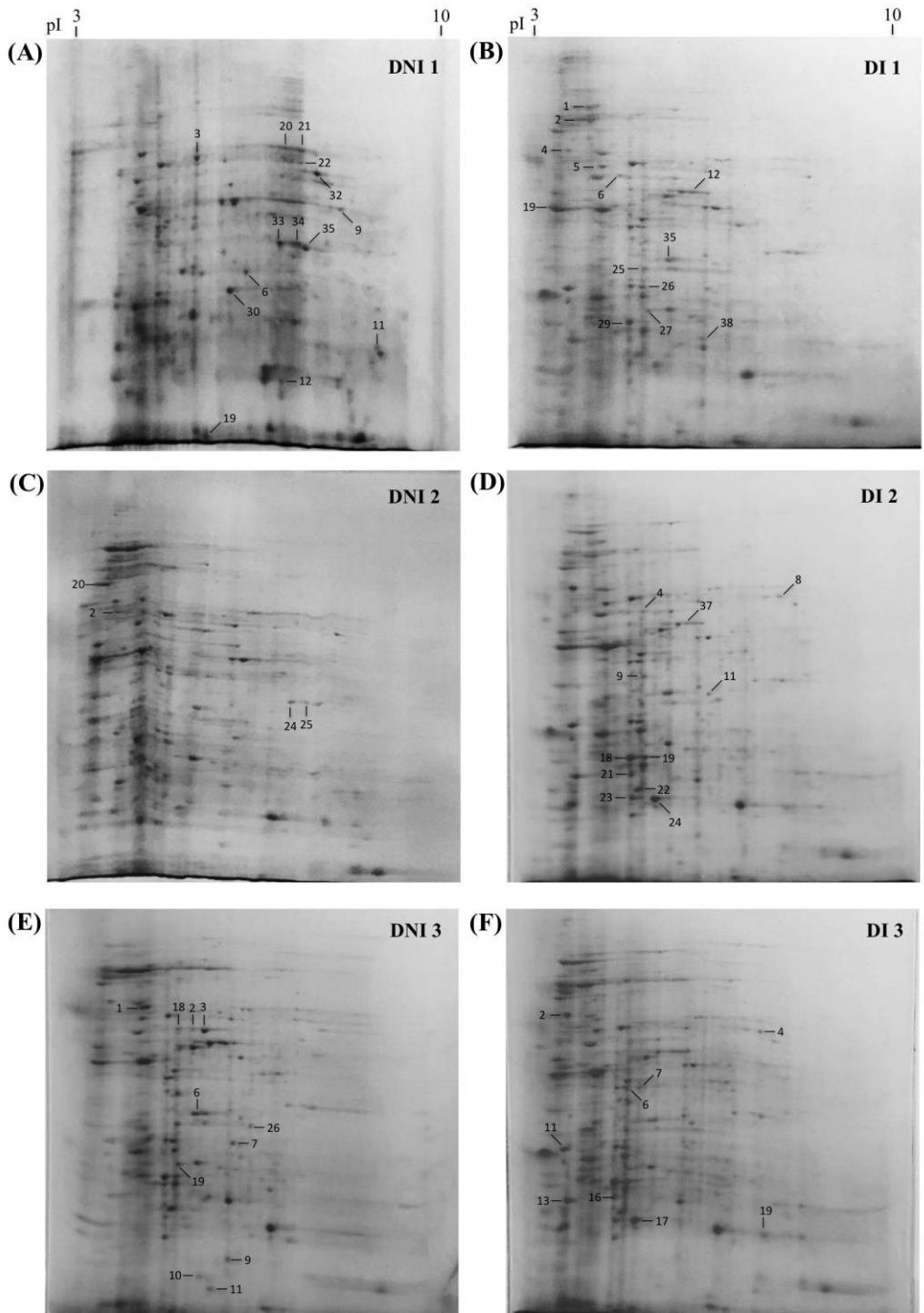


Figure 2.2. Comparison of 2D-GE profiles of *in vitro* antimony resistant lines (DI) with their corresponding sensitive pairs (DNI)

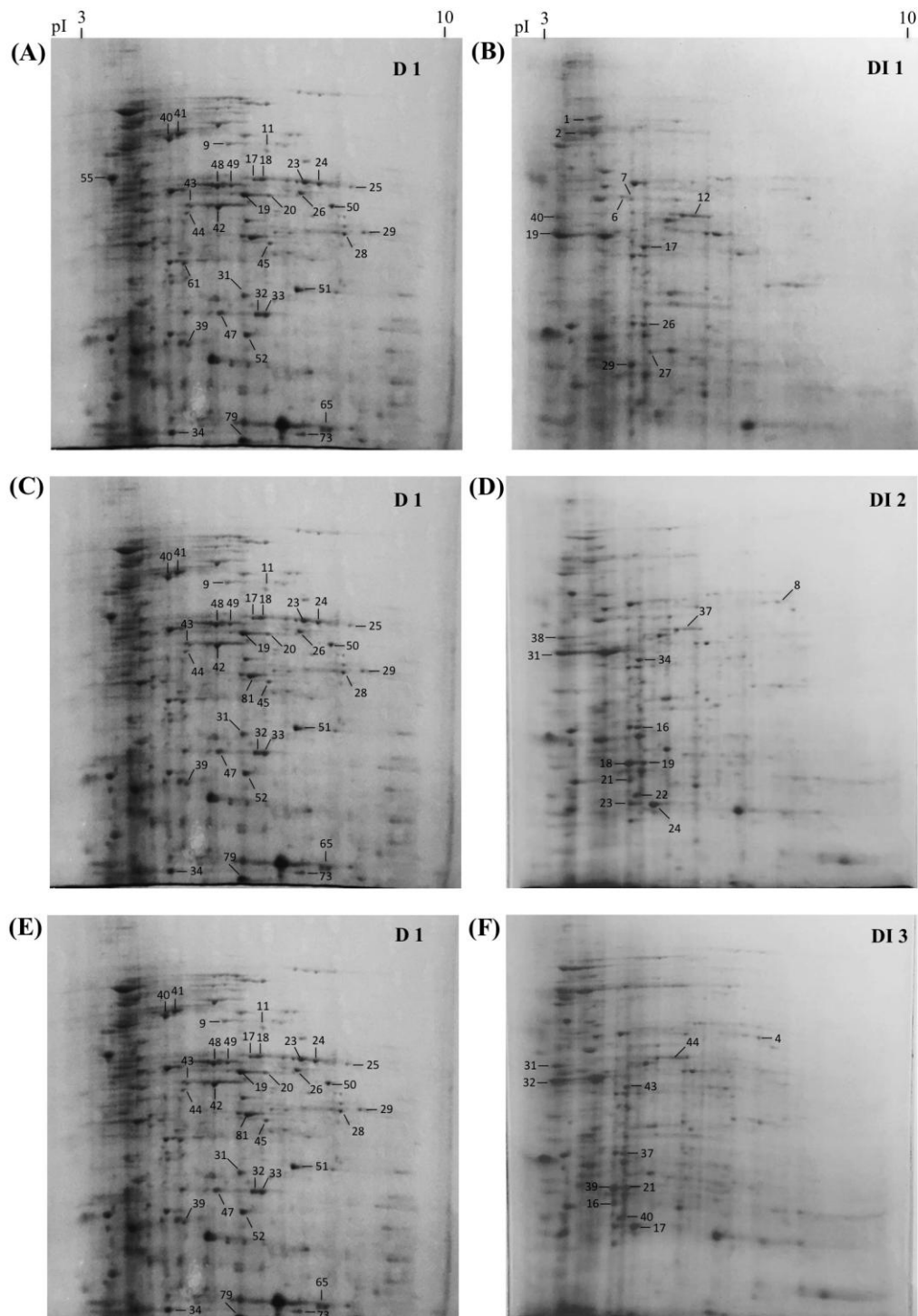


Figure 2.3. Comparison of 2D-GE profiles of *in vitro* antimony resistant lines (DI) with resistant clinical isolate (DI1)

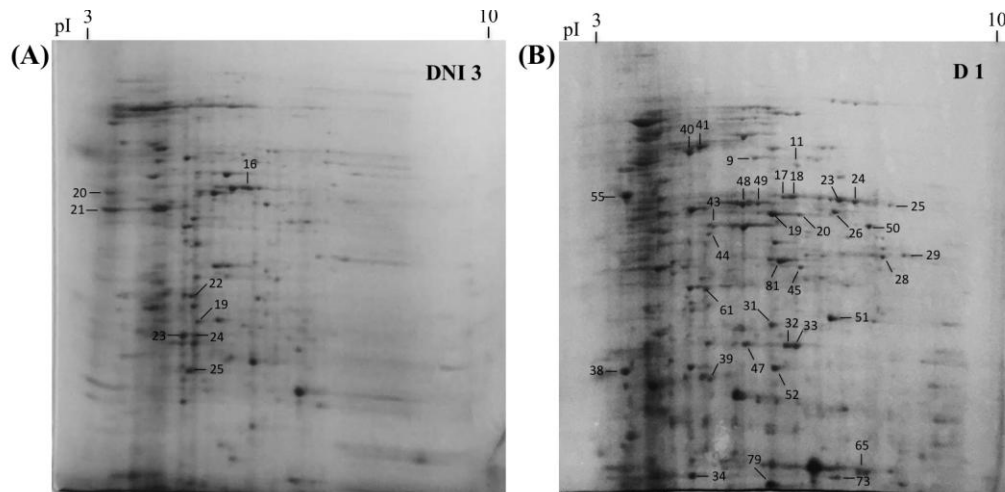


Figure 2.4. Comparison of 2D-GE profiles of antimony resistant clinical isolate, D1 with sensitive isolate, DNI3

2.3.1. Cytoskeletal Proteins

Paraflagellar rod protein 1 (PFR1), a cytoskeletal protein, exhibited higher abundance only in D1 resistant line when compared to its sensitive pair, DNI1. This protein is a component of paraflagellar rod of kinetoplastid flagella, which is involved in several functions in parasites including motility, developmental regulation and signaling (64). On the other hand, comparative proteomic analysis between D1 and *in vitro* resistant lines, D11, D12, D13, and DNI3 sensitive isolate showed that cytoskeletal proteins were more abundant in D1 clinical resistant line. Spots corresponding to PFR1D and PFR2D, another component of paraflagellar rod, was identified from D1 line. Previous studies have reported changes in the expression levels of various cytoskeletal proteins and among them PFR1 was shown to be down regulated in *in vitro* antimony resistant *L. braziliensis* isolates (57, 58). Although a direct association between increased levels of paraflagellar rod proteins and acquisition of drug resistance has not been reported in previous genomic and proteomic studies, findings of this study reveal that generation of resistance to antimony, mainly in host cell, causes an increase in the abundance of paraflagellar rod proteins. Regarding the physiological roles of paraflagellar rod proteins, they may be implicated in resistance mechanism via functioning in signal transduction cascades.

Table 2.2. MS/MS identification of differentially expressed proteins in DNI1 and DI1 by database search

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Antioxidant/Detoxification</i>						
DNI1-12	DR	putative iron superoxide dismutase [<i>L. infantum</i>]	XP_001467866.1	21742	6.49	77
DI1-38	N	iron superoxide dismutase [<i>L. chagasi</i>]	AAC38829.1	21846	6.29	67
<i>Cytoskeletal</i>						
DI1-6	N	paraflagellar rod protein 1 [<i>L. infantum</i>]	AAU09406.1	69598	5.30	98
<i>Metabolic Enzymes</i>						
DNI1-3	DR	metallo-peptidase, Clan MA(E), Family M32 [<i>L. infantum</i>]	XP_001468312.1	57449	5.44	243
DNI1-32	DR	metallo-peptidase, Clan ME, Family M16 [<i>L. major</i>]	XP_003722531.1	55015	6.36	80
DNI1-6	DR	activated protein kinase c receptor, partial [<i>L. tropica</i>]	AGL40001.1	32067	6.06	192
DNI1-9	DR	putative pyruvate dehydrogenase E1 component alpha subunit [<i>L. major</i>]	XP_001682547.1	43359	7.98	62
DNI1-20	DR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
DNI1-21	DR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
DNI1-11	DR	putative dihydrolipoamide acetyltransferase precursor [<i>L. major</i>]	XP_001686843.1	49062	7.01	68
DNI1-33	DR	putative methylthioadenosine phosphorylase [<i>L. major</i>]	XP_001687560.1	33730	6.26	67
DNI1-34	DR	putative methylthioadenosine phosphorylase [<i>L. major</i>]	XP_001687560.1	33730	6.26	154
DNI1-35	DR	fructose-1,6-bisphosphate aldolase [<i>L. major</i>]	XP_001686698.1	41201	8.82	125
DI1-25	UR	putative reiske iron-sulfur protein precursor [<i>L. major</i>]	XP_003722546.1	34062	5.93	73
DI1-26	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	250
DI1-35	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	455
<i>Protein Biosynthesis</i>						
DNI1-19	DR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	172
DNI1-30	DR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	238
DI1-27	UR	elongation factor 2 [<i>L. mexicana</i>]	XP_003874361.1	94852	5.77	190
DI1-12	N	putative eukaryotic initiation factor 4a [<i>L. infantum</i>]	XP_001462692.1	45356	5.83	104

(cont. on next page)

Table 2.2 (cont).

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Protein Folding/Chaperons and Stress Proteins</i>						
DI1-29	UR	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	416
DI1-4	N	calreticulin [<i>L. donovani</i>]	AAB17728.2	44604	4.55	45
DI1-19	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	319
<i>Hypothetical proteins</i>						
DNI1-22	DR	conserved hypothetical protein [<i>L. mexicana</i>]	XP_003876029.1	55693	6.34	232
DI1-1	UR	uncharacterized protein, partial [<i>L. mexicana</i>]	XP_003886477.1	73438	5.10	253
DI1-2	UR	uncharacterized protein, partial [<i>L. mexicana</i>]	XP_003886477.1	73438	5.10	95
DI1-5	N	conserved hypothetical protein [<i>L. major</i>]	XP_001681173.1	65707	5.32	131

Table 2.3. MS/MS identification of differentially expressed proteins in DNI2 and DI2 by database search

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Antioxidant/Detoxification</i>						
DI2-22	UR	mitochondrial trypanredoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	76
DI2-23	UR	mitochondrial trypanredoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	176
DI2-24	UR	mitochondrial trypanredoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	296
<i>Metabolic Enzymes</i>						
DNI2-2	DR	carboxypeptidase, putative [<i>L. donovani</i>]	XP_003864091.1	57440	5.40	95
DNI2-20	DR	2,3-bisphosphoglycerate-independent phosphoglycerate mutase [<i>L. major</i>]	XP_001687258.1	60903	5.21	219
DNI2-24	DR	putative methylthioadenosine phosphorylase [<i>L. major</i>]	XP_001687560.1	33730	6.26	67
DNI2-25	DR	putative methylthioadenosine phosphorylase [<i>L. major</i>]	XP_001687560.1	33730	6.26	154
DI2-4	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	130
DI2-9	UR	putative transaldolase [<i>L. major</i>]	XP_001682155.1	37225	5.4	191
DI2-11	UR	nonspecific nucleoside hydrolase [<i>L. major</i>]	XP_001682567.1	34545	5.85	52

(cont. on next page)

Table 2.3 (cont).

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Protein Biosynthesis</i>						
DNI2-23	DR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	238
DI2-21	UR	elongation factor 2, partial [<i>L. major</i>]	XP_001686588.1	94928	5.77	162
DI2-37	N	putative eukaryotic initiation factor 4a [<i>L. infantum</i>]	XP_001462692.1	45356	5.83	104
<i>Protein Folding/Chaperons and Stress Proteins</i>						
DI2-18	UR	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	338
DI2-19	N	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	257
<i>Hypothetical proteins</i>						
DI2-8	UR	conserved hypothetical protein [<i>L. mexicana</i>]	XP_003876029.1	55693	6.34	141

Table 2.4. MS/MS identification of differentially expressed proteins in DNI3 and DI3 by database search

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Antioxidant/Detoxification</i>						
DNI3-9	DR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21594	6.10	109
DI3-17	UR	mitochondrial trypanredoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	370
DI3-19	UR	peroxidoxin 1 [<i>L. aethiopica</i>]	AAZ23599.1	21563	6.43	136
<i>Metabolic Enzymes</i>						
DNI3-1	DR	putative ATPase beta subunit [<i>L. major</i>]	XP_001683872.1	56541	5.14	400
DNI3-2	DR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	100
DNI3-3	DR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	495
DNI3-6	DR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	455
DNI3-18	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	130
DNI3-26	DR	activated protein kinase c receptor [<i>L. tropica</i>]	AFR77659.1	34905	6.05	288
DI3-6	UR	cytochrome c oxidase subunit IV [<i>L. major</i>]	XP_001681672.1	39581	5.63	175

(cont. on next page)

Table 2.4 (cont).

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Protein Biosynthesis</i>						
DNI3-7	DR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	156
DNI3-11	DR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	172
DNI3-19	UR	elongation factor 2 [<i>L. mexicana</i>]	XP_003874361.1	94852	5.77	190
DI3-16	UR	elongation factor 2 [<i>L. major</i>]	XP_001686588.1	94928	5.77	360
DI3-7	N	elongation factor 2 [<i>L. major</i>]	XP_001686588.1	94928	5.77	171
<i>Protein Folding/Chaperons and Stress Proteins</i>						
DI3-2	UR	calreticulin [<i>L. donovani</i>]	AAB17728.2	44604	4.55	45
DI3-11	UR	putative glucose-regulated protein 78 [<i>L. major</i>]	XP_001684402.1	72067	5.05	162
<i>RNA/DNA Processing</i>						
DNI3-10	DR	putative ribonucleoprotein p18, mitochondrial precursor [<i>L. major</i>]	XP_001681946.1	21629	6.74	91
<i>Hypothetical proteins</i>						
DI3-4	UR	conserved hypothetical protein [<i>L. mexicana</i>]	XP_003876029.1	55693	6.34	232
DI3-13	UR	conserved hypothetical protein [<i>L. major</i>]	XP_001687618.1	22485	4.63	206

Table 2.5. MS/MS identification of differentially expressed proteins in D1 and DI1 by database search

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Antioxidant/Detoxification</i>						
D1-17	UR	trypanothione reductase [<i>L. amazonensis</i>]	ABQ57410.1	53697	5.93	94
D1-18	UR	trypanothione reductase [<i>L. amazonensis</i>]	ABQ57410.1	53697	5.93	94
D1-65	UR	peroxidoxin 1 [<i>L. aethiopica</i>]	AAZ23599.1	21563	6.43	136
D1-73	UR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21266	6.10	85
D1-79	UR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21266	6.10	94

(cont. on next page)

Table 2.5 (cont).

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Cytoskeletal</i>						
D1-40	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	105
D1-41	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	69
D1-39	N	paraflagellar rod protein 2C [<i>L. major</i>]	XP_001682225.1	69257	5.40	120
DI1-6	UR	paraflagellar rod protein 1 [<i>L. infantum</i>]	AAU09406.1	69598	5.30	98
DI1-7	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	98
<i>Metabolic Enzymes</i>						
D1-48	UR	metallo-peptidase, Clan MA(E), Family M32 [<i>L. infantum</i>]	XP_001468312.1	57449	5.44	243
D1-50	UR	metallo-peptidase, Clan ME, Family M16 [<i>L. major</i>]	XP_003722531.1	55015	6.36	80
D1-20	N	metallo-peptidase, Clan MH, Family M20 [<i>L. infantum</i>]	XP_001466769.1	50132	6.02	63
DI1-17	UR	metallo-peptidase, Clan MA(E), Family M32 [<i>L. major</i>]	XP_001686015.1	57470	5.55	245
D1-19	UR	S-adenosylhomocysteine hydrolase [<i>L. major</i>]	XP_001686974.1	48425	5.72	150
D1-23	UR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
D1-24	UR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
D1-26	UR	putative succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor [<i>L. major</i>]	XP_001685995.1	53145	7.94	180
D1-31	UR	nonspecific nucleoside hydrolase, partial [<i>L. tropica</i>]	AAS48369.1	32576	5.80	96
D1-33	UR	activated protein kinase c receptor, partial [<i>L. tropica</i>]	AGL40001.1	32067	6.06	223
D1-42	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	225
D1-43	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	130
D1-44	UR	S-adenosylhomocysteine hydrolase [<i>L. major</i>]	XP_001686974.1	48425	5.72	128
D1-47	UR	coproporphyrinogen III oxidase [<i>L. major</i>]	XP_001680900.1	34690	5.32	161
D1-49	UR	putative vacuolar ATP synthase subunit b [<i>L. major</i>]	XP_001684527.1	55826	5.56	165
D1-51	UR	putative methylthioadenosine phosphorylase [<i>L. major</i>]	XP_001687560.1	33730	6.26	67
D1-61	UR	cytochrome c oxidase subunit IV [<i>L. major</i>]	XP_001681672.1	39581	5.63	175
D1-32	N	prostaglandin f2-alpha synthase, partial [<i>L. tropica</i>]	ABI17868.1	30805	6.45	212

(cont. on next page)

Table 2.5 (cont).

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
D1-9	N	putative carnitine/choline acetyltransferase [<i>L. infantum</i>]	XP_001466657.1	68949	5.73	62
D1-11	N	glucose-6-phosphate dehydrogenase, partial [<i>L. donovani</i>]	AFK80083.1	60130	6.15	92
D1-25	N	Aldehyde dehydrogenase, mitochondrial [<i>L. tarentolae</i>]	Q25417.1	54787	7.51	80
D1-29	N	putative glutamate dehydrogenase [<i>L. major</i>]	XP_001684577.1	49669	6.29	174
D1-45	N	putative succinyl-CoA ligase [GDP-forming] beta-chain [<i>L. major</i>]	XP_001686874.1	44616	6.53	132
DI1-26	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	250
<i>Protein Biosynthesis</i>						
D1-52	UR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	238
DI1-27	UR	elongation factor 2 [<i>L. mexicana</i>]	XP_003874361.1	94852	5.77	190
DI1-12	N	putative eukaryotic initiation factor 4a [<i>L. infantum</i>]	XP_001462692.1	45356	5.83	104
<i>Protein Folding/Chaperons and Stress Proteins</i>						
D1-55	UR	calreticulin [<i>L. donovani</i>]	AAB17728.2	44604	4.55	45
D1-28	N	putative heat shock protein DNAJ [<i>L. mexicana</i>]	XP_003876821.1	44223	7.49	106
DI1-29	UR	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	416
DI1-19	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	319
DI1-40	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	97
<i>Transport</i>						
D1-34	UR	GTP-binding protein homologue [<i>L. major</i>]	AAA18826.1	22438	5.55	284
<i>Hypothetical proteins</i>						
DI1-1	UR	uncharacterized protein, partial [<i>L. mexicana</i>]	XP_003886477.1	73438	5.10	253
DI1-2	UR	uncharacterized protein, partial [<i>L. mexicana</i>]	XP_003886477.1	73438	5.10	95

Table 2.6. MS/MS identification of differentially expressed proteins in D1 and DI2 by database search

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Antioxidant/Detoxification</i>						
D1-17	UR	trypanothione reductase [<i>L. amazonensis</i>]	ABQ57410.1	53697	5.93	94
D1-18	UR	trypanothione reductase [<i>L. amazonensis</i>]	ABQ57410.1	53697	5.93	94
D1-65	UR	peroxidoxin 1 [<i>L. aethiopica</i>]	AAZ23599.1	21563	6.43	136
D1-73	UR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21266	6.10	85
D1-79	UR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21266	6.10	94
DI2-22	UR	mitochondrial tryparedoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	76
DI2-23	UR	mitochondrial tryparedoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	176
DI2-24	UR	mitochondrial tryparedoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	296
<i>Cytoskeletal</i>						
D1-40	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	105
D1-41	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	69
D1-39	N	paraflagellar rod protein 2C [<i>L. major</i>]	XP_001682225.1	69257	5.40	120
<i>Metabolic Enzymes</i>						
D1-48	UR	metallo-peptidase, Clan MA(E), Family M32 [<i>L. infantum</i>]	XP_001468312.1	57449	5.44	243
D1-50	UR	metallo-peptidase, Clan ME, Family M16 [<i>L. major</i> strain]	XP_003722531.1	55015	6.36	80
D1-20	N	metallo-peptidase, Clan MH, Family M20 [<i>L. infantum</i>]	XP_001466769.1	50132	6.02	63
DI2-34	UR	metallo-peptidase, Clan MA(E), Family M32 [<i>L. major</i>]	XP_001686015.1	57470	5.55	245
D1-19	UR	S-adenosylhomocysteine hydrolase [<i>L. major</i>]	XP_001686974.1	48425	5.72	150
D1-23	UR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
D1-24	UR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
D1-26	UR	putative succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor [<i>L. major</i>]	XP_001685995.1	53145	7.94	180
D1-31	UR	nonspecific nucleoside hydrolase, partial [<i>L. tropica</i>]	AAS48369.1	32576	5.80	96
D1-33	UR	activated protein kinase c receptor, partial [<i>L. tropica</i>]	AGL40001.1	32067	6.06	223
D1-42	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	225

(cont. on next page)

Table 2.6 (cont).

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
D1-43	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	130
D1-44	UR	S-adenosylhomocysteine hydrolase [<i>L. major</i> strain Friedlin]	XP_001686974.1	48425	5.72	128
D1-45	UR	succinyl-CoA ligase [GDP-forming] beta-chain [<i>L. major</i>]	XP_001686874.1	44616	6.53	132
D1-47	UR	coproporphyrinogen III oxidase [<i>L. major</i>]	XP_001680900.1	34690	5.32	161
D1-49	UR	putative vacuolar ATP synthase subunit b [<i>L. major</i>]	XP_001684527.1	55826	5.56	165
D1-51	UR	putative methylthioadenosine phosphorylase [<i>L. major</i>]	XP_001687560.1	33730	6.26	67
D1-81	UR	putative aldose 1-epimerase [<i>L. major</i>]	XP_003722492.1	41428	5.95	113
D1-32	N	prostaglandin f2-alpha synthase, partial [<i>L. tropica</i>]	ABI17868.1	30805	6.45	212
D1-9	N	putative carnitine/choline acetyltransferase [<i>L. infantum</i>]	XP_001466657.1	68949	5.73	62
D1-11	N	glucose-6-phosphate dehydrogenase, partial [<i>L. donovani</i>]	AFK80083.1	6.130	6.15	92
D1-25	N	Aldehyde dehydrogenase, mitochondrial [<i>L. tarentolae</i>]	Q25417.1	54787	7.51	80
D1-29	N	putative glutamate dehydrogenase [<i>L. major</i>]	XP_001684577.1	49669	6.29	174
DI2-16	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	235
<i>Protein Biosynthesis</i>						
D1-52	UR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	238
DI2-21	UR	elongation factor 2, partial [<i>L. major</i>]	XP_001686588.1	94928	5.77	162
DI2-37	N	putative eukaryotic initiation factor 4a [<i>L. infantum</i>]	XP_001462692.1	45356	5.83	104
<i>Protein Folding/Chaperons and Stress Proteins</i>						
D1-28	N	putative heat shock protein DNAJ [<i>L. mexicana</i>]	XP_003876821.1	44223	7.49	106
DI2-18	UR	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	338
DI2-19	N	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	257
DI2-31	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	97
DI2-38	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	97
<i>Transport</i>						
D1-34	UR	GTP-binding protein homologue [<i>L. major</i>]	AAA18826.1	22438	5.55	284
<i>Hypothetical proteins</i>						
DI2-8	UR	conserved hypothetical protein [<i>L. mexicana</i>]	XP_003876029.1	55693	6.34	141

Table 2.7. MS/MS identification of differentially expressed proteins in D1 and DI3 by database search

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Antioxidant/Detoxification</i>						
D1-17	UR	trypanothione reductase [<i>L. amazonensis</i>]	ABQ57410.1	53697	5.93	94
D1-18	UR	trypanothione reductase [<i>L. amazonensis</i>]	ABQ57410.1	53697	5.93	94
D1-65	UR	peroxidoxin 1 [<i>L. aethiopica</i>]	AAZ23599.1	21563	6.43	136
D1-73	UR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21266	6.10	85
D1-79	UR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21266	6.10	94
DI3-40	UR	mitochondrial trypanredoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	76
DI3-17	UR	mitochondrial trypanredoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	370
<i>Cytoskeletal</i>						
D1-40	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	105
D1-41	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	69
D1-39	N	paraflagellar rod protein 2C [<i>L. major</i>]	XP_001682225.1	69257	5.40	120
<i>Metabolic Enzymes</i>						
D1-50	UR	metallo-peptidase, Clan ME, Family M16 [<i>L. major</i>]	XP_0037222531.1	55015	6.36	80
D1-20	N	metallo-peptidase, Clan MH, Family M20 [<i>L. infantum</i>]	XP_001466769.1	50132	6.02	63
D1-48	N	metallo-peptidase, Clan MA(E), Family M32 [<i>L. infantum</i>]	XP_001468312.1	57449	5.44	243
DI3-43	UR	metallo-peptidase, Clan MA(E), Family M32 [<i>L. major</i>]	XP_001686015.1	57470	5.55	245
D1-11	N	glucose-6-phosphate dehydrogenase, partial [<i>L. donovani</i>]	AFK80083.1	6.130	6.15	92
D1-19	UR	S-adenosylhomocysteine hydrolase [<i>L. major</i>]	XP_001686974.1	48425	5.72	150
D1-23	UR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
D1-24	UR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
D1-26	UR	putative succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor [<i>L. major</i>]	XP_001685995.1	53145	7.94	180
D1-31	UR	nonspecific nucleoside hydrolase, partial [<i>L. tropica</i>]	AAS48369.1	32576	5.80	96
D1-33	UR	activated protein kinase c receptor, partial [<i>L. tropica</i>]	AGL40001.1	32067	6.06	223
D1-42	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	225

(cont. on next page)

Table 2.7 (cont).

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
D1-43	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	130
D1-44	UR	S-adenosylhomocysteine hydrolase [<i>L. major</i> strain Friedlin]	XP_001686974.1	48425	5.72	128
D1-47	UR	coproporphyrinogen III oxidase [<i>L. major</i> strain Friedlin]	XP_001680900.1	34690	5.32	161
D1-49	UR	putative vacuolar ATP synthase subunit b [<i>L. major</i>]	XP_001684527.1	55826	5.56	165
D1-51	UR	putative methylthioadenosine phosphorylase [<i>L. major</i>]	XP_001687560.1	33730	6.26	67
D1-81	UR	putative aldose 1-epimerase [<i>L. major</i>]	XP_003722492.1	41428	5.95	113
D1-32	N	prostaglandin f2-alpha synthase, partial [<i>L. tropica</i>]	ABI17868.1	30805	6.45	212
D1-9	N	putative carnitine/choline acetyltransferase [<i>L. infantum</i>]	XP_001466657.1	68949	5.73	62
D1-25	N	Aldehyde dehydrogenase, mitochondrial [<i>L. tarentolae</i>]	Q25417.1	54787	7.51	80
D1-29	N	putative glutamate dehydrogenase [<i>L. major</i>]	XP_001684577.1	49669	6.29	174
D1-45	N	putative succinyl-CoA ligase [GDP-forming] beta-chain [<i>L. major</i>]	XP_001686874.1	44616	6.53	132
DI3-37	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	235
<i>Protein Biosynthesis</i>						
D1-52	UR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	238
DI3-16	UR	elongation factor 2 [<i>L. major</i>]	XP_001686588.1	94928	5.77	360
DI3-44	N	putative eukaryotic initiation factor 4a [<i>L. infantum</i>]	XP_001462692.1	45356	5.83	104
<i>Protein Folding/Chaperons and Stress Proteins</i>						
D1-28	N	putative heat shock protein DNAJ [<i>L. mexicana</i>]	XP_003876821.1	44223	7.49	106
DI3-21	UR	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	257
DI3-39	UR	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	338
DI3-31	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	97
DI3-32	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	97
<i>Transport</i>						
D1-34	UR	GTP-binding protein homologue [<i>L. major</i>]	AAA18826.1	22438	5.55	284
<i>Hypothetical proteins</i>						
DI3-4	UR	conserved hypothetical protein [<i>L. mexicana</i>]	XP_003876029.1	55693	6.34	232

Table 2.8. MS/MS identification of differentially expressed proteins in D1 and DNI3 by database search

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
<i>Antioxidant/Detoxification</i>						
D1-17	UR	trypanothione reductase [<i>L. amazonensis</i>]	ABQ57410.1	53697	5.93	94
D1-18	UR	trypanothione reductase [<i>L. amazonensis</i>]	ABQ57410.1	53697	5.93	94
D1-65	UR	peroxidoxin 1 [<i>L. aethiopica</i>]	AAZ23599.1	21563	6.43	136
D1-73	UR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21266	6.10	85
D1-79	UR	peroxidoxin 1 [<i>L. tropica</i>]	AAZ23600.1	21266	6.10	94
DNI3-25	UR	mitochondrial tryparedoxin peroxidase [<i>L. amazonensis</i>]	AAX47429.1	25556	6.90	76
<i>Cytoskeletal</i>						
D1-40	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	105
D1-41	UR	putative paraflagellar rod protein 1D [<i>L. major</i>]	XP_003722257.1	69390	5.30	69
D1-39	N	paraflagellar rod protein 2C [<i>L. major</i>]	XP_001682225.1	69257	5.40	120
<i>Metabolic Enzymes</i>						
D1-50	UR	metallo-peptidase, Clan ME, Family M16 [<i>L. major</i>]	XP_003722531.1	55015	6.36	80
D1-20	N	metallo-peptidase, Clan MH, Family M20 [<i>L. infantum</i>]	XP_001466769.1	50132	6.02	63
D1-48	N	metallo-peptidase, Clan MA(E), Family M32 [<i>L. infantum</i>]	XP_001468312.1	57449	5.44	243
D1-11	N	glucose-6-phosphate dehydrogenase, partial [<i>L. donovani</i>]	AFK80083.1	6.130	6.15	92
D1-19	UR	S-adenosylhomocysteine hydrolase [<i>L. major</i>]	XP_001686974.1	48425	5.72	150
D1-23	UR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
D1-24	UR	pyruvate kinase [<i>L. major</i>]	XP_003722399.1	57833	6.43	114
D1-26	UR	putative succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor [<i>L. major</i>]	XP_001685995.1	53145	7.94	180
D1-31	UR	nonspecific nucleoside hydrolase, partial [<i>L. tropica</i>]	AAS48369.1	32576	5.80	96
D1-33	UR	activated protein kinase c receptor, partial [<i>L. tropica</i>]	AGL40001.1	32067	6.06	223
D1-42	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	225
D1-43	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	130
D1-44	UR	S-adenosylhomocysteine hydrolase [<i>L. major</i>]	XP_001686974.1	48425	5.72	128

(cont. on next page)

Table 2.8 (cont).

Spot no	Expression	Protein identity	Accession no	Mr	pI	Score
D1-47	UR	coproporphyrinogen III oxidase [<i>L. major</i>]	XP_001680900.1	34690	5.32	161
D1-49	UR	putative vacuolar ATP synthase subunit b [<i>L. major</i>]	XP_001684527.1	55826	5.56	165
D1-51	UR	putative methylthioadenosine phosphorylase [<i>L. major</i>]	XP_001687560.1	33730	6.26	67
D1-61	UR	cytochrome c oxidase subunit IV [<i>L. major</i>]	XP_001681672.1	39581	5.63	175
D1-81	UR	putative aldose 1-epimerase [<i>L. major</i>]	XP_003722492.1	41428	5.95	113
D1-32	N	prostaglandin f2-alpha synthase, partial [<i>L. tropica</i>]	ABI17868.1	30805	6.45	212
D1-9	N	putative carnitine/choline acetyltransferase [<i>L. infantum</i>]	XP_001466657.1	68949	5.73	62
D1-25	N	Aldehyde dehydrogenase, mitochondrial [<i>L. tarentolae</i>]	Q25417.1	54787	7.51	80
D1-29	N	putative glutamate dehydrogenase [<i>L. major</i>]	XP_001684577.1	49669	6.29	174
D1-45	N	putative succinyl-CoA ligase [GDP-forming] beta-chain [<i>L. major</i>]	XP_001686874.1	44616	6.53	132
DNI3-22	UR	enolase [<i>L. donovani</i>]	ACE74540.1	46662	5.41	235
<i>Protein Biosynthesis</i>						
D1-52	UR	elongation factor 2, partial [<i>L. major</i>]	AAG33264.1	72576	7.25	238
DNI3-19	UR	elongation factor 2 [<i>L. mexicana</i>]	XP_003874361.1	94852	5.77	190
DNI3-16	N	putative eukaryotic initiation factor 4a [<i>L. infantum</i>]	XP_001462692.1	45356	5.83	104
<i>Protein Folding/Chaperons and Stress Proteins</i>						
D1-38	UR	putative glucose-regulated protein 78 [<i>L. major</i>]	XP_001684402.1	72067	5.05	178
D1-28	N	putative heat shock protein DNAJ [<i>L. mexicana</i>]	XP_003876821.1	44223	7.49	106
DNI3-23	UR	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	338
DNI3-24	UR	putative heat-shock protein hsp70 [<i>L. major</i>]	XP_001684563.1	71893	5.30	257
DNI3-20	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	97
DNI3-21	N	heat shock protein 83-1 [<i>L. major</i>]	XP_001685759.1	80997	5.08	97
<i>Transport</i>						
D1-34	UR	GTP-binding protein homologue [<i>L. major</i>]	AAA18826.1	22438	5.55	284

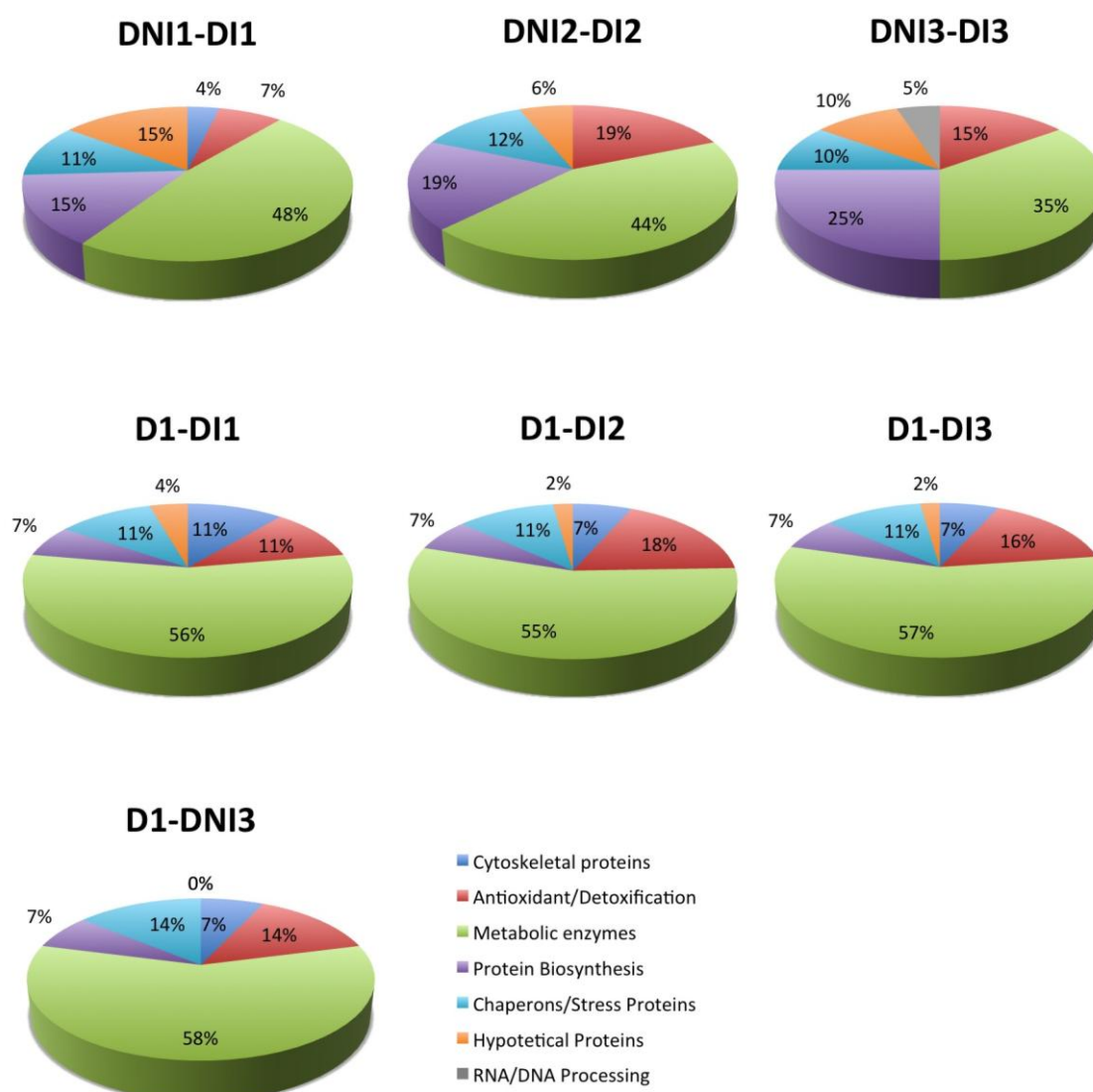


Figure 2.5. Relative distribution of identified proteins according to their biological roles. Compared pairs are indicated at top of each diagram

2.3.2. Protein Biosynthesis

Elongation factor 2 (EF2) is a protein that is involved in the translation machinery. It promotes translocation of nascent polypeptide chains within ribosome. EF2 is also implicated in signal transduction and apoptosis. EF2 was less abundant in all three *in vitro* resistant lines than their sensitive pairs in this study. Similarly, decreased amounts of EF2 in *in vitro* antimony resistant *L. braziliensis* lines has previously been reported and associated with mechanism of drug resistance and

response to high concentrations of drug (57). During the identification of protein spots corresponding to EF2, it was observed that several spots selected from different regions of the gel, which have different *Mr* and *pI* values, were identified as EF2. This finding can be accounted for post-translational modifications and truncated or degraded portions of the proteins that might have occurred either in the living system or during sample preparation. Moreover, some of these protein spots were found to be up regulated in resistant lines, when compared with their corresponding sensitive pairs. Previous studies also demonstrated the overexpression of EF2 in *in vitro* antimony resistant *L. panamensis* (63) and *L. infantum* (58) lines. Besides, an increase in the abundance of eukaryotic initiation factor 4a (eIF4A), which is required for mRNA binding to ribosome, was observed in DI1 and DI2 resistant lines. However, recent data reported the down regulation of eIF4A in *in vitro* resistant *L. braziliensis* (57).

As a result of comparison of D1 resistant clinical isolate with sensitive isolate and three *in vitro* resistant lines, it was observed that EF2 was more abundant and eukaryotic initiation factor 4a was less abundant in D1 isolate. EF2 has previously been shown to up regulated in resistant clinical isolate of *L. donovani* compared to its sensitive pair (61), corroborating the finding of this study. Identification of different spots as EF2 was also observed in this comparison. Such that, protein spots corresponding to EF2 were found to be overexpressed from different parts of the gel, indicating an increase in the level of EF2 variants in *in vitro* resistant lines. In brief, EF2 was less abundant and eukaryotic initiation factor 4a was more abundant in *in vitro* lines compared to their sensitive pairs and clinical isolate. Increased abundance of proteins that have important roles in protein synthesis is correlated with large-scale modulations of protein synthesis that is required for resistance phenotype. On the other hand, down regulation of proteins involved in protein synthesis might be due the direct effect of drug on specific proteins

2.3.3. Antioxidant and Detoxification

Antioxidants and detoxification process are the primary biological processes that are implicated in oxidative stress created by drugs. Trypanosomatids represent a unique mechanism for detoxification of peroxides, which is mediated by trypanothione. Peroxidoxins comprise a family of antioxidants and play important roles in the defense

against oxidative stress in prokaryotes and eukaryotes. In this study, protein spot corresponding to peroxidoxin 1 was found overexpressed in DI3 resistant line, compared to DNI3 sensitive pair. Similarly, overexpression of peroxidoxin was observed in *in vitro* antimony resistant *L. braziliensis* (57) and *L. infantum* (58) lines. Tryparedoxin peroxidase is a key enzyme in antioxidant defense mechanism for the detoxification of hydroperoxides. Increased amounts of tryparedoxin peroxidase was found in DI2 and DI3 resistant lines, compared to their sensitive isolates. In agreement with these results, increased expression of this protein has already been reported in antimony resistant *L. braziliensis* (57, 58) and *L. infantum* (32, 58) lines. Another protein that is involved in detoxification process, iron superoxide dismutase, exhibited an increase in abundance only in DI1 line. These data reveal that increased level of antioxidant defense mechanisms in resistant lines help parasite for circumventing toxic effects of the drugs and can be associated with the resistance of parasite to antimonial drugs.

Several enzymes that fall into this class were found increased in abundance in D1 resistant clinical isolate, in comparison to sensitive isolate and *in vitro* resistant lines. These proteins include trypanothione reductase and peroxidoxin. Similar to peroxidoxin, trypanothione reductase is another key enzyme of the trypanothione dependent defense mechanism of *Leishmania*. Peroxidoxin and trypanothione reductase were more abundant in D1 isolate than sensitive isolate and *in vitro* resistant lines. According to previous genomic and proteomic studies on drug resistance, it was shown that redox reactions play a major role in the mechanism of resistance. In this study, overexpression of several antioxidants and detoxification enzymes were observed in *in vitro* resistant lines compared to their sensitive lines. Further increase in the abundance of these enzymes and other new ones in clinical isolate suggest the increased level of defense mechanism in parasite due to the host organism defense mechanism. Reduction in oxidative stress is required for survival of the parasite in host organism and may accompany the acquisition of resistance.

2.3.4. Protein Folding/Chaperones and Stress Proteins

Chaperones are a class of proteins whose function is to aid folding of nascent polypeptides, transport across membranes, disassembly of macromolecular complexes

or aggregates and ensure correct folding of proteins. As stated before, these proteins are involved in programmed cell death (PCD), as well. Because most of the chaperones were first identified as being induced by heat shock or other stresses, they are also known as heat shock proteins (HSPs) (65). When cells are exposed to stress conditions such as heat and drug, HSPs protect cells from deleterious environmental effects. Overexpression of HSPs, HSP70 and/or HSP83, have already been reported in various *in vitro* and clinical isolates of antimony resistant *Leishmania spp.* (32, 56, 58, 60, 61, 63) and associated with drug resistance. It has been suggested that expression of HSPs is a primary nonspecific response of parasite to toxic effects of drugs. They do not confer to resistance directly but rather help cells to develop more efficient and specific mechanisms (58). For example, it has been found that HSP70 and HSP83 are acting as negative regulators of mitochondria dependent apoptosis pathway (66), thereby overexpression of these proteins may increase resistance by protecting parasite from drug induced PCD. In this study, increased amounts of HSP83 in DI2 line, and HSP70 in DI1 and DI2 lines were found when compared to their sensitive pairs. In DI3 line, glucose-regulated protein 78, a member of HSP70 molecular chaperone family, was found overexpressed. Besides these proteins, a different type of chaperone, calreticulin, was more abundant in DI1 and DI3 lines compared to their corresponding sensitive lines. Calreticulin is a calcium-binding chaperone and is involved in folding and assembly of glycoproteins. Increased amounts of calreticulin have been found in *in vitro* antimony resistant *L. braziliensis* (57). However, overexpression of this protein has not been directly linked to drug resistance in *Leishmania*.

When compared with D1 resistant line, HSP70 and HSP83-1 were more abundant in *in vitro* resistant lines and sensitive line. On the other hand, another heat shock protein, HSP DNAJ, was only detectable in D1 line. This protein is also known as HSP40 and regarded as a crucial partner for HSP70. Moreover, increased amount of glucose-regulated protein 78 was observed in D1 line, compared to DNI3 sensitive line. Previous studies reported the down regulation of glucose-regulated protein 78 in *in vitro* resistant *L. braziliensis* (58) and HSP DNAJ in *in vitro* resistant *L. infantum* (32), compared to their sensitive isolates. However, these findings have not been directly associated with drug resistance and also these studies were not conducted on antimony resistant clinical isolates. The results of this study showed that, HSPs are more abundant in *in vitro* resistant lines than their sensitive pairs and less abundant in D1 resistant clinical isolate compared to sensitive isolate and *in vitro* resistant lines. The observation

of decreased abundance in HSPs in resistant clinical isolate might be originating from a combined effect of drug and host cell defense mechanism on parasite, mainly on HSPs. In addition, the acquisition of antimony resistance of *L. tropica* used in this study may follow different pathways in host organism and in *in vitro* conditions.

2.3.5. RNA/DNA Processing

Comparative proteomic analysis identified one differentially expressed protein involved in RNA processing. Decreased amount of ribonucleoprotein p18, a RNA-binding protein, was found in DI3 resistant line. This protein has not been reported as a differentially expressed protein in other antimony resistant *Leishmania spp.* before. Ribonucleoproteins are ribosomal proteins that conjugate with RNA and are involved in translation. Whereas, ribonucleoprotein p18 identified in this study is the mitochondrial precursor of the protein and down regulation of this protein may be associated with damaging effects of high concentrations of antimony on specific mitochondrial enzymes.

2.3.6. Metabolic Enzymes

Energy generation in *Leishmania* promastigotes is mainly achieved through glycolysis and amino acid metabolism. Previous studies on the mechanism of action of pentavalent antimony suggested that this drug inhibits biosynthesis of macromolecules by inhibition of energy metabolism, glycolysis and fatty acid oxidation. In this study, altered abundance of enzymes for various metabolic pathways were observed. Decreased amounts of pyruvate kinase, pyruvate dehydrogenase E1 component alpha subunit and dihydrolipoamide acetyltransferase precursor in DI1 line and 2,3-bisphosphoglycerate-independent phosphoglycerate mutase in DI2 line were found, compared to their sensitive pairs. Pyruvate kinase and phosphoglycerate mutase are enzymes of glycolytic pathway. Pyruvate dehydrogenase E1 component and dihydrolipoamide acetyltransferase are enzymes of pyruvate dehydrogenase complex, which links glycolysis to tricarboxylic acid cycle via decarboxylation of pyruvate. Down regulation of pyruvate kinase has already been reported in *in vitro* antimony resistant *L. braziliensis* (57) and *L. infantum* (59), corroborating the results of this study.

On the other hand, up regulation of pyruvate dehydrogenase E1 component (58) and phosphoglycerate mutase (32, 57) in *L. infantum* have been reported in previous studies. Another enzyme involved in glycolysis, enolase, was found over expressed in both DI1 and DI2 lines and down regulated in DI3 line, compared to their sensitive pairs. Differential abundance of enolase in *Leishmania spp.* was observed in previous studies, too (58, 60). Protein spots corresponding to enolase were identified from distinct regions of the gels, indicating the possibility of post-translational modifications or degraded/truncated forms of the enzyme, as mentioned previously for another protein, elongation factor-2. A decrease in abundance of fructose-1,6-bisphosphate aldolase, an enzyme involved in gluconeogenesis, was observed in DI1 line. In contrast to the finding of this study, increased levels of this enzyme were observed in antimony resistant isolates of *L. donovani* (60) and *L. infantum* (61). A spot corresponding to transaldolase, which links pentose phosphate pathway to glycolysis, was found over expressed and was only detectable in DI2 line. Two proteins involved in electron transport chain, reiske iron-sulfur protein and cytochrome c oxidase, were increased in abundance in DI1 and DI3 lines, respectively. At the same time, down regulation of ATPase beta subunit, a proton transporter for the synthesis of ATP, was observed in DI3 line. Protein spots identified as activated protein kinase-c receptor (LACK) were less abundant in DI1 and DI3 lines than their sensitive isolates. LACK is involved in several signal transduction cascades and already been shown to be down regulated in *in vitro* resistant *L. infantum* amastigotes (59). Changes in the abundance of some enzymes that have roles in salvage pathway were observed, as well. Protozoan parasites lack the *de novo* synthesis of purine bases and are dependent on the host cell nucleosides and nucleotides for synthesis of purines. Nucleoside hydrolase is involved in the purine salvage pathway and considered as a prime target for the development of anti-parasitic agents (64). This enzyme was only detectable in DI2 line with increased abundance. Another enzyme involved in salvage pathway of purines is methylthioadenosine phosphorylase and was found down regulated in DI1 line. Comparative proteomic analysis also identified some proteins involved in proteolysis. Metallo-peptidase Clan MA(E) Family M32, and carboxypeptidase were less abundant in DI1 and DI2 lines than their corresponding sensitive pairs. In contrast, previous studies have reported the overexpression of enzymes involved in peptide catabolism from *in vitro* resistant *Leishmania spp.* (57, 63).

A number of metabolic enzymes having altered abundance were identified when proteome of resistant clinical isolate was compared to sensitive and *in vitro* resistant lines. The amounts of some of the glycolytic pathway enzymes, such as pyruvate kinase and enolase, were found increased in D1 line. The spot corresponding to S-adenosylhomocysteine hydrolase (SAHH) was more abundant in D1 line than sensitive isolate and *in vitro* resistant lines. SAHH is involved in the conversion of S-adenosylhomocysteine to homocysteine, which is then used in the generation of cysteine, a precursor of glutathione and trypanothione. Increased amounts of SAHH reveals that the trypanothione dependent detoxification mechanism of antimony resistant *Leishmania* is regulated not only at the level of antioxidant and detoxification enzymes but also by modulation of the level of enzymes involved in the biosynthetic pathway for trypanothione. Overexpression of this enzyme has previously been reported in clinical isolate of resistant *L. donovani* (61) and *in vitro* resistant *L. panamensis* (63). Increased amounts of nucleoside hydrolase and methylthioadenosine phosphorylase, enzymes of the salvage pathway of purines, were observed in D1 line. LACK and vacuolar ATP synthase beta subunit were also more abundant in D1. Similar to the findings of this study, increased amount of LACK has been reported in antimony resistant clinical isolate of *L. tropica* (62). Besides these proteins, some other enzymes (with their specific biochemical roles indicated in parenthesis) which increased in abundance in D1 line include; succinyl-coA:3-ketoacid-coenzyme A transferase (ketone body catabolism), coproporphyrinogen III oxidase (heme biosynthesis), putative carnitine/choline acetyltransferase (fatty acid oxidation), glucose-6-phosphate dehydrogenase (pentose phosphate pathway), aldehyde dehydrogenase (aldehyde metabolism), glutamate dehydrogenase (alpha-ketoglutarate synthesis), succinyl-CoA ligase [GDP-forming] beta-chain (carbohydrate metabolism), aldose 1-epimerase (carbohydrate metabolism), prostaglandin f2-alpha synthase (prostaglandin synthesis) and GTP binding protein (signal transduction). Increased amounts of aldehyde dehydrogenase, glucose-6-phosphate dehydrogenase and aldose 1-epimerase have already been observed in antimony resistant *L. donovani* clinical isolate (61). Overexpression of prostaglandin f2-alpha synthase has also been shown in *in vitro* resistant *L. panamensis* line (63). However, changes in the abundance of other proteins in relation to drug resistance have not been reported in previous studies. Proteins involved in proteolysis, metallo-peptidase Clan MA(E) Family M32, metallo-peptidase Clan ME Family M16 and metallo-peptidase Clan MH Family M20 exhibited higher

abundance in D1 resistant clinical isolate compared to *in vitro* resistant lines and sensitive line. Overexpression of proteolytic enzymes in antimony resistant clinical isolates has been reported in previous studies, as well (60, 61).

These results showed that some of the metabolic enzymes were up regulated, whereas some others were down regulated in *in vitro* resistant lines and resistant clinical isolate. Increased level of enzymes of glucose metabolism and ATP synthesis might be associated with the energy requirement of parasite for proliferation and survival in the presence of drug. Moreover, it has been suggested that the increased level of glycolytic enzymes not only results in the increased level of glycolysis for the production of more energy but also increases the level of pyruvate. Pyruvate is a glycolysis product, at the same time it acts as a scavenger of peroxides. Therefore, increased level of glycolysis makes contributions to defense mechanism. In addition, increase in the abundance of enzymes of salvage pathway shows elevation of DNA and RNA synthesis, hence might be related with parasite proliferation. Increased abundance of proteolytic enzymes is linked to survival of the parasite via increased proteolytic activities for the degradation of damaged proteins. The deleterious effect of antimony on metabolic enzymes is observed as diminished levels of these enzymes in resistant lines, thereby resulting in the perturbation of metabolic activities. It is important to note that, metabolic enzymes of various pathways were more abundant in resistant clinical isolate than in *in vitro* resistant lines, revealing the elevated levels of energy production in resistant clinical isolate.

2.3.7. Hypothetical Proteins

A number of hypothetical and uncharacterized proteins were identified in clinical and *in vitro* resistant lines and sensitive line. A conserved hypothetical protein with NCBIprot accession number XP_003876029.1 was found overexpressed in DI2 and DI3 lines, compared to their sensitive pairs and D1 isolate. An uncharacterized protein, XP_003886477.1, was increased in abundance in DI1 compared to DNI1 and D1 isolates. Another hypothetical protein, XP_001681173.1, with increased abundance was observed in DI1 and DI3 lines in comparison to sensitive pairs. Lastly, a conserved hypothetical protein having XP_001687618.1 accession number was more abundant in DI3 line than DNI3 sensitive isolate. The results reveal that these hypothetical proteins

were more abundant in *in vitro* resistant lines, suggesting their association with resistance phenotype. According to UniProt database (<http://www.uniprot.org>), some of these proteins are involved in calcium ion binding (XP_001681173.1), isoprenoid biosynthesis (XP_003876029.1) and has a domain thioredoxin like fold (XP_001687618.1). Previous data reported that isoprenoid biosynthesis generates a number of end products that are essential for cell viability and this pathway provides potential targets for drug discovery. The other protein with thioredoxin like fold is probably involved in detoxification process.

2.4. Conclusion

Pentavalent antimony based compounds are used as first line drugs in the treatment of different forms of leishmaniasis. Unfortunately, the emergence of antimony resistant *Leishmania spp.* resulted in a rapid increase in the number of unresponsive treatments. The mechanism of action of antimony and generation of antimony resistance in *Leishmania* are not completely understood. Therefore, understanding how antimonials work and why they sometimes fail is very important to the optimal use of existing drugs and pave the way in the development of new therapeutics. Up to date, a large number of genomic and proteomic studies were conducted with *in vitro* generated resistant lines as well as with clinical isolates from antimony unresponsive patients in order to find new insights into the resistance mechanisms. These studies were focused mainly on the easily cultured, *in vitro* generated resistant promastigotes. Moreover, regarding the differences in the environmental conditions that the parasites are exposed in host organism and culture conditions, differences in the accumulation of these two resistance types are expected, as well. Therefore, this study was designed to investigate; *in vitro* generated antimony resistance, antimony resistance gained in host organism and differences between these two resistance types.

The comparative proteomic analysis of *in vitro* antimony resistant lines and resistant clinical isolate resulted in the identification of 48 differentially expressed proteins with some retrieved from several spots, revealing post-translational modifications, isoforms or degraded forms of the original proteins. The results showed that, accumulation of antimony resistance in *Leishmania*, either *in vitro* or in host organism, is associated with the altered abundance of proteins involved in different

cellular functions such as; metabolic processes, antioxidant and detoxification mechanism, protein biosynthesis, protein folding/chaperons and stress response. Majority of the identified proteins (about 50%) that exhibited changes in their abundance are enzymes of metabolic processes in both resistant types. However, this change was not in the same direction in *in vitro* resistant lines and clinical isolate. Such that, when compared to sensitive isolate most of the identified metabolic enzymes were decreased in abundance in *in vitro* resistant lines, whereas increased in abundance in resistant clinical isolate. Increased amounts of metabolic enzymes were also observed when *in vitro* resistant lines and clinical isolate was compared. Although expression levels of cytoskeletal proteins did not show a significant change in *in vitro* resistant lines, they were more abundant in clinical isolate. Besides, proteins involved in detoxification processes showed a similar trend in their expression levels in both resistant types. Increased amounts of these proteins were observed, with more abundance in clinical isolate than *in vitro* resistant lines. In addition, chaperons were less abundant in clinical isolate when compared to sensitive isolate and *in vitro* resistant lines. All these changes reveal that *Leishmania* alter the abundance of proteins involved in these different pathways for its protection, proliferation and reduction of the effects caused by drug pressure. More importantly, the mechanism of antimony resistance in host cell and *in vitro* conditions probably follow different strategies. Accumulation of resistance in host organism is associated with increased levels of metabolic processes and defense mechanism, as well as changes in cytoskeletal proteins. Increased levels of proteins in defense mechanism shows that the parasite in host organism is trying to cope with the host organism immune system together with drug pressure. On the other hand, *in vitro* antimony resistance resulted in diminished levels of metabolic processes and increased abundance of chaperons. Therefore, *in vitro* antimony resistance may possibly be mediated through circumventing deleterious effects of drug via increased levels of proteins involved in protein folding and defense mechanism. Although, overexpression of chaperons or some of the metabolic enzymes are mostly regarded as nonspecific primary responses against antimony, this response may help cells to develop more efficient and specific reaction mechanisms and delay cell death. Moreover, some of the identified specific proteins or a combination of some of these proteins may be required for the accumulation of drug resistance.

All in all, it is clear that drug resistance in *Leishmania* is complex and multifactorial. Studies with resistant clinical isolates and *in vitro* selected lines would

together pave the way for a better understanding of resistance mechanisms in *Leishmania* and mechanism of action of antimony. It is believed that these results can make important contributions for the development of more efficient therapeutics against leishmaniasis.

CHAPTER 3

PROTEOMIC ANALYSIS OF BENCE JONES PROTEINS ISOLATED FROM URINE OF PATIENTS WITH MULTIPLE MYELOMA

3.1. Introduction

Comparative proteomic analysis in health sciences allows the identification of biomarkers by comparing the protein profiles of biological samples, such as tissues, cells, urine, blood plasma/serum and saliva, obtained from healthy individuals and patients (67). The combination of protein separation technologies, such as high-resolution 2D-gel electrophoresis or liquid chromatography, with protein identification and sequencing techniques, such as mass spectrometry, enabled the discovery of biomarkers for various diseases including renal disorders. Concerning renal diseases, tissue biopsy may seem to be an ideal sample for biomarker discovery studies. However, body fluids are more appropriate than tissue samples due the advantages they provide such as low invasiveness, low cost, easy sample collection and processing (68). Moreover, some patients with bleeding disorders, obesity, end stage renal disease and severe hypertension may not be suitable for renal tissue biopsy (69). Therefore, human body fluid analysis has become a preferred method for the discovery of renal disease biomarkers. Among body fluids, urine is a better proteomic sample because it can be easily obtained in large amounts, urine proteins/peptides are less complex and more stable, and changes in kidney function can directly affect the composition of the urine (70).

In normal conditions, kidney filters and degrades most of the plasma proteins and a trace amount of protein (150 mg/day) is observed in urine. Albumin, retinol binding protein, α 1-microglobulin, β 2-microglobulin, cystatin C, immunoglobulin heavy and light chains, transferrin, α 2-microglobulin, α 1 antitrypsin, and lysozyme are some of the proteins observed in certain amounts in urine of a healthy individual (71). Among these proteins, immunoglobulins (Ig) have primary roles in immune system, by

recognizing foreign molecules and activation of a response to pathogen invasions. Basically, Ig molecules are formed from two identical heavy chains and two identical light chains, joined by inter-chain disulfide bonds in the hinge region. The general structure of an Ig molecule is shown in Figure 3.1. Each light chain and heavy chain is composed of a constant (C) and variable (V) domain. Variable domains on the light and heavy chains (VL and VH, respectively) contain four conserved framework regions (FR) and three hypervariable complementarity determining regions (CDR). The FRs are involved in the folding of Ig molecules and CDRs are responsible for antigen recognition and binding (72). Production of heavy and light chains as well as incorporation of light chains into Ig molecule occurs in plasma cells (also known as B-cells) of bone marrow. In normal conditions, excess amount of light chains is produced over heavy chains and the majority of light chains are bound to heavy chains. The unbound, free light chains can be detected in serum and urine at low levels (73). These free light chains are filtered through the glomeruli, reabsorbed in the proximal tubules and degraded. The urine of a healthy human is generally free of light chains and the observation of less than 3-5 mg/day light chains in urine is regarded as normal. However, in several disease conditions that affect the bone marrow, increased amounts of free light chains in bloodstream and urine are observed. When the metabolizing capacity of the nephrons is exceeded by the excessive amounts of light chains in bloodstream, these proteins appear in urine at elevated amounts. This generally occurs when the production of light chains exceeds 10-30 g/day (74).

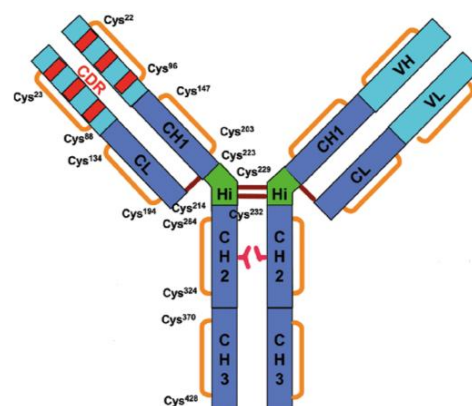


Figure 3.1. Schematic representation of an immunoglobulin structure (75)

Multiple myeloma and other immunoproliferative disorders such as chronic lymphocytic leukemia and Walden-ström's microglobulinemia are characterized by the excessive proliferation of monoclonal plasma cells (76). These cells secrete abnormally high amounts or dysfunctional forms of light chains into the circulation. The light chains in bloodstream can deposit in various tissues including kidney, hearth and liver and lead to organ failure. Kidney is the most frequently effected organ and majority of renal disorders as well as some other diseases are caused by the excess amount of light chains. Formation of amyloid fibrils and accumulation of these fibrils in vital organs causes organ failure and associated with a disease called light chain amyloidosis (AL). Another one, light chain deposition disease (LCDD) is characterized by the formation of granular amorphous aggregates in the basement of kidney, as well as in hearth and liver (77). Moreover, formation of light chain crystals results in Fanconi's syndrome, a disease of renal proximal tubule cells. Cast nephropathy, also called as myeloma kidney, is another disease caused by light chains in which certain type of light chains bind to Tamm-Horsfall proteins of kidney to form casts and lead to further renal complications (74, 78). These light chains, which are produced in excess amounts and associated with immunoproliferative diseases, are also referred to as Bence Jones proteins (BJP). The name is owed to Dr. Henry Bence Jones, who received a urine sample with unusual properties and reported his observations in 1840s (79). He detected an abnormal protein that was later defined as Bence Jones protein (BJP), and it was found that the unusual properties of urine arise from the spill over of excess free monoclonal immunoglobulin light chains (80). Observation of elevated levels of light chains in urine is often called as Bence Jones proteinuria and arises mainly from the overproduction, increased filtration or decreased proximal tubular reabsorption of light chains. All in all, excess amount of BJP in urine is a concomitant phenomenon of the mentioned immunoproliferative diseases and became a widely used biomarker in the diagnosis of these diseases and renal failure (74).

BJP have molecular weight of 22-24 kDa and are characterized as lambda (λ) or kappa (κ) chains (81). In humans, κ chains are produced in approximately twice the amount of lambda λ chains. In normal conditions, the ratio of λ/κ light chains in urine is approximately 0.5, but this ratio was found to have increased to 3 in certain disease conditions like light chain amyloidosis (82). Moreover, λ and κ chains can be found in various forms in bloodstream and urine. These include low molecular weight fragments, monomers, dimers and multimeric complexes (trimers, tetramers or trimer of a dimer)

(83). Association of BJP monomers into dimers depends on the chain type and arises from the ability of these molecules to mimic the Fab portion of an antibody in which heavy and light chain are held together via interactions (81, 83). Dimerization of BJP are mainly stabilized by covalent and non-covalent interactions. In non-covalent dimerization, hydrogen bonds, intermolecular static and hydrophobic interactions may play role, whereas in covalent dimerization disulfide bond between cysteine residues located in the C-terminal region of light chains stabilizes the structure. The non-covalent dimerization of BJP is achieved through the conserved amino acid residues in the framework region of the variable domain. The amino acid residues that take part in hydrogen bond formation are tyrosine 36, glutamine 38, and tyrosine 87, whereas leucine 46 and phenylalanine 98 take part in hydrophobic interactions. On the other hand, residues in the CDR region can have significant effects on the geometry and formation of dimers, as well (73).

The renal disorders, which are associated with BJP proteinuria, mainly result from the direct toxic effects of these proteins on the proximal tubule cells or interaction of BJP with Tamm-Horsfall proteins in the renal tubule to form myeloma casts. However, there is a considerable variability among BJP that are associated with renal disorders. Such that, in some patients even small amounts of BJP in urine give rise to severe renal dysfunctions, while others with BJP proteinuria up to 8-9 g/day or greater are associated with minimal nephrotoxicity. Though it has been shown that the variability of BJP nephrotoxicity was determined by variable region of the light chain, the precise determinants of variable toxicity have not been fully characterized (74). Previous studies demonstrated that λ chains are mainly found in the amyloid structures of AL amyloidosis (84), and κ chains are more frequently observed in Fanconi's syndrome compared to λ chain (85). However, it is now widely accepted that both λ and κ chains can be equally nephrotoxic (74). The studies on light chain nephrotoxicity revealed important mechanisms concerning direct toxicity of these proteins. It has been shown that, light chains at concentrations found in the tubule fluid of multiple myeloma patients inhibit glucose, amino acid and phosphate transport in cultured proximal tubule cells (86, 87). Moreover, Na-K-ATPase on proximal tubule cells has been inhibited on both gene and protein level (88). Another study demonstrated that light chains cause apoptosis and necrosis, associated with cytoskeletal injury and DNA damage in cultured human proximal tubule cells (89). Further investigations noted the production of inflammatory and proinflammatory cytokines through activation of nuclear factor-kappa

B (NF- κ B) (90) and mitogen activated protein kinases (MAPKs) induced by light chain exposure (91).

The tendency of BJP to exist as structures of different molecular weights makes the electrophoretic analysis suitable for detection of these proteins. In clinical diagnosis, urinary protein electrophoresis (UPE), immunoelectrophoresis (IE) and immunofixation electrophoresis (IFE) are used as routine tests. The urine is screened for light chains using UPE on cellulose acetate strips or agarose gels and the identity of the light chain is confirmed using IE or IFE. Moreover, a great number of publications have been released in literature for the analysis of BJP using 1D-gel electrophoresis, 2D-gel electrophoresis, isoelectric focusing, and capillary electrophoresis. These methods are widely employed in research laboratories for molecular characterization of light chain and its association with nephrotoxicity. All these methodologies have been extensively reviewed by Marshall et. al (78). Numerous publications demonstrated the primary structure of light chains using conventional protein sequencing and amino acid analysis methods (92-94). In a previous work by Kishida et al., an additional cysteine residue was identified in variable region of BJP following acid hydrolysis and amino acid determination (92). In another work, mass spectrometry analysis of BJP in combination with enzymatic digestions allowed the identification of a cysteinyl post-translational modification on a cysteine residue in variable region of light chain (95). In recent years, mass spectrometry has been widely used as a powerful tool in the clinical analysis of BJP in order to monitor (96) and quantify (97) free light chains in serum, identify λ and κ chains in serum (98, 99), as well as to classify amyloid deposits in biopsy samples (100).

During investigation the cytotoxicity of light chains obtained from the urine of patients with multiple myeloma, who had significant light chain proteinuria and modest to moderate renal insufficiency, it was observed that the light chains showed varying degrees of *in vitro* renal toxic effects on proximal tubule cells (89). Additional studies using these light chains pointed various toxic effects including inhibition of Na-dependent glucose, amino acid, and phosphate transport (86, 87), inhibition of activity and gene expression of Na-K-ATPase in kidney cells (88), inducing inflammatory and pro-inflammatory cytokines and activation of nuclear factor kappa B, (NF- κ B) (90), and activation of mitogen activated protein kinases (MAPKs) (91). However, the proteomic analyses of these samples have not been performed up to date. Therefore, the aim of this study is to perform electrophoretic and mass spectrometric analysis of BJP isolated

from the urine of six different patients, so as to demonstrate similarities and differences in the physicochemical properties and structures of BJP, which can be associated with the varying degrees of nephrotoxicity.

3.2. Experimental Methods

3.2.1. Isolation and Purification of Light Chains

Protein samples were obtained from Dr. Vecihi Batuman in Tulane University Medical School. Light chains were formerly isolated and purified from the urine of patients with multiple myeloma as reported by Pote et al. (89). Briefly, light chains were precipitated from urine with ammonium sulfate (55% to 90% saturation), dialyzed against distilled water and lyophilized into a dry powder form. Then, lyophilized crude protein was dissolved in buffer at pH 6.0, applied to a carboxymethyl-Sephadex (C-50) column and eluted with 0.6 mol/L NaCl. The fraction containing light chains was again dialyzed against distilled water and lyophilized. Six protein samples used in this study were enumerated from 1 to 6 and prepared by dissolving lyophilized powder in ultra pure water at a concentration of 40 μ M.

3.2.2. One- and Two-Dimensional Gel Electrophoresis

One-dimensional gel electrophoresis (1D-GE) was performed under both reducing and non-reducing conditions and using 15% polyacrylamide gels. The gels were prepared according to the Laemmli Buffer System (Appendix A). For reducing electrophoresis, 5 μ L of each sample was mixed with 20 μ L of reducing sample buffer (see Appendix A for composition of sample buffer), heated at 95 $^{\circ}$ C for 4 min and a total of 5 μ g of each sample was loaded onto gel. For non-reducing electrophoresis, SDS and 2-mercaptoethanol were excluded from sample buffer. Five μ L of each sample was mixed with 20 μ L of non-reducing sample buffer and loaded onto gel skipping the heating step. Electrophoresis was performed using Mini-Protean Electrophoresis System (Bio-Rad), running at 60 V for 1 h and then at 100 V until the dye front reached

the bottom of the gel. Following electrophoresis, gels were washed with distilled water and stained with colloidal Coomassie Brilliant Blue G-250.

For two-dimensional gel electrophoresis experiments (2D-GE), protein concentration of each sample was adjusted in rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 65 mM DTT, 2.5% pH 3-10 ampholyte solution) to a final concentration of 1 mg/mL. The first dimension, isoelectric focusing (IEF), was performed with non-linear pH 3-10, 17 cm IPG strips (Bio-Rad) using a Protean IEF cell (Bio-Rad). 350 μ g of total protein in 350 μ L of rehydration buffer was loaded on a strip; passive rehydration was performed for 2 h at room temperature followed by an active rehydration at 50 V for 16 h. The strips were then subjected to a total of 60000 Vh of electrophoresis. IEF steps were as follows; step 1: linear increase, 200 V, 300 Vh; step 2: linear increase, 500 V, 500 Vh; step 3: linear increase, 1000 V, 1000 Vh; step 4: linear increase, 4000 V, 4000 Vh; step 5: rapid increase, 8000 V, 24000 Vh; step 6: rapid increase, 8000 V, 30000 Vh. The second dimension, SDS-PAGE, was performed using Protean II XL System (Bio-Rad). Prior to SDS-PAGE, the strips were equilibrated first in equilibration buffer I (6 M urea, 0.375 M Tris-HCl pH 8.8, 2% SDS, 20% glycerol, 2% DTT) for 15 min and then in equilibration buffer II (6 M urea, 0.375 M Tris-HCl pH 8.8, 2% SDS, 20% glycerol, 2.5% iodoacetamide) for an additional 15 min. The strips were placed onto 12% polyacrylamide gels and then sealed with an agarose solution (0.5% agarose, 0.003% bromophenol blue in 1X Tris-Glycine-SDS Buffer). Electrophoresis was carried out at 16 mA for the first hour and then at 180 V until the dye front reached the bottom of the gel. Following electrophoresis, gels were washed with distilled water and stained with colloidal Coomassie Brilliant Blue G-250.

3.2.3. In-Gel Digestion and Peptide Extraction

Protein spots were excised from gel with a clean pipette tip, placed into clean micro centrifuge tubes and subjected to in-gel digestion protocol. First, each gel was divided into smaller pieces and 200 μ L of wash solution (50% methanol, 5% acetic acid in ultra pure water) was added on to each sample. The gel pieces were rinsed with wash solution on a shaker at room temperature, by several changes of the solution until all the dye is removed from gel pieces. The wash solution was removed, 200 μ L of acetonitrile was added and gel pieces were dehydrated at room temperature for 5 min. Acetonitrile

was removed and gel pieces were completely dried in a vacuum centrifuge for 2-3 min. Then, 30 μ L of 10 mM DTT solution, prepared in 100 mM ammonium bicarbonate solution, was added on each sample and proteins were reduced for 30 min at room temperature. Following incubation, DTT solution was replaced with 30 μ L of 100 mM iodoacetamide solution, prepared in 100 mM ammonium bicarbonate solution, and alkylation was performed for an additional 30 min at room temperature. Iodoacetamide solution was removed from each sample, 200 μ L of acetonitrile solution was added and gel pieces were dehydrated for 5 min at room temperature. Rehydration of gel pieces was achieved by addition of 100 mM ammonium bicarbonate solution and incubation of gel pieces for 10 min at room temperature. Dehydration step was repeated for the last time and then gel pieces were completely dried in a vacuum centrifuge for 2-3 min. Trypsin solution was prepared in 50 mM ammonium bicarbonate solution at a final concentration of 20 μ g/mL and 20 μ L of this solution was added to each sample. Digestion was carried out overnight at 37 °C. Peptides, obtained from digestion were extracted from gel pieces in a three-step procedure. First, 30 μ L of 50 mM ammonium bicarbonate solution was added to each sample and incubated for 10 min with occasional gentle vortex mixing. The sample was centrifuged and the supernatant was transferred into a clean micro centrifuge tube. Then, 30 μ L of extraction solution (50% acetonitrile, 5% formic acid in ultra pure water) was added to each sample and incubated for 10 min with occasional gentle vortex mixing. The supernatant after centrifugation was combined with the former extract. The last step was repeated once more, and the volume of peptide extract was reduced to 10 μ L in a vacuum centrifuge.

3.2.4. MALDI-MS/MS Analysis and Database Search

Before mass spectrometry analysis, each sample was desalted with C18-ZipTip (Millipore). Mass spectrometry experiments were conducted on an Autoflex III smartbeam MALDI TOF/TOF MS (Bruker Daltonics, Germany) instrument. Two-layer matrix preparation method was employed, using alpha-cyano-4-hydroxycinnamic acid (CHCA) as matrix. The first layer consisted of 0.6-1.2 mg CHCA dissolved in 100 μ L solution of 20% methanol in acetone and the second layer consisted of 1 mg CHCA dissolved in 100 μ L of 40% methanol in 0.1% TFA solution. One μ L of first layer was spotted onto MALDI target plate (MTP-384 massive target gold plated T, Bruker

Daltonics), followed by spotting 1 μL of peptide-matrix mixture, mixed at a ratio of 1:5, onto first layer and left air drying. The mass spectra of each sample were recorded in reflectron positive ion mode and from each mass spectrum; peptides were selected for fragmentation in MS/MS mode using argon as a collision gas. The data obtained from MS/MS analysis were searched against NCBIprot database (selecting Homo sapiens taxonomy) using Mascot search engine (V2.4, Matrix Science) for identification of the corresponding proteins. The search parameters were set as follows; 200 ppm mass tolerance in MS mode, 0.5 Da mass tolerance in MS/MS mode, cysteine carbamidomethylation as fixed modification, methionine oxidation as variable modification and allowing up to 1 missed cleavage.

3.2.5. LC-ESI-MS and MALDI-MS Analysis of Proteins

For LC-ESI-MS experiments, protein samples were prepared by dissolving lyophilized proteins in a solution of 0.1% TFA in ultra pure water. Five μg of each sample was injected onto a 10 cm x 2.1 mm Kromasil C₁₈ column coupled to a Dionex Ultimate 3000 Nano/Cap HPLC system (Dionex, Camberley, UK). Solvent A was 0.1% TFA in water and solvent B was 0.1% TFA in acetonitrile. The column temperature was set to 25 °C. Proteins were eluted from the column with a multistep gradient of B with the following method: 0 min, 0% B; 8 min, 12% B; 13 min, 27% B; 21 min, 39% B; 25 min, 44% B; 25.01 min, 45% B; 27 min, 45% B; 35 min, 60%B; 40 min, 65% B. The detection wavelength was set at 280nm. The solvent flow rate was 200 $\mu\text{L}/\text{min}$ and outlet of detector was directed to the ESI source of LTQ XL linear ion trap mass spectrometer (ThermoFisher Scientific, San Jose, CA), maintaining a flow rate of 10 $\mu\text{L}/\text{min}$ *via* a post-detector split. Data acquisition was performed in positive mode, scanning over the mass range m/z 400 to 2000. MALDI-MS analyses of intact proteins were performed on an Autoflex III smartbeam MALDI TOF/TOF MS, using sinapinic acid as matrix. The mass spectra were recorded in linear positive ion mode, scanning over the mass range m/z 5000 to 80000. The recorded data from both ESI- and MALDI-MS instruments were plotted using Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

3.3. Results and Discussion

The formation of different forms of BJP with distinct molecular weights enables the determination of these proteins via gel electrophoresis. Therefore, first of all, 1D-GE was performed under both reducing and non-reducing conditions so as to determine the differences in electrophoretic mobilities of the different BJP forms present in each sample. In reducing conditions, both covalent and non-covalent interactions are disrupted by heating the sample in the presence of a reducing agent, 2-mercaptoethanol. However, in non-reducing conditions, any kind of intermolecular interactions that may stabilize the possible dimer or multimeric complexes of light chains were not disturbed through excluding reducing agent from sample buffer and skipping the heating step. By this way, the relative amounts of monomers, dimers or other forms of light chains present in each sample can be estimated due to the differences in the mobilities of high and low molecular weight species. The electrophoretic patterns of the samples following colloidal coomassie staining are reflected in Figure 3.2.

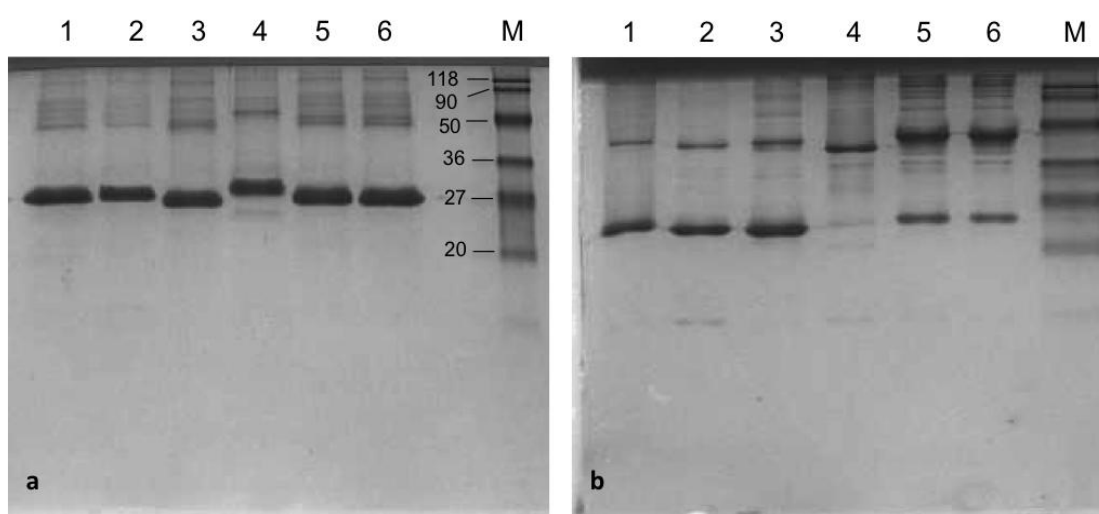


Figure 3.2. 1D-GE of samples performed under (a) reducing conditions, (b) non-reducing conditions. Lane numbers (written on top of each well) also reflect the sample numbers. M: Protein marker. Apparent molecular weight of marker bands (kDa) were given in (a)

In Figure 3.2a, it can be clearly seen that under reducing and denaturing conditions, light chains in each sample were reduced to their corresponding monomers and gave a single prominent band around 27 kDa. On the other hand, electrophoretic pattern of each sample under non-reducing conditions (Figure 3.2b) clearly

demonstrated that light chains in each sample exist in different forms in native conditions. Samples 1, 2, and 3 mainly consist of monomers with a molecular weight around 23 kDa, and a small amount of these samples exists as dimers with an estimated molecular weight of 46 kDa. In contrast to these samples, the abundance of light chain dimers was higher in samples 5 and 6, with a molecular weight around 48 kDa, and a little fraction exist as monomer form. Apart from those, sample 4 exhibited a totally distinct composition among the samples tested. Almost all the light chains in this sample exist as dimer form, confirmed by the single prominent band around 46 kDa, and a very little amount of monomer can be seen as a faint band around 23 kDa. In addition to highly abundant monomer and dimer light chain bands, some other protein bands with relatively lower color intensity were also detected in polyacrylamide gels. The other protein bands, observed in the upper parts of the gel in each sample (above 50 kDa), can be attributed to the association of light chains into multimeric complexes, whereas protein bands in the low molecular weight region (below 20 kDa) of the gel, observed mainly in samples 2 and 4, can be attributed to the fragments or truncated forms of light chains, as well as some contaminating proteins which were co-purified with BJP. As a result, on the basis of differences in their electrophoretic mobilities under non-reducing conditions, it was observed that the light chains obtained from different patients have different tendency to form multimeric complexes.

The distribution of monomer, dimer and other forms of light chains in samples were further demonstrated by MALDI-MS analysis and the recorded spectra are shown in Figure 3.3. Similar to the non-reducing 1D-GE results, these spectra clearly indicate the presence of multiple forms of light chains in each sample. The apparent molecular weights of each species in the samples are given in Da units in MS spectra. In addition to monomer and dimer forms (with peaks observed around 24 kDa and 46 kDa, respectively), light chain fragments (11.7 kDa), monomer-fragment complexes (approx. 35 kDa) and dimer-fragment complexes (approx. 58 kDa and 70 kDa) were also present in all samples at varying levels. These fragments and multimeric complexes were also observed as additional protein bands in 1D-GE, confirming the presence of these species via both electrophoresis and mass spectrometry. Likewise, the formation of multimeric complexes of light chains has been reported in a previous study, by identification of trimeric aggregates of light chain dimers in serum of a patient with multiple myeloma (83). Moreover, the intensity of peaks corresponding to different light chain forms in MALDI-MS spectra correlates with the intensity of protein bands

observed in 1D-GE, except for sample 4. Although non-reducing 1D-GE analysis revealed that the light chains in this sample are mainly in the dimer form, MALDI-MS spectrum shows that the most abundant species with highest peak intensity is light chain fragments. This observation could be explained by the dissociation of dimer into its fragments during ionization in the ion source or a relatively low ionization cross-section of dimer form of this light chain compared to its fragment.

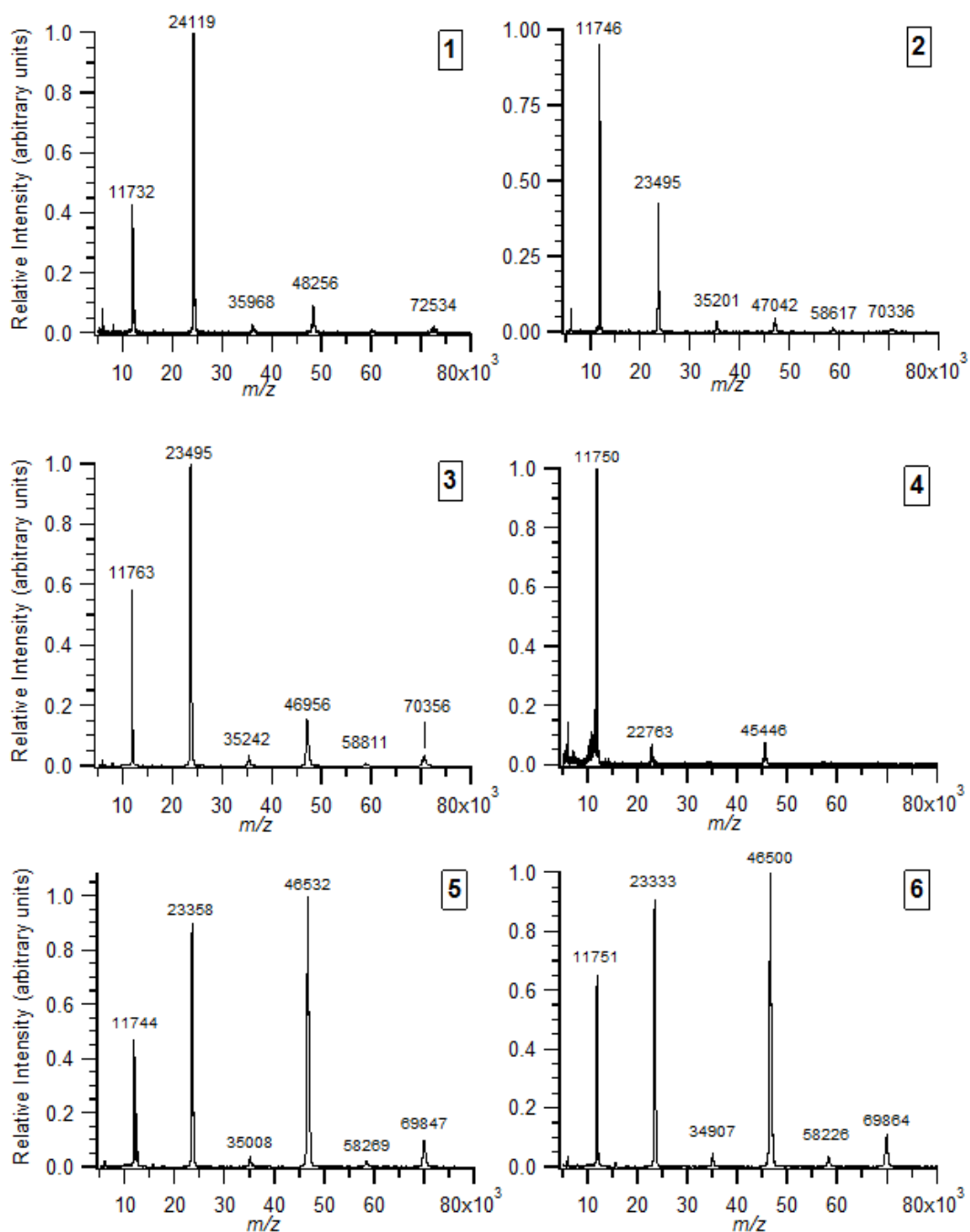


Figure 3.3. MALDI-MS spectra of samples. The apparent molecular weights of peaks are given in Da units

In determining the molecular weight of most abundant forms of light chains in each sample, LC-MS analyses were also performed. The samples were injected onto a reversed phase protein column, and the detector outlet was connected to ESI source of mass spectrometer. LC chromatogram and the MS spectrum of each sample were recorded simultaneously. The recorded LC profiles of the samples are given in Figure 3.4. The highest peak with the same retention time observed in each chromatogram was found to be BJP in each sample and confirmed by MS analysis. The MS spectra of these LC peaks are given in Figure 3.5 for each sample, reflecting the multiply charged ion distribution of light chains. It is known that ESI-MS analysis of large molecules like proteins produces multiply charged ions, opposed to singly or doubly-charged ions observed in MALDI-MS, and the number of charges tends to increase with increasing molecular weight of protein. This is due to the increase in the number of protonated sites on a protein during ionization in an ESI source, as the molecule gets bigger. When MS spectra of samples 1, 2 and 3 were compared with 4, 5 and 6, it is easy to notice the closer peaks in same mass region of the spectra and higher charged state ions in samples 4, 5 and 6, indicating the presence of high molecular weight species of light chains in these samples. Using these multiply charged ion series, molecular weights of BJP were calculated as 24.1 kDa for sample 1, 23.5 kDa for sample 2, 23.4 kDa for sample 3, 45.3 kDa for sample 4 and 46.4 kDa for samples 5 and 6. The calculated molecular weights of the monomer and dimer structures in each sample are in a good agreement with non-reducing 1D-GE and especially with MALDI-MS results, and further confirms the presence of monomers in samples 1, 2, and 3 and dimers in samples 4, 5 and 6 as highly abundant forms of BJP.

In order to demonstrate differences in both charge and molecular weight of samples, 2D-GE was performed and protein profiles of each sample are shown in Figure 3.6. The results revealed that light chains obtained from different patients have different net charges. The protein spots in samples 1, 2 and 4 are distributed mainly in neutral pH region of the gel, whereas protein spots in samples 3, 5 and 6 accumulated in acidic pH region of the gel. According to a graph of pH vs length of IPG strip, provided by the manufacturer, pI values of light chains in each sample were estimated. The samples 1, 2, and 4 have pI around 6-7 and samples 3, 5 and 6 have pI around 4-5. Earlier reports proposed that the physical properties of BJP are important in determining whether renal impairment develops and BJP with relatively high pI are more likely to be associated with renal dysfunction (101, 102). However, this association was opposed in later

studies and shown that renal impairment is not related to the pI of the BJP (103, 104). On the other hand, one can easily note the appearance of multiple protein spots arranged horizontally through a range of pH in 2D-GE of all samples. This kind of electrophoretic pattern of BJP under reducing 2D-GE was similar to that of observed in previous studies (76, 105). Harrison et al. reported the occurrence of multiple, equally spaced, light chain bands, with same molecular weight, in IFE studies of urine (106), later on in 2D-GE studies (107) and termed this as ladder light chain (or tiger stripe or pseudo-oligoclonal) pattern. This phenomenon reflects the presence of charge variants of light chain with apparently same molecular weight. Another study performed 2D-GE under non-reducing conditions and demonstrated the appearance of charge heterogeneity of both monomers and dimers of light chains in urine of patients with multiple myeloma and Bence Jones proteinuria (108). As a result, this pattern appears to be a common feature of BJP when analyzed by 2D-GE.

The protein spots in each sample were subjected to an in-gel digestion procedure and further analyzed by MALDI-TOF/TOF MS. Following tandem MS analysis, the data were searched against Homo sapiens database and the results are given in Table 3.1. The database search resulted in the identification of the proteins spots with different pI as well as molecular weight as Ig light chain in all samples. Although 2D-GE was performed in denaturing and reducing conditions, the protein spots in the upper part of the gel and in the middle part of the gel (dark and big protein spots aligned horizontally) were identified as same protein, reflecting the existence of light chain dimers still in reducing conditions. In addition, charge variants of light chains dimers, as well as monomers were clearly demonstrated in this study. A detailed examination of results shown in Table 3.1 highlights the identification of λ type light chain only in sample 4, whereas the light chains identified in other samples were all κ type. Previous studies have shown that, λ light chains mainly exist as covalent dimers and κ light chains usually exist as stable monomers or non-covalent dimers (83, 99). Taken this into consideration, it is now possible to explain why nearly all of the light chains in sample 4 are in the dimer form, as observed in non-reducing 1D-GE. From database search results, the given sequences of best matching proteins for each sample are shown in Figure 3.7. In this figure, the full sequence of each protein is given with matching peptides shown in bold red. Although MS/MS analysis and database search identified the same type of light chain (κ type) with the same accession number in samples 1, 5 and 6, the light chains in samples 5 and 6 appears to have high tendency to form dimers,

and the light chains in sample 1 mainly exist as monomers as revealed in 1D-GE. Accordingly, the light chains in sample 5 and 6 should bear some differences in terms of amino acid sequence in variable region. This may be an increase in the number of cysteine residues or an increase the number of amino acid residues, which take part in non-covalent interactions. These differences could in turn result in the formation of dimers stabilized through covalent or non-covalent interactions. However, with the available data, only partial sequences (about 23 % coverage) of the light chains were obtained, and the whole sequence could not be obtained. Therefore, considering amino acid sequences, it is not possible to figure out the factors that force dimerization of light chains in samples 5 and 6. It is necessary to note here that samples 5 and 6 were obtained from same patient at different stages of disease. By electrophoretic and mass spectrometry analysis, it was shown that these samples were almost the same. Therefore, in the conclusion of following results, only sample 5 will be considered.

Besides, numerous additional proteins were identified in the samples by MS/MS analysis. These are, β 2-microglobulin, protein AMBP (α 1-microglobulin), heparan sulfate proteoglycan, retinol binding protein and cystatin C. These are the urinary proteins that are found in certain amounts in normal conditions, but they may also increase in some diseases. The identification of these proteins is not a point of this study, and reflects the co-purification of these samples with light chains during precipitation and purification steps.

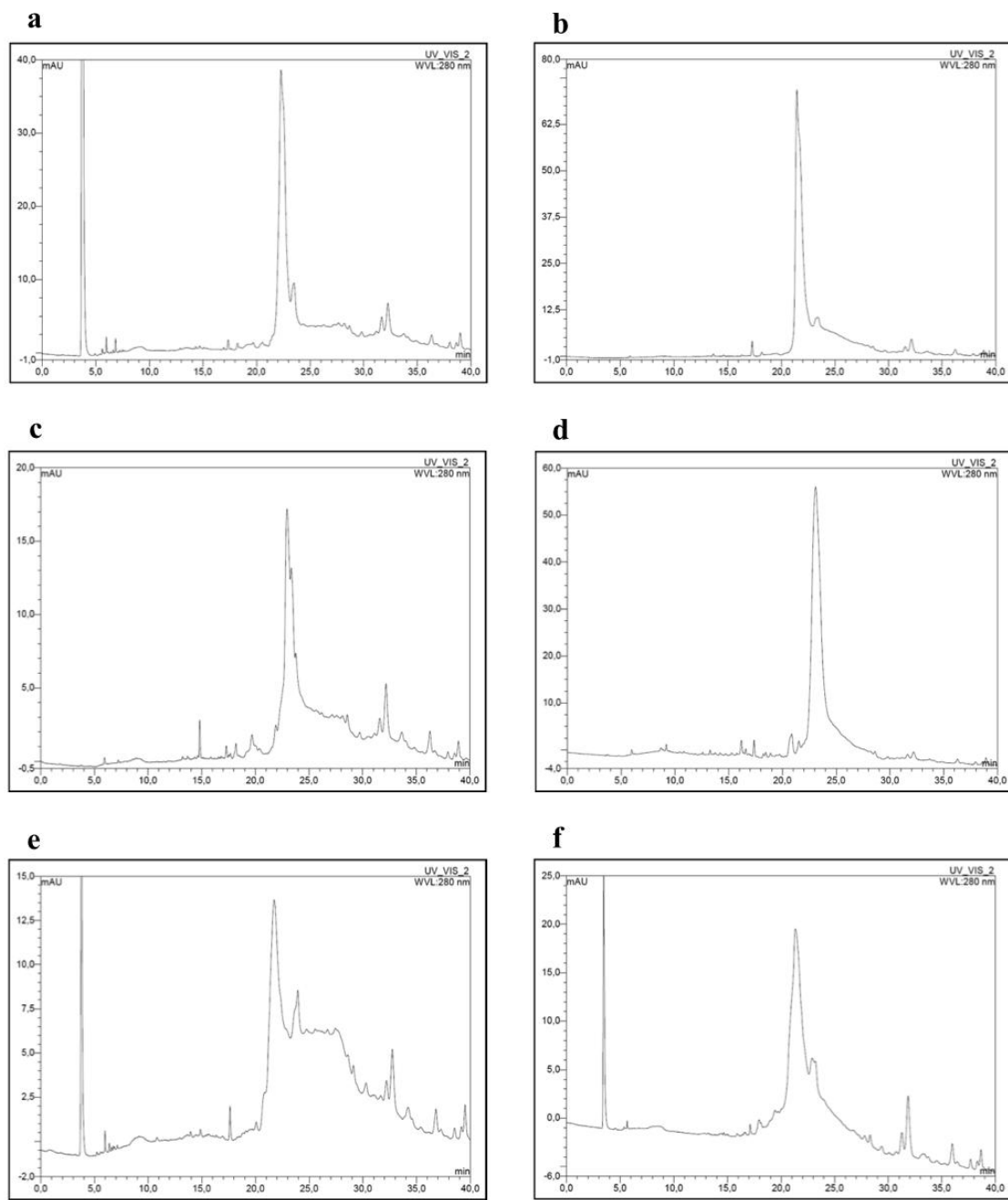


Figure 3.4. RP-LC profiles of samples (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, (f) 6

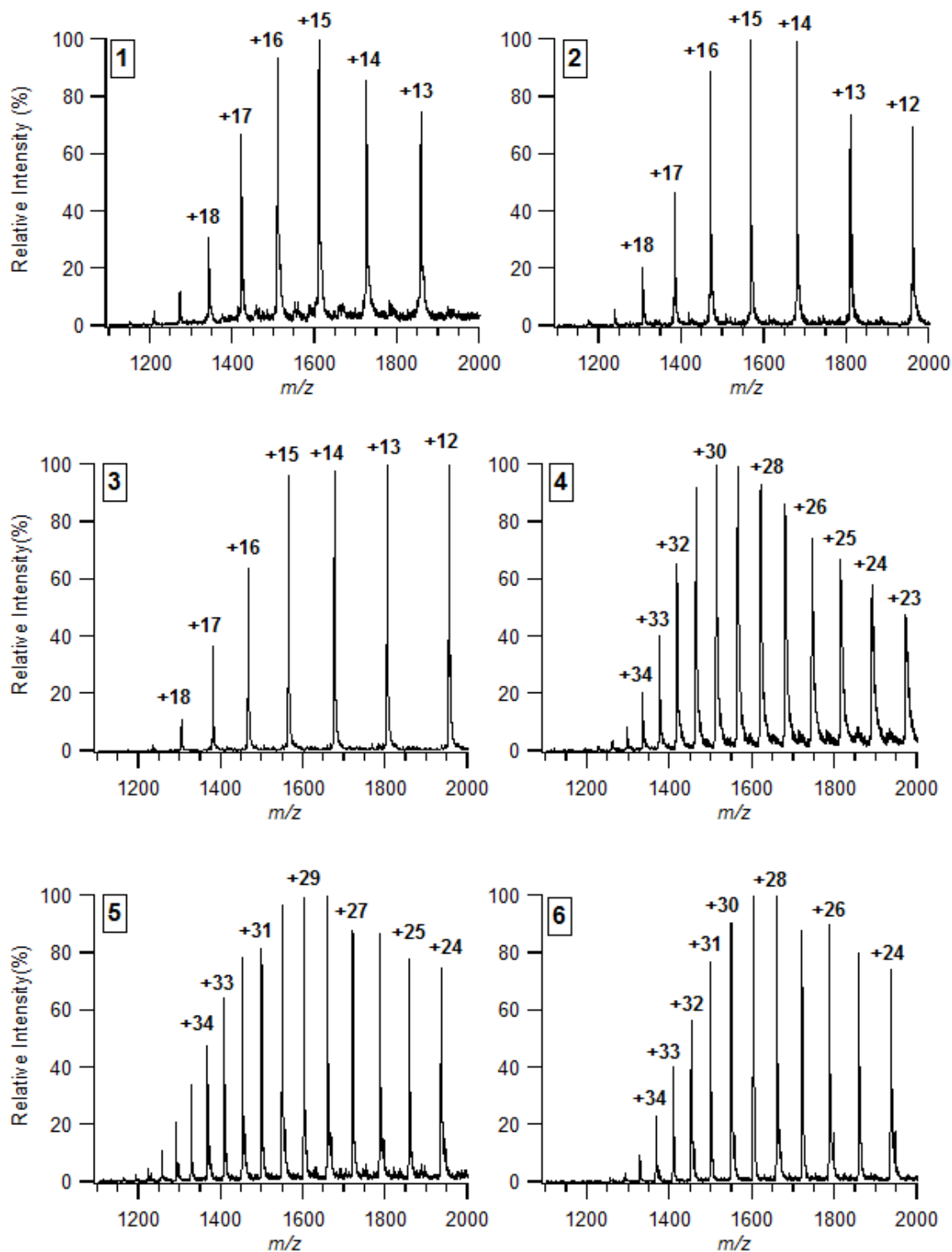


Figure 3.5. ESI-MS spectra of samples. The numbers given on top of each peak indicate the charge state of that peak.

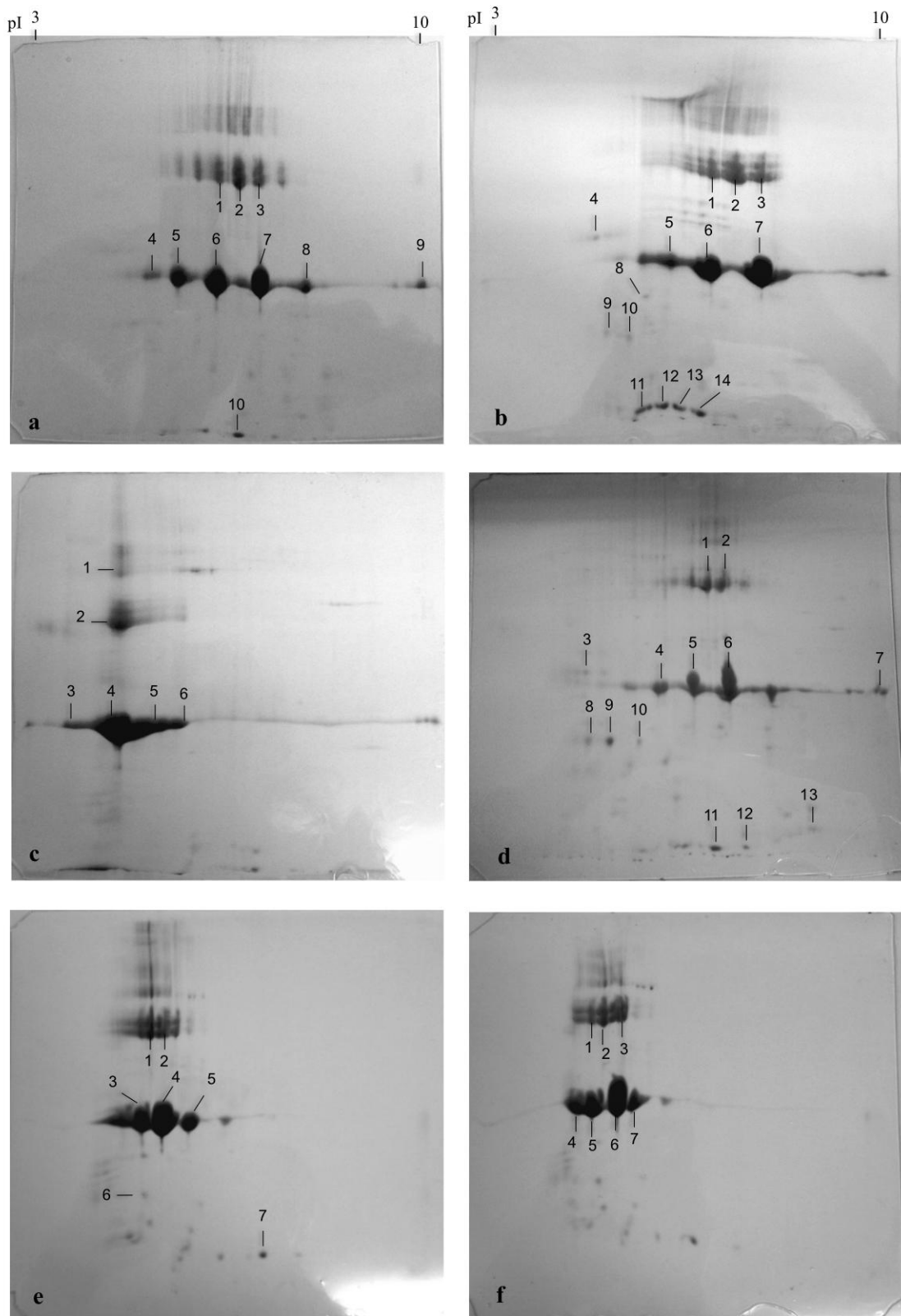


Figure 3.6. 2D-GE analysis of samples (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, (f) 6

Table 3.1. MS/MS identification of protein spots by database search

	Spot no	Protein Identity	Accession no	Mr	pI	Score
SAMPLE 1	Spot 1	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	51
	Spot 2	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	73
	Spot 3	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	478
	Spot 4	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	272
	Spot 5	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	411
	Spot 6	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	371
	Spot 7	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	317
	Spot 8	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	443
	Spot 9	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	134
	Spot 10	Beta-2-microglobulin precursor	ABI63355.1	10506	7.82	163
SAMPLE 2	Spot 1	Anti-rabies SOJA immunoglobulin kappa light chain	AAO17825.1	26116	6.15	134
	Spot 2	Anti-rabies SOJA immunoglobulin kappa light chain	AAO17825.1	26116	6.15	372
	Spot 3	Anti-rabies SOJA immunoglobulin kappa light chain	AAO17825.1	26116	6.15	238
	Spot 4	Protein AMBP preproprotein	NP 001624.1	39886	5.95	292
	Spot 5	Anti-rabies SOJA immunoglobulin kappa light chain	AAO17825.1	26116	6.15	328
	Spot 6	Anti-rabies SOJA immunoglobulin kappa light chain	AAO17825.1	26116	6.15	311
	Spot 7	Anti-rabies SOJA immunoglobulin kappa light chain	AAO17825.1	26116	6.15	229
	Spot 8	Heparan sulfate proteoglycan	AAA52700.1	47924	6.05	167
	Spot 9	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	196
	Spot 10	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	167

(cont. on next page)

Table 3.1 (cont).

	Spot no	Protein Identity	Accession no	Mr	pI	Score
SAMPLE 2	Spot 11	Beta-2-microglobulin	AAB25312.1	11481	5.86	104
	Spot 12	Beta-2-microglobulin	AAB25312.1	11481	5.86	61
	Spot 13	Beta-2-microglobulin precursor	ABI63355.1	10506	7.82	161
	Spot 14	Beta-2-microglobulin precursor	ABI63355.1	10506	7.82	163
SAMPLE 3	Spot 1	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	356
	Spot 2	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	339
	Spot 3	immunoglobulin light chain, partial	ACF34415.1	23883	5.35	330
	Spot 4	immunoglobulin light chain, partial	ACF34415.1	23883	5.35	413
	Spot 5	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	372
	Spot 6	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	269
SAMPLE 4	Spot 1	immunoglobulin lambda light chain	BAA00861.1	25870	7.60	48
	Spot 2	immunoglobulin lambda light chain	BAA00861.1	25870	7.60	69
	Spot 3	Protein AMBP preproprotein	NP 001624.1	39886	5.95	199
	Spot 4	immunoglobulin lambda light chain	BAA00861.1	25870	7.60	144
	Spot 5	immunoglobulin lambda light chain	BAA00861.1	25870	7.60	57
	Spot 6	IGL@ protein	AAH30983.1	25347	8.14	208
	Spot 7	immunoglobulin lambda light chain	BAA00861.1	25870	7.60	265
	Spot 8	Retinol binding protein 4 isoform a precursor	NP006735.2	23337	5.76	198
	Spot 9	Retinol binding protein 4 isoform a precursor	NP006735.2	23337	5.76	303
	Spot 10	Retinol binding protein 4 isoform a precursor	NP006735.2	23337	5.76	116
	Spot 11	Beta-2-microglobulin, partial	CAA23830.1	12905	5.77	268
	Spot 12	Beta-2-microglobulin, partial	CAA23830.1	12905	5.77	40
	Spot 13	Cystatin-C precursor	NP000090.1	16017	9.00	235

(cont. on next page)

Table 3.1 (cont).

	Spot no	Protein Identity	Accession no	Mr	pI	Score
SAMPLE 5	Spot 1	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	235
	Spot 2	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	168
	Spot 3	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	316
	Spot 4	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	263
	Spot 5	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	403
	Spot 6	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	45
	Spot 7	Beta-2-microglobulin, partial	CAA23830.1	12905	5.77	150
SAMPLE 6	Spot 1	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	197
	Spot 2	immunoglobulin light chain variable region, partial	AHZ09131.1	10614	4.33	76
	Spot 3	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	168
	Spot 4	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	295
	Spot 5	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	316
	Spot 6	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	293
	Spot 7	immunoglobulin kappa light chain VLJ region, partial	BAC01670.1	29116	6.71	403

The light chains analyzed in this study were obtained from the urine of multiple myeloma patients with light chain proteinuria and with minimal or no albuminuria. The patients suffered from renal dysfunction at varying degrees. However, kidney biopsies were not performed in any of these patients. In a previous study from Dr. Batuman's group (89), the effect of these light chains on thymidine incorporation by cultured human kidney proximal tubule cells was performed. Thymidine incorporation assay is a way of expressing DNA synthesis, through measuring the level of incorporation of radioactive nucleoside tritium (^3H)-thymidine into new strands of DNA during cell division. Therefore, decreased amount of measured ^3H -thymidine indicated a decrease in cell viability. The results of this study revealed that, all light chains inhibited thymidine incorporation, thereby DNA synthesis, in human proximal tubule cells at

varying levels (Figure 3.8a). Samples 1 (not shown), 3 and 5 were found to have minimal effect on thymidine incorporation, whereas sample 4 was found to be a more effective inhibitor than others. These results demonstrated the variability of light chain toxicity among different species. However, it was stated that the variability observed among different light chains did not show a correlation with either severity of clinical disease or immunologic type. In addition, proximal tubule cells were incubated with different concentrations of light chains from samples 3, 4 and 5 and the cell viability was determined again by thymidine incorporation assay. The results of this experiment are shown in Figure 3.8b and indicated that toxicity of light chains is dose-dependent. None of the light chains resulted in a significant level of cell death at concentrations up to 100 μ M. However, above this concentration they exhibited varying levels of inhibition of thymidine incorporation. According to this finding, it was shown that the amount of light chain is an important parameter of nephrotoxicity and increased amount of light chains might be related with renal involvement in multiple myeloma.

Sample 1					Sample 2						
1	MKYLLPTAAA	GLLLAAQPA	MADIQMTQSP	SSLASVGDGR	VTITCRASQS	1	MEAPQLLFL	LLLWLPDITG	EIVLTQSPAT	LSLSPGERAT	LACRASQTAS
51	ITNYLNWYQQ	KPGKAPNLLI	YAASSLQSGV	PSRFGSGSGG	TDFTLTISL	51	RYLAWYQKPK	GQAPRLLIYD	TSNRATGIPA	RFGSGSGGTD	FTLSISSLEP
101	QPEDFATYYC	QQSYSTPTTF	GGGTKVEIKR	TVAAPSVFIF	PPSDEQLKSG	101	EDFAVYYCQQ	RFNWPWTFGQ	GTKVEFKR TV	AAPSVFIFPP	SDEQLKSGTA
151	TASVVCLLNN	FYPREAKVQW	KVDNALQSGN	SQESVTEQDS	KDSTYLSLST	151	SVVCLLNIFY	PREAKVQWKV	DNALQSGNSQ	ESVTEQDSK D	STYLSLSTLT
201	LTLISKADYEK	HKVYACEVTH	QGLSSPVTKS	FNRGEC SARQ	STPFVCEYQG	201	LSKADYEKHK	VYACEVTHQG	LSSPVTKSFN	RGEC	
251	QSSDLPQPPV	NAGGSGGGGS	G								
Sample 3					Sample 4						
1	DIQMTQSPSS	LSASVGRVTV	ITCQASQDIS	NYLWNFQKPK	GKAPNLLIYD	1	MAWTPDLLLF	LSHCTGSLSQ	AVLTQPSSLS	ASPGTSASLT	CTLRSDINVG
51	ASNLETGVPV	RFGSGSGGTN	FTFTIHSLQP	EDIATYYCQQ	YDNLPPFTFG	51	SPFIYWYQK	PGSPPQFLLR	YKSDSDNQQG	SGVPSRFGSG	KDASANAGIL
101	PGTKVDFKRT	VAAPSVFIFP	PSDEQLKSGT	ASVVCLLNFF	YPREAKVQWK	101	LISGLQSEDE	ADYYCMINRG	TAVVFGGGTK	LTVLGGPK AA	PSVTLFPPSS
151	VDNALQSGNS	QESVTEQDSK	DSTYLSLSTL	TLISKADYEK	KLYACEVTHQ	151	EELQANKATL	VCLISDFYPG	AVTVAMKADS	SPVKAGVETT	TPSKQSNKY
201	GLSSPVTKSF	NRGEC				201	AASSYLSLTP	EQWKSHRSYS	CQVTHEGSTV	EKTVPARTERS	
Sample 5					Sample 6						
1	MKYLLPTAAA	GLLLAAQPA	MADIQMTQSP	SSLASVGDGR	VTITCRASQS	1	MKYLLPTAAA	GLLLAAQPA	MADIQMTQSP	SSLASVGDGR	VTITCRASQS
51	ITNYLNWYQQ	KPGKAPNLLI	YAASSLQSGV	PSRFGSGSGG	TDFTLTISL	51	ITNYLNWYQQ	KPGKAPNLLI	YAASSLQSGV	PSRFGSGSGG	TDFTLTISL
101	QPEDFATYYC	QQSYSTPTTF	GGGTKVEIKR	TVAAPSVFIF	PPSDEQLKSG	101	QPEDFATYYC	QQSYSTPTTF	GGGTKVEIKR	TVAAPSVFIF	PPSDEQLKSG
151	TASVVCLLNN	FYPREAKVQW	KVDNALQSGN	SQESVTEQDS	KDSTYLSLST	151	TASVVCLLNN	FYPREAKVQW	KVDNALQSGN	SQESVTEQDS	KDSTYLSLST
201	LTLISKADYEK	HKVYACEVTH	QGLSSPVTKS	FNRGEC SARQ	STPFVCEYQG	201	LTLISKADYEK	HKVYACEVTH	QGLSSPVTKS	FNRGEC SARQ	STPFVCEYQG
251	QSSDLPQPPV	NAGGSGGGGS	G			251	QSSDLPQPPV	NAGGSGGGGS	G		

Figure 3.7. The full sequence of light chains obtained by database search from; spot 3 of sample 1, spot 2 of sample 2, spot 4 of sample 3, spot 7 of sample 4, spot 4 of sample 5 and spot 7 of sample 6. The matched peptides are shown in bold red

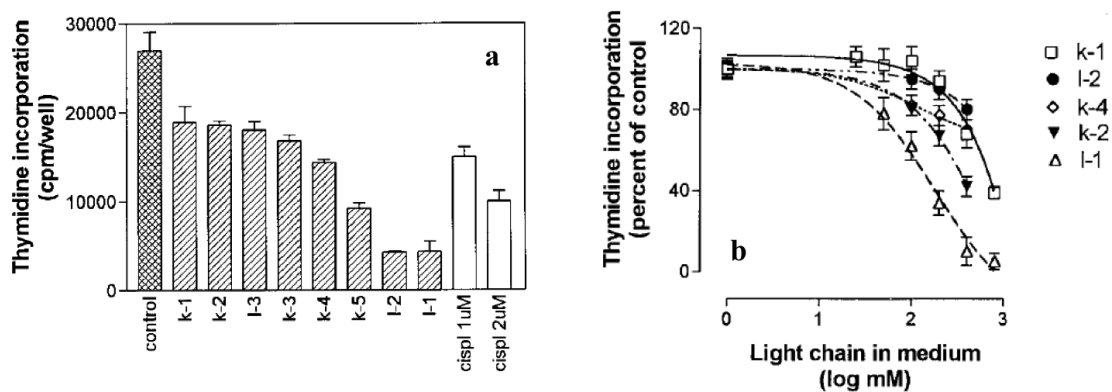


Figure 3.8. Effect of (a) light chains (400 μ M) and (b) different concentration of light chains on 3 H-thymidine incorporation by human proximal tubule cells (89). k-1 represents sample 5, k-2 represent sample 3 and l-2 represents sample 4

Other studies from Dr. Batuman's group made important contributions to elucidate mechanism of renal involvement in light chain proteinuria. It has been shown that light chains in proximal tubule cells are endocytosed through binding to endocytic receptors megalin and cubilin and degraded in lysosomes. Increased levels of light chains cause excessive endocytosis (109, 110), which then activates the transcription factors triggering nuclear transcription of inflammatory and pro-inflammatory cytokines (90). The nuclear transcription factor NF- κ B mediates inflammation reactions through its role in the expression of pro-inflammatory cytokine genes (111) and these cytokines are responsible for the generation of inflammation (112). Sengul et al. have shown that increased light chain endocytosis activates NF- κ B and production of inflammatory cytokines IL-6 (interleukin-6), IL-8 and MCP-1 (monocyte inflammatory protein) in cultured human proximal tubule cells (90). The effect of different light chains on production of IL-6 is shown in Figure 3.9. This figure clearly demonstrated that all tested light chains resulted in the production of IL-6, while human serum albumin (HSA), which is observed at high levels in proteinuria patients along with light chains, had no significant effect. This result highlights the involvement of light chains in pro-inflammatory cytokine production. In this figure, LC-1 represents sample 5, and LC-2 represent a lambda type of light chain. The increase in IL-6 production is clearly observed in these samples among the other light chains tested and again points the variability among light chains. It has also been demonstrated that light chains at concentration as low as 25 μ g/ml, concentrations that can be easily achieved in multiple myeloma or proteinuria patients, resulted in the production of cytokines. Authors have

concluded that light chain induced cytokine production by proximal tubule cells is an important mechanism of myeloma kidney observed in multiple myeloma patients.

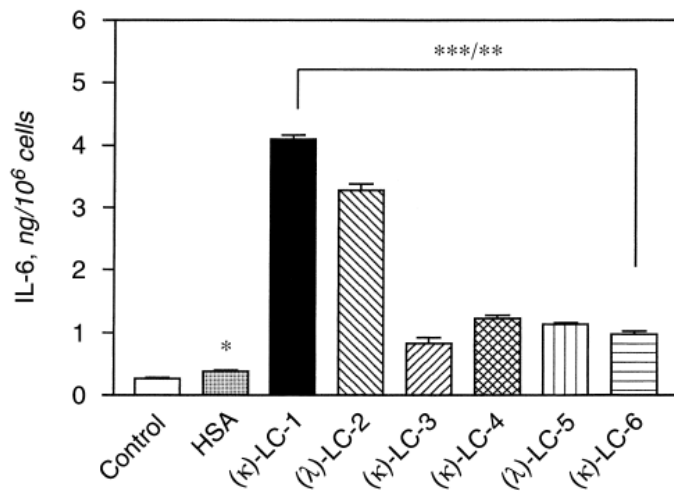


Figure 3.9. Effect of different light chains and HSA on IL-6 production in proximal tubule cells (90). LC-1 represents samples 5

3.4. Conclusion

In this study, BJPs isolated from urine of multiple myeloma patients were investigated by electrophoresis and mass spectrometry. 1D-GE of samples in non-reducing conditions, where the samples exhibit their native structures, revealed that samples 1, 2 and 3 exist mainly as monomers and samples 4, 5 and 6 exist mainly as dimers. This observation was further confirmed by MALDI-MS and ESI-MS analyses of samples. These analyses also allowed the calculation of apparent molecular weights of major forms of BJPs in each sample. Moreover, MALDI-MS analysis highlighted the association of different forms of light chains into high molecular weight structures, such as monomer-fragment or dimer-fragment complexes. Later on, 2D-GE analyses demonstrated the existence of charge variants of monomers and dimers in each sample by observation of horizontally arranged multiple protein spots through a range of pH values. Isoelectric points of BJPs in each sample were estimated by 2D-GE. Light chain monomers, as well as dimers, in sample 1, 2 and 4 were distributed in neutral pH range, whereas those of samples 3, 5 and 6 were distributed in acidic pH range. The appearance of charge variants of BJPs is a common pattern observed in 2D-GE of these

proteins. However, neither observation of charge variants nor differences in pI were associated with differential nephrotoxicity of BJPs. The type and partial amino acid sequences of BJPs in each sample were identified by MS/MS analysis of protein spots from 2D-GE gels and according to these results only sample 4 contained λ type chain and other samples contained κ type chain. Given the tendency of λ light chains to exist as stable dimers, this finding was expected for sample 4, because almost all of the light chains in this sample exist as dimer structure. However, it was not possible to figure out factors that force dimerization of BJPs in samples 5 and 6, still it can be assumed that variations in amino acid sequence in variable region of these light chains may trigger formation of covalent or non-covalent dimers. In addition, previous studies that used the same samples shed light on differential nephrotoxicity of these samples. Following incubation of proximal tubule cells with light chains, thymidine incorporation assay was performed and the results revealed that samples 1, 3 and 5 do not exhibit considerable effects on DNA synthesis in proximal tubule cells, whereas sample 4 resulted in the observation of higher level of DNA synthesis inhibition. Following experiments highlighted the importance of BJPs concentration on nephrotoxicity; such that elevated concentrations of each sample resulted in a high level of inhibition of DNA synthesis. Lastly, previous studies demonstrated that light chains induce production of inflammatory cytokines, such as IL-6, and sample 5 have the ability to induce production of IL-6 in proximal tubule cells. As a conclusion, it was shown that nephrotoxicity of BJPs is a complex phenomenon observed in multiple myeloma patients. The toxic effects of these proteins cannot be directly associated with observed physicochemical properties, such as pI, molecular weight or ability to form multimeric structures. All types of light chains used in this study were shown to have toxic effects at high concentrations. It is evident that, concentration of the light chains that the proximal tubule cells are exposed or the increased levels of endocytosis followed by production of inflammatory cytokines are important parameters in the development of renal failure in multiple myeloma patients. Among the samples tested, light chains in samples 4 and 5 exhibited some distinct properties in terms of inhibition of DNA synthesis and inducing the production of IL-6, which could help in explaining the high nephrotoxic potential of these samples.

CHAPTER 4

MASS SPECTROMETRY OF INTACT PROTEINS REVEALS +98 U CHEMICAL ARTIFACTS FOLLOWING PRECIPITATION IN ACETONE

4.1. Introduction

Appropriate sample preparation and purification is a key step in proteomic analysis. The success of proteomic studies largely depends on the purity and the integrity of the analyte. Tissue and cell extracts are complex in composition, and may include nucleic acids, lipids, salts, detergents, reducing agents, etc. (113). These components have an adverse impact on protein separation (gel electrophoresis, HPLC) and MS analysis. Therefore, suitable methods as a means to concentrate and purify proteins should be employed. Protein precipitation is routinely employed for small or large-scale sample purification, being favored ahead of proteome analysis by gel electrophoresis, and LC-MS analysis. Several strategies exist to induce precipitation, including changing the solution temperature (114) or pH (115), through addition of salts (116) (e.g., ammonium sulphate), polymers (117) (e.g., polyethylene glycol), or organic solvents (118) (e.g., methanol, acetone, ethanol). An optimal precipitation strategy reliably recovers purified proteins in high yield, but without altering the sample, which would result in the formation of unintentional chemical modifications. Such artifacts would propagate throughout the detection workflow, and finally impact the success of sample characterization.

Precipitation of proteins is mainly based on altering the solution conditions, thereby creating an environment of low solubility and increased protein-protein interactions, which then results in protein aggregation and precipitation. Normally, proteins in aqueous solutions adopt a conformation that exposes hydrophilic regions to the surroundings, enabling polar interactions with the aqueous solvent and formation of a hydration layer around the protein. Protein solubility in aqueous solution results from repulsive electrostatic forces between like charged ions and the hydration layer that

shields the protein-protein interactions (119, 120). Addition of organic solvents, as a precipitation method, lowers the dielectric constant of the solution. In a solution of low dielectric constant, proteins tend to aggregate and precipitate driven by the increased attractive forces between oppositely charged ions on proteins.

With potential for exceptional recovery and purity, organic solvent precipitation has steadily risen in popularity as a front-end tool for proteomic analysis. Several groups have demonstrated the advantages of acetone precipitation in bottom-up and top-down proteomics workflows and made improvements in order to obtain higher protein recovery. The importance of ionic strength in maximizing protein recovery with acetone precipitation has been noted (119). Crowell *et al.* reported that protein recovery through acetone precipitation is improved by including 1-30 mM NaCl together with acetone. Under the proper conditions, near quantitative recovery is obtained with high repeatability, including for dilute protein samples. Protein precipitation in acetone has also been shown to be an effective SDS depletion approach prior to MS analysis. SDS is routinely used in biochemical studies for total solubilization of cells and tissues. However, detergents such as SDS are detrimental for MS analysis, such that they dominate mass spectra due to their higher ionization ability and higher abundance compared to individual peptides (121). In addition, SDS results in the observation of adducts spaced by 266 u in deconvoluted protein mass spectra. Puchades *et al.* evaluated acetone precipitation for SDS depletion ahead of MALDI-MS analysis and reported 100-fold reduction of SDS (122). Acetone precipitation has also been employed for SDS depletion in top-down (123) and middle-down (124) proteomic workflows. An alternative approach for SDS depletion is filter-aided sample preparation method (FASP), in which molecular weight cut-off filters are used for the removal of low molecular weight components and detergents from protein mixtures, as well as performing protein digestion and peptide elution in the same filter unit (121). Acetone precipitation approach outperforms the popular FASP method for SDS removal, affording greater identification of peptides by MS (125). It has been demonstrated that acetone-precipitated proteins initially containing SDS provide higher sequence coverage than an equivalent tryptic digest of the non-precipitated, surfactant-free proteome mixture (125, 126).

Although analyte recovery and purity are important considerations of protein sample preparation in proteomics workflow, one must also consider *in vitro* protein modifications induced as artifacts of sample handling. A number of chemical

modifications formed during sample handling have been reported in literature and additional types of such modifications may exist. These modifications are common to several proteomic workflows, including acrylamide adduct formation on cysteine residues during SDS-PAGE (127), methylation of aspartate and glutamate residues by methanol present in staining/destaining solutions (128), oxidation of methionine, tryptophan and cysteine residues (129), esterification of aspartate and glutamate residues by glycerol (130), urea carbamylation (131), and formylation of serine, threonine and lysine residues by formic acid (132, 133). These artifacts increase sample complexity and challenge quantitative analysis, particularly if the mass shift matches that of *in vivo* protein modification. If the modifications are known, steps can be taken to minimize the reaction; alternatively the resulting mass shifts can be accounted for during data analysis. Unknown chemical modifications are detrimental to proteome analysis, as the unknown mass shifts will hamper proper MS identification and increase the false discovery rates.

Although acetone precipitation has been employed by a great number of researchers for a long time, it was only recently reported that acetone precipitation could manifest artifacts in the resulting MS spectra. Acetone is not chemically inert; primary amines are reactive towards ketones to yield imines (134, 135). A reaction with acetone at lysine or the N-terminus of a protein or peptide would form a ketimine, with a hypothetical mass shift of +40 u. Schiff base formation is pH sensitive, being labile under the acidic conditions typically encountered during LC-ESI-MS. However, as reported by Simpson *et al.*, a +40 u artifact was observed in the mass spectrum of peptides exposed to traces of acetone under acidic conditions (136). The modified peptides contained glycine at the second position (XGX n) and account for some 5% of the total proteome, though the exact nature of this chemical reaction was not determined. To date this is the only reported unintentional modification of peptides or proteins induced by acetone. Still other reactions with acetone may be possible. Aldol condensation of acetone forms mesityl oxide (Mr 98 u), which proceeds through the intermediate diacetone alcohol (Mr 116 u) (137). Each of these compounds may in turn react with proteins.

While investigating acetone precipitated intact proteins by ESI-MS, multiple satellite ions, spaced +98 u from the expected protein signal were observed. This type of modification induced by acetone has never been reported in literature. Therefore, the aim of this study is to investigate this novel chemical modification occurred during

precipitation of proteins in acetone. To that end, several factors such as incubation time, temperature and solution pH that are believed to influence the formation of this artifact are investigated. Moreover, the extent of modification is quantified for various proteins and peptides using mass spectrometry. The origin of this novel type of acetone modification as well as its implications in proteome analysis workflows were discussed. The motivation is to determine how this product forms, such that the unwanted modification can be eliminated from conventional acetone precipitation protocols.

4.2. Experimental Methods

4.2.1. Protein Precipitation

Stock solutions of standard proteins (cytochrome c, myoglobin, ubiquitin, bovine serum albumin, hemoglobin) were prepared in water from lyophilized powder to a concentration of 1 g/L. Peptide samples were prepared in water at 100 μ M concentration. Working solutions of the proteins were diluted to 0.1 g/L in water, or water with one of the following pH adjusters: 0.1% TFA, 0.1% formic acid, 0.1% acetic acid, 50 mM ammonium bicarbonate, 50 mM Tris-HCl (pH 8), 10 mM NH_4OH , or 2 M NH_4OH . The sample was then mixed with four volumes of ice-cold acetone and incubated at a specified temperature (-20 $^{\circ}\text{C}$, unless otherwise indicated) for a defined period (1 hour, unless otherwise stated). Samples were then fully dried in a SpeedVac vacuum concentrator.

Proteins were also precipitated in methanol, acetonitrile, or acetonitrile containing either 0.1 % (v/v) mesityl oxide (MO) or 0.1 % diacetone alcohol (DAA). The conditions for precipitation in these solvents, including protein concentration, solvent ratio, temperature and time are as described above.

4.2.2. *E. coli* Proteome Extraction and Precipitation

E. coli was inoculated into LB media and grown overnight at 37 $^{\circ}\text{C}$ with shaking. Cells were harvested at an OD_{600} of 1.0 by centrifugation (15 min, 5000 $\times g$) then washed with PBS buffer, and finally with water. The cells were suspended in 50

mM Tris-HCl buffer (pH 8.0), then snap frozen in liquid nitrogen and ground via mortar and pestle. The resulting cell lysate was heated to 95 °C for 5 minutes, and then centrifuged (30 min, 16000 ×g, 4 °C). The protein content in the supernatant, as measured using BCA assay, was adjusted in Tris buffer (pH 8) to a final concentration of 0.5 g/L. One hundred microliters of the extracted *E.coli* proteins was subjected to acetone precipitation protocol as described above (4:1 volume ratio, -20 °C, 1 hour).

4.2.3. Preparation of Apo-Cytochrome c

Removal of heme from cytochrome c was achieved by a modification of the method of Fisher *et al.* (138). To 60 µL of 10 mg/ml cytochrome c, 1 mg of AgNO₃ in 113 µL of water and 10 µL of acetic acid were added. The mixture was incubated in the dark for 4 hours at 40 °C then centrifuged to pellet the heme (15 min, 10,000 ×g). The protein in the supernatant was purified by HPLC using a self-packed 10 cm × 1 mm POROS R2 column (20 µm beads, Applied Biosystems) and collected as a single fraction by a rapid solvent ramp of acetonitrile from 5 to 50 %. The solvent was evaporated in a SpeedVac vacuum concentrator and the protein was dissolved in 100 µL of 50 mM ammonium acetate (pH 5.0), which contains 6 M guanidine HCl and 1 M DTT. The solution was incubated at room temperature for 2 hours in the dark and then centrifuged (15 min, 10,000 ×g). The supernatant was again injected onto a POROS R2 column, collecting the purified apo-cytochrome c as a single fraction, wherein the protein has no measurable absorbance at 410 nm.

4.2.4. Protein Digestion

50 µg of *E. coli* proteome or 50 µg BSA were subjected to trypsin digestion following precipitation. The pellet was resuspended in 20 µL of 8 M urea with repeat pipetting followed by 5 min sonication to aid in sample dispersion. Next, 80 µL of 50 mM ammonium bicarbonate was added and proteins were reduced through addition of 4.75 µL of 200 mM DTT (20 min, 60 °C), then alkylated with 10.5 µL of 200 mM iodoacetamide (20 min, 20 °C). Digestion occurred following addition of 1 µg trypsin to the *E. coli* extract, or 2.5 µg trypsin for BSA, with incubation at 37 °C overnight.

Digestion was terminated through addition of 10% TFA, then peptides were desalted by reversed phase HPLC cleanup using a self-packed 10 cm × 1 mm C18 column (5 μm beads, Waters, Mississauga, CA) and collected as a single fraction by a rapid solvent ramp of acetonitrile from 5% to 85%.

4.2.5. LC-MS Analysis of Intact Proteins or Peptides

Following precipitation, the dried protein pellets were resolubilized in 100 μL 15 % acetonitrile, 0.1% formic acid in water by incubation in ultrasonic bath for 15 min. The entire sample was immediately injected onto a self-packed 30 cm × 0.5 mm POROS R2 column coupled to an Agilent 1200 HPLC system. Solvent A was 0.1% formic acid in water and B was 0.1% formic acid in acetonitrile. The following gradient was employed: 0 min, 15% B; 5 min, 15% B, 5.01 min, 50% B; 15 min, 50% B; 15.01 min 15% B. The solvent flow rate (100 μL/min) was split post-column to direct 10 μL/min to the ESI source of a ThermoFisher LTQ linear ion trap mass spectrometer (San Jose, CA). Data acquisition was in positive “MS only” mode, scanning over the range m/z 400 to 2000. For exact mass determination of the mass adduct on intact protein, acetone-modified cytochrome c was subject to analysis on an Orbitrap Velos Pro (ThermoFisher Scientific), employing a self-packed 30 cm × 75 μm spray tips (New Objective, Woburn, MA) containing 4 μm C12 beads (Phenomenex, Torrance, CA) and coupling to nanospray at a flow rate of 0.25 μL/min. The isocratic solvent system was 50% acetonitrile in water with 0.1% formic acid.

Peptide experiments were conducted on a LTQ XL linear ion-trap mass spectrometer (ThermoFisher Scientific) for MS and MS/MS analysis. The peptides were further analyzed on an Orbitrap Exactive mass spectrometer (ThermoFisher Scientific) for exact mass determination of artifacts observed on peptides. The samples were introduced into the ESI source of both instruments *via* syringe pump and the systems were operated in the positive mode, scanning over the mass range m/z 150 to 2000.

The data obtained for intact protein from Orbitrap Velos Pro was deconvoluted using Xcalibur software which accompanies the Orbitrap platform, whereas other data for intact protein were deconvoluted using Microsoft Excel. MS spectra of intact proteins and peptides were plotted using Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

4.2.6. LC-MS/MS of Tryptic Peptides and Database Search

For BSA, 1 pmol of desalted peptides was analyzed on the LTQ linear ion trap mass spectrometer using dual capillary LC-MS/MS system (139), employing self-packed 30 cm × 75 μm spray tips (New Objective, Woburn, MA) containing 4 μm C12 beads (Phenomenex, Torrance, CA) and coupling to nanospray at a flow rate of 0.25 μL/min. The same solvent system was employed as for intact protein analysis, through with the following gradient: 0 min, 5% B; 0.1 min, 7.5% B; 45 min, 20% B; 57.5 min, 25% B; 60 min, 35% B; 61 min, 80% B; 64.9 min, 80% B; 65 min, 5% B. The LTQ operated in data-dependent mode (MS followed by MS/MS of top three peaks) with 30 s dynamic exclusion. The *E. coli* digest was analyzed on an Orbitrap Velos Pro. One microgram of the desalted peptide digest was injected onto a self-packed 15 cm × 0.1 mm monolithic C18 column (Phenomenex) using a Dionex Ultimate 3000 Rapid Separation LC nanosystem (Bannockburn, IL). The solvent gradient was as follows: 0 min, 3% B; 3 min, 3% B; 5 min, 5% B; 69 min, 35% B; 72 min, 95% B; 77 min, 95% B; 80 min, 3% B. The LC column was coupled to a 10 μm New Objective PicoTip non-coated Emitter Tip (Woburn, MA). The Orbitrap operates at a resolution of 30,000 FWHM in MS mode, scanning in data-dependent mode (MS/MS of top ten peaks) at a scan rate of 66,666 u/s, <0.6 u FWHM. Dynamic exclusion was applied for 25 s over a range of ±5 ppm. MS/MS data were searched using Proteome Discoverer software against the *E. coli* K12 UniProt database (downloaded on May 2014, 4269 entries) with a mass tolerance of 3 ppm (high res MS mode) and 0.8 u (MS/MS mode), or against a BSA sequence with a mass tolerance of 0.8 u (low res MS mode) and 1 u (MS/MS mode). The searches allowed for up to two missed cleavages, assigning a peptide false-positive rate of 1%, and minimum two peptides per protein. Peptide modifications included oxidation of methionine and carbamidomethylation at cysteine. Protein modification caused by incubation in acetone was first searched by assigning +98.0732 u dynamic modification to all amino acid residues. The data were searched again, confining +98.0732 u to histidine, lysine or arginine residues. Finally, data were searched for a +40 u dynamic modification on all amino acids as well as the N-terminus.

4.3. Results and Discussion

Acetone precipitation of proteins is favored for purification and concentration of proteins prior to chromatographic and electrophoretic separation as well as MS analysis, due to the detrimental effects of contaminants (such as detergents) in these analyses. Obtaining a high quality MS spectrum is critical to the success of proteomics workflow, through proper identification of proteins. For that purpose, the methods used in the proteomic workflow should provide minimal loss of proteins and not introduce modifications on proteins, at least in an unintentional way.

During characterization of acetone precipitated intact proteins by ESI-MS, a set of satellite peaks, successively spaced 98 u above the mass of the unmodified protein were observed. Figure 4.1 shows the multiply charged ESI-MS spectrum of cytochrome c following overnight incubation in 80% acetone at -20 °C (Figure 4.1a) relative to the non-precipitated control (Figure 4.1b). The deconvoluted spectrum of cytochrome c precipitated in acetone (Figure 4.1c) depicts the intense satellite ion series spaced successively at +98 u. These multiple peaks impact the overall signal intensity and lower spectral quality. In this figure, the artifacts account for 56% of the total protein signal. Besides the increased complexity in sample MS spectrum, the intensity of the intact protein was decreased, such that the absolute intensity of the unmodified cytochrome c represents only 20% of the non-precipitated control. Unlike incubation in acetone, non-precipitated proteins (Figure 4.1b and Figure 4.1d) and samples precipitated with other organic solvents such as methanol or acetonitrile (Figure 4.2) are free of +98 u artifacts. These results demonstrated that the modification is specific to acetone. The +98 u artifacts were also observed using multiple batches of acetone, on different MS platforms, and across multiple labs.

The MS system, LTQ ion trap MS, used throughout the characterization of +98 u artifact did not enable the determination of exact mass of modification. Therefore, a high-resolution MS system, Orbitrap MS, was used for this analysis. Orbitrap MS data demonstrated that the exact mass of the artifact is 98.08 ± 0.01 u (Appendix B, Figure B.1). A second, less intense series of ions, spaced 84 u (exact mass 84.06 ± 0.01 u) above the mass of intact proteins, was observed. This second artifact was also observed from the data acquired using LTQ ion trap MS, as the shoulder on the red peaks in

Figure 4.1c. These series of ions correlate with inclusion of acetone for protein precipitation.

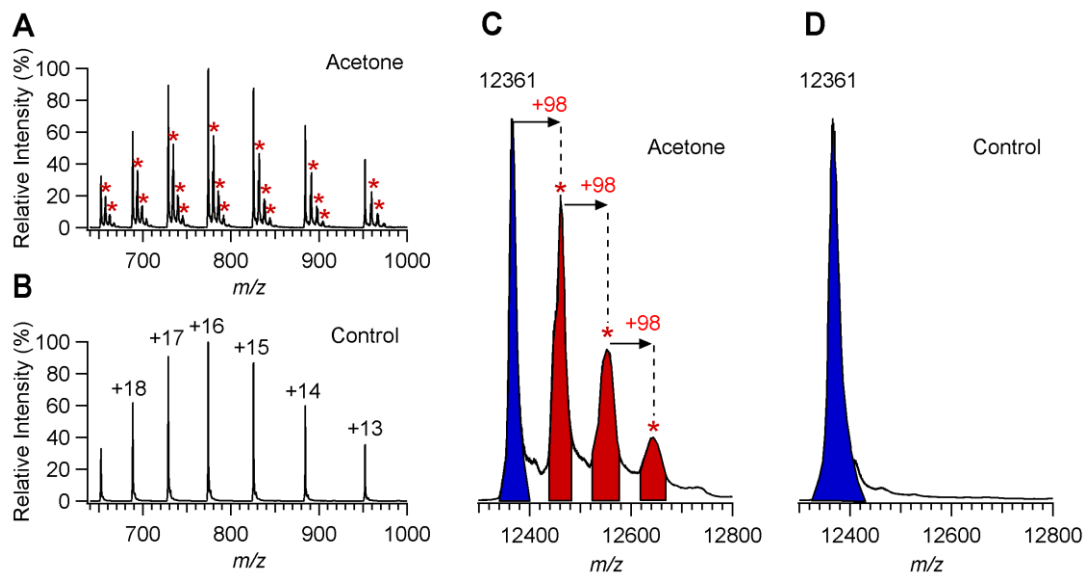


Figure 4.1. (a) Multiply charged and (c) deconvoluted ESI-MS spectra of cytochrome c following precipitation in 80% acetone (-20 °C, overnight). (b) Multiply charged and (d) deconvoluted MS spectra of non-precipitated control

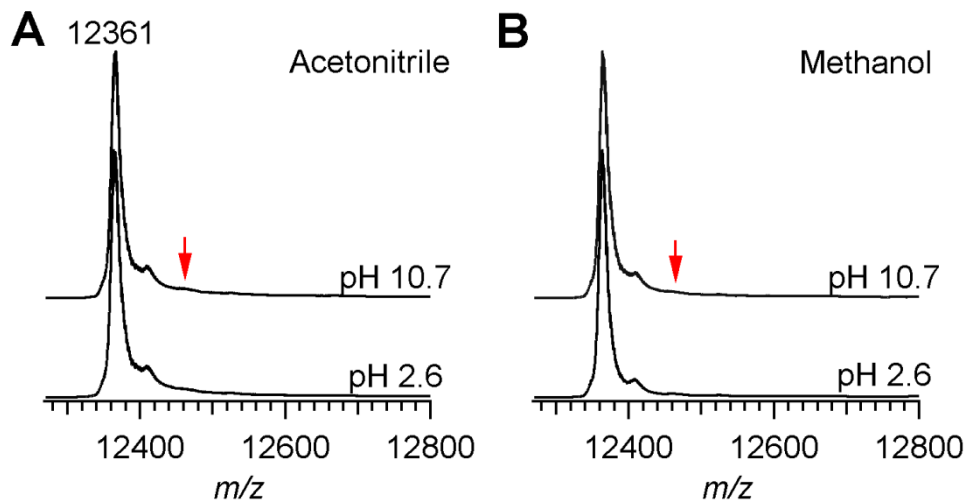


Figure 4.2. Deconvoluted MS spectra of cytochrome c following incubation in (a) 80% acetonitrile, or (b) 80% methanol, for 1 hour at -20 °C

In proteomic analyses, the observation of +98 u artifacts in MS spectra is generally attributed to the formation of sulfate ($M + H_2SO_4$, $\Delta m/z$ 97.967) or phosphate ($M + H_3PO_4$, $\Delta m/z$ 97.977) adducts as contaminants of the system. In localizing the origin of the +98 u artifacts observed in this study, first of all formation of sulfate or phosphate adducts were considered. However, the difference between the expected mass shift of these adducts to that observed on an Orbitrap MS analysis ($\Delta m/z$ 98.08 \pm 0.01) of acetone precipitated sample was over 1000 ppm, which excludes sulfates and phosphates as possible origins of the 98 u adduct. In addition, as initially reported by a previous study, these weakly associated ion complexes can be removed via in-source dissociation (140). Attempts to reduce the intensity of the +98 u ions in this way were unsuccessful; further varying the CID collision energy on a linear ion trap mass spectrometer also failed to eject neutral adducts and ultimately resulted in fragmentation of the full protein.

Analysis of acetone-precipitated proteins by liquid chromatography coupled to ESI-MS also confirms the presence of +98 u artifacts. Furthermore, modification of protein induced by acetone resulted in a chromatographic shift. From Figure 4.3, protein with increased modification is seen to elute later in the chromatogram. Although MS spectra extracted from the later portion of the eluting peak show a distinctly higher degree of modification, distinctly resolved chromatographic peaks could not be obtained. The net effect of this artifact of acetone precipitation is a broadening of the chromatographic profile for the protein. Given the existing difficulties of intact protein fractionation with reversed phase chromatography, further reduction of chromatographic performance is a significant problem. This shift in chromatographic retention also demonstrated that the 98 u artifacts exist in solution, eliminating the possibility of formation of gas phase adducts of neutral or ionic species.

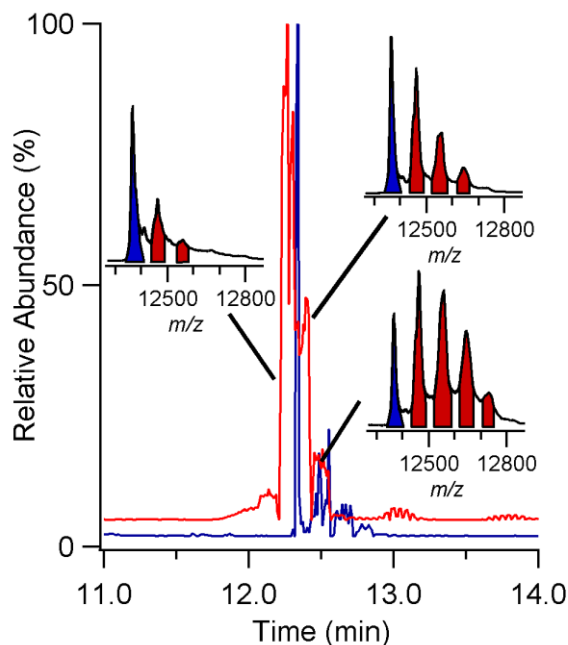


Figure 4.3. LC-MS analysis of cytochrome c (blue trace), or following acetone precipitation (-20 °C, overnight) (red trace)

According to the results thus far, the possibility of a reaction between the protein and acetone was deduced. Therefore, variables, which would likely influence the rate of reaction, including the incubation temperature and time, as well as the pH of the solution were investigated. Figures 4.4 and 4.5 summarize the results of experiments regarding the effect of time and temperature, respectively. For Figure 4.4, cytochrome c was incubated in acetone at reduced temperature (-20 °C) over a range of incubation times, with complete evaporation of the solvent via vacuum concentrator following this incubation period. It should first be noted that the continued exposure of protein to acetone, which occurs during solvent evaporation is not reflected in the reported incubation times. The complete evaporation of acetone lasts approximately 30 min in the vacuum concentrator. In the preparation of a 'time zero' sample, acetone was added to protein solution and then immediately placed in vacuum concentrator for solvent evaporation. Thus, this sample is still exposed to acetone, albeit for a brief period of time. As seen in Figure 4.4a, a noticeable +98 u signal is readily observed at time zero, accounting for 12% of the total protein signal. The level of modification attributed to the drying phase is minor in comparison to that observed following overnight incubation (16 hours) of the protein in acetone. In the latter case, over 90% of the MS signal is attributed to +98 u artifacts. Figure 4.4a reveals a gradual increase in the level

of modification with increased incubation times. Moreover, the degree of protein modification as a function of time was quantified in Figure 4.4b. This was measured through the relative intensity of the unmodified protein signal with respect to the total cytochrome c peak (sum of modified + unmodified peak areas). It was estimated that this reaction does not proceed at constant rate with respect to protein concentration. Modification proceeds more quickly over the first two hours of reaction (calculated pseudo first-order rate constant of $\sim 0.1 \text{ h}^{-1}$). The rate slows considerably, to $\sim 0.01 \text{ h}^{-1}$ (~ 10 fold difference), from 2 to 16 hours. Considering the precipitation process, it is likely that protein aggregation impacts its availability to react. As the incubation time increases, the number of protein molecules that are available for interaction with acetone decreases, as a result of increase in aggregation. Moreover, the change in higher order protein structure on exposure to acetone should be considered. Finally, the observation of multiple modification sites of a given protein (3 distinct artifact peaks are seen in Figure 4.4a) suggests certain sites to be more reactive than others.

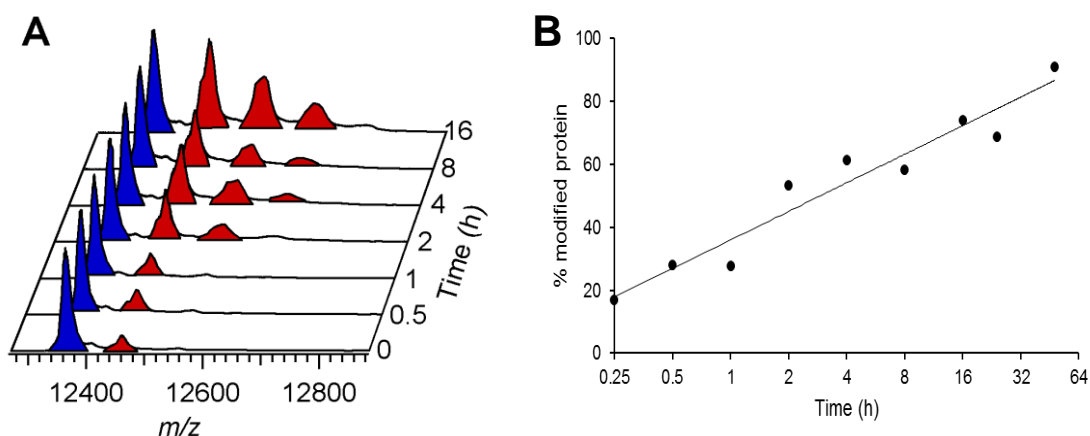


Figure 4.4. (a) Deconvoluted MS spectra of cytochrome c following incubation in 80% acetone ($-20 \text{ }^{\circ}\text{C}$) for defined incubation periods. (b) The degree of protein modification as a function of time

The effect of incubation temperature on protein precipitation in acetone is highlighted in Figure 4.5. In these experiments, incubation of cytochrome c in acetone was performed for a constant 1 hour period, though at varying temperatures. The results reveal that exposure to acetone at higher temperature dramatically increased the level of protein modification. While over 70% of the protein signal is observed in unmodified form following incubated at $-20 \text{ }^{\circ}\text{C}$, the level reduced to 10% unmodified protein once

the temperature is raised to 0 °C. Higher temperatures also caused a significant increase in the intensity for the multiply modified protein (Figure 4.5b), with up to five +98 u artifacts observed in the MS spectrum (incubation at 20 °C). It is feasible that far more reactive sites are involved in the modification; the signals may constitute a distribution of reaction sites, and may also incorporate more than one type of protein chemical moiety. The impact of incubation temperature caused more dramatic changes on the level of modification compared to impact of incubation time. This might be due to the increased rate of reaction between acetone and protein molecules at higher temperatures than -20 °C and highlights the necessity of incubation at -20 °C.

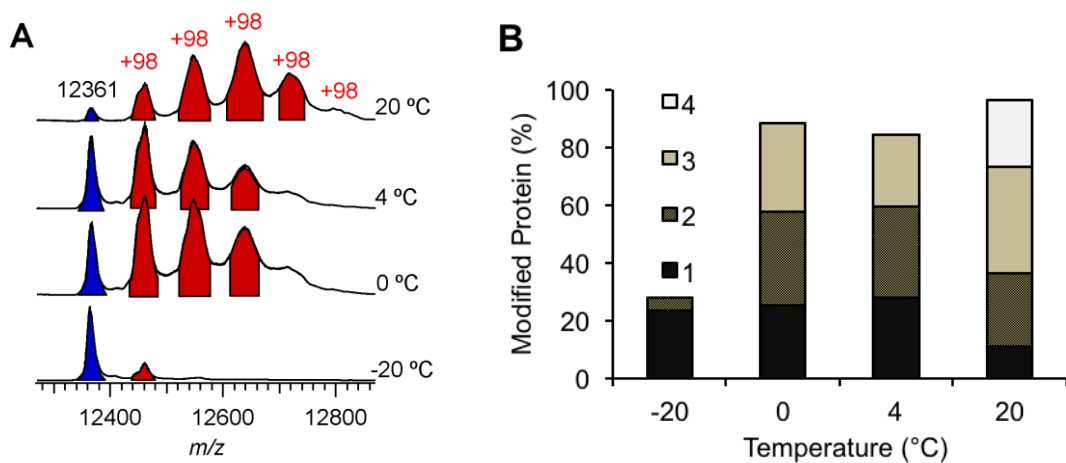


Figure 4.5. (a) Deconvoluted MS spectra of cytochrome c following incubation in 80% acetone for 1 hour at specified temperatures (b) The level of modification is quantified as a function of temperature, noting peak areas for protein with 1 to 4 modifications

Acetone precipitation is routinely employed by a great number of researchers, though conditions for precipitation are not standardized. Typically, samples are precipitated at reduced temperatures (-20 °C to 4 °C), but precipitation at room temperature is also possible. Incubation periods also vary considerably (1 hour or overnight incubations are typical). The results of this study demonstrated that longer incubation at higher temperature lead to a significantly greater amount of modification of protein. In practice, during conventional precipitation the protein is also exposed to acetone throughout the centrifugation step to isolate the protein pellet. The centrifugation step may take place at reduced temperatures, but -20 °C is rarely

employed. Following centrifugation and decanting of the solvent, the presence of residual acetone is unavoidable, and the final drying stage for the evaporation of residual acetone also exposes the protein to acetone, for varying times and temperatures. Some of the researchers allow the protein pellet to stand at room temperature for final drying; others may heat the sample to enhance drying. The differences in the employment of the method surely impact the level of modification induced by acetone. To standardize this variability and employ a controlled method of assessing incubation temperature and time, the results reported here was obtained using an unusual practice, as such complete evaporation of the solvent via vacuum concentrator. To confirm that the +98 u peaks are not simply an artifact of the solvent evaporation stage and to demonstrate the effect of residual acetone in protein pellet, the conventional acetone precipitation was performed, as well (1 hour, -20 °C). In this method, the bulk of the acetone is removed *via* pipetting of the supernatant following centrifugation and the residual acetone was evaporated in a fume hood at room temperature. Figure 4.6 depicts the comparison of these two evaporation approaches. Compared to the complete solvent evaporation protocol, conventional precipitation resulted in a higher degree of protein modification, in terms of intensity and number of +98 u artifact peaks. This was to be expected, since the higher level of modification can be attributed to exposure of protein to residual acetone at a high temperature (20 °C) during the final drying of the solvent.

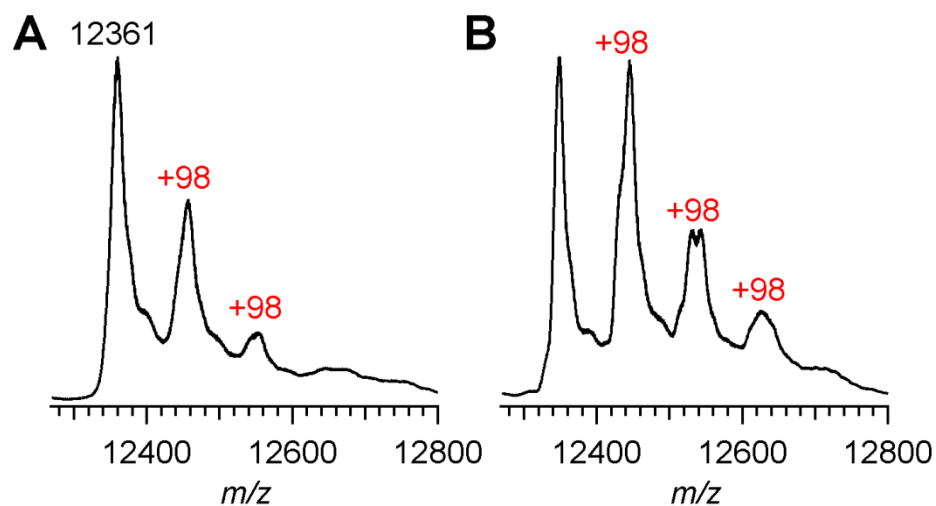


Figure 4.6. Deconvoluted MS spectra of cytochrome c following overnight incubation in 80% acetone (-20 °C) and subsequent evaporation of the solvent via (a) vacuum concentrator or (b) conventional protocol

Based on these results, while it could be recommended to maintain reduced temperatures and brief incubation periods for decreased level of modification, it is evident that conventional acetone precipitation still presents the potential for a significant level of protein modification, as observed through +98 u satellite peaks on cytochrome c. Without further reducing the level of protein modification, acetone would be viewed as an undesirable solvent choice for proteomics workflows.

The pH of a solution is an important variable for proteins. For the investigation of the impact of pH on the formation of +98 u artifact, cytochrome c was prepared in acidic, neutral or basic solutions. These solutions and their measured pH values in parenthesis are as follows: 0.1% TFA (pH 1.9), 0.1% formic acid (pH 2.6), 0.1% acetic acid (pH 3.2), water (pH 6.3), 50 mM Tris-HCl buffer (pH 8.0), 50 mM ammonium bicarbonate buffer (pH 8.3), 10 mM NH₄OH (pH 10.7) and 2 mM NH₄OH (pH 12.3). As shown in Figure 4.7, the pH of the initial solution has a controlling influence on the reaction. Formation of +98 u artifacts is suppressed when acetone is added to acidified cytochrome c solutions near or below pH 3. This is true regardless of the acid used. By contrast, incubation of protein in acetone at neutral or basic pH solutions resulted in an increased level of protein modification. A typical buffer solution (Tris-HCl pH 8 or ammonium bicarbonate pH 8.3) shows minimal difference in the level of modification compared to a purely aqueous solution. However, adjusting the solution pH to higher values with ammonium hydroxide further increased the level of modification. This observation might result from the adaptation of different conformational structures of protein in solutions at acidic, neutral or basic pH. As a result of such changes in the conformation of protein, the exposure of more reactive sites to interact with acetone is possible. Based on these findings, it can be concluded that protein precipitation is best conducted on solutions adjusted to acidic pH prior to addition of the acetone. Although not tested, a commonly employed trichloroacetic acid (TCA)/acetone precipitation method would meet these conditions due to the fact that the modification is not dependent on the type of acid used. This method may minimize the risk of unexpected protein modification and allow the use of acetone in a proteomics workflow.

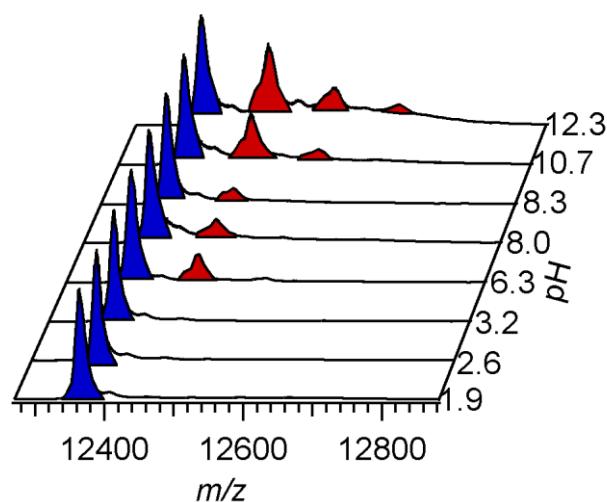


Figure 4.7. Deconvoluted MS spectra of cytochrome c prepared in solutions of varying pH and precipitated in 80% acetone (-20 °C, 1 hour)

In further investigation of the effect of solution pH, it was attempted to reverse the modification by incubating an acetone-precipitated protein in acidic solution following removal of the solvent. Moreover, it should be noted that all MS spectra reported in this study incorporated an acidic solution (0.1% formic acid) to resolubilize the protein pellet prior to LC-ESI-MS analysis with further use of this solvent as a mobile phase for LC. The protein pellet was also incubated in 80% concentrated formic acid, though in all cases the +98 u artifacts remain present. Therefore, it is clear that the final reaction product is stable in acidic environments, despite the controlling influence of solution pH on the initial reaction. Once the acetone induced modification is formed on the protein, it cannot be reversed using acidic solutions.

The formation of +98 u artifacts was primarily observed in a study of intact protein MS analysis of cytochrome c. For ease of comparison, the data presented thus far constitute the modification of cytochrome c. To further investigate the extent of modification, different standard proteins were used. Figure 4.8 displays MS spectra following acetone precipitation of standard proteins including myoglobin, ubiquitin, hemoglobin, as well as apo-cytochrome c (in the absence of a heme group). Artifact peaks were observed on other proteins, and to highly varying degrees depending on the protein. These proteins were precipitated in acidic, neutral and basic solutions to confirm the impact of solution pH on the reaction with acetone. According to the results, it was observed that the highest degree of modification occurred on heme-containing proteins (cytochrome c, myoglobin, hemoglobin). Therefore it was suggested

that the heme group may directly be involved in the formation of modification either catalyzing the reaction, or as the chemical moiety involved in the modification. To test this, a sample of apo-cytochrome c was prepared, thereby removing the heme group from the system. As seen in Figure 4.8, a significant level of modification was also present on apo-cytochrome c following incubation in acetone, indicating that the amino acid side chains or peptide backbone of the protein is directly modified in acetone. The data suggests that the heme group does not appear to be required for the formation of modification, though it may still involve in modification process by accelerating the reaction.

According to the data represented thus far, it was observed that incubation of proteins in acetone results in the observation of +98 u artifact peaks in MS spectra and this modification is stable in both acidic and basic environments. The next step is to localize the reaction site(s) on the protein. To that end, bottom up LC-MS/MS analyses of tryptic digests of BSA, as a test protein, and *E. coli* total protein extract, as a complex protein mixture, were performed following incubation in acetone. In processing the LC-MS/MS data, +98.0732 u modification was included in database search parameters, without restriction to any specific amino acid. The results from both tryptic digests showed that the modification was localized to amine-containing amino acids (K, H and R). While bottom up LC-MS/MS analysis of *E. coli* proteome identified nearly 1500 unique peptides, only 17 were observed as being tentatively modification by acetone (Appendix B, Table B.1). Of these, 7 were assigned to histidine while 10 were assigned to lysine residues. This is a rather low rate of modification, and may be a consequence of the level of modification being protein or sequence specific, or possibly that the modification is unstable during the digestion process and MS/MS analysis. Nonetheless, sample MS/MS spectra is provided in Figure 4.9, showing the localization of the modification to a histidine residue of the BSA peptide hLVDEPQNLIK, as well as a lysine residue on the tryptic peptide SLDDFLIKQ, isolated from the *E.coli* DNA-binding protein H-NS (Uniprot accession number P0ACF8). During the preparation of these samples, acetone was fully removed via vacuum concentrator prior to tryptic digestion, thus the N-terminus of the resulting peptides would not be available for reaction. These results do not preclude the possibility of modification at other protein residues (*e.g.* cysteine).

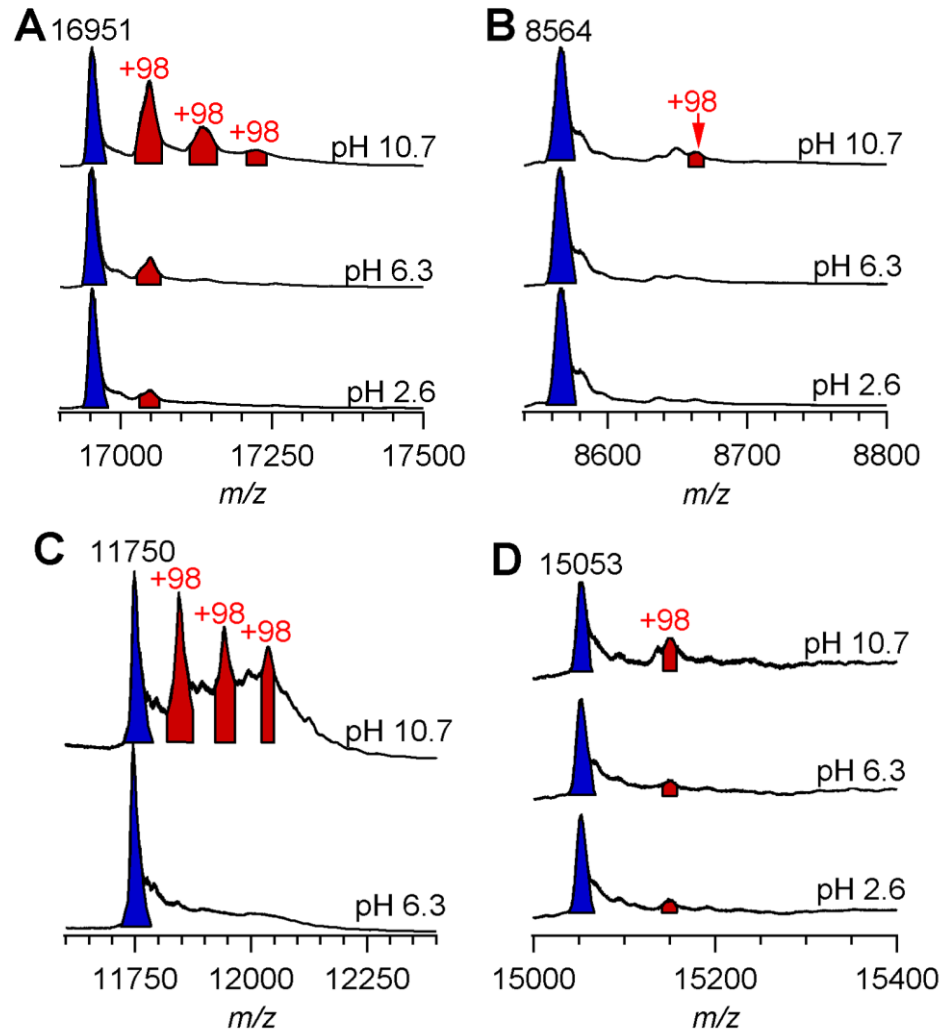


Figure 4.8. Deconvoluted MS spectra of acetone-precipitated proteins prepared in acidic, neutral and basic solutions. (a) Myoglobin, (b) ubiquitin, (c) apocytochrome c, (d) hemoglobin (alpha chain shown)

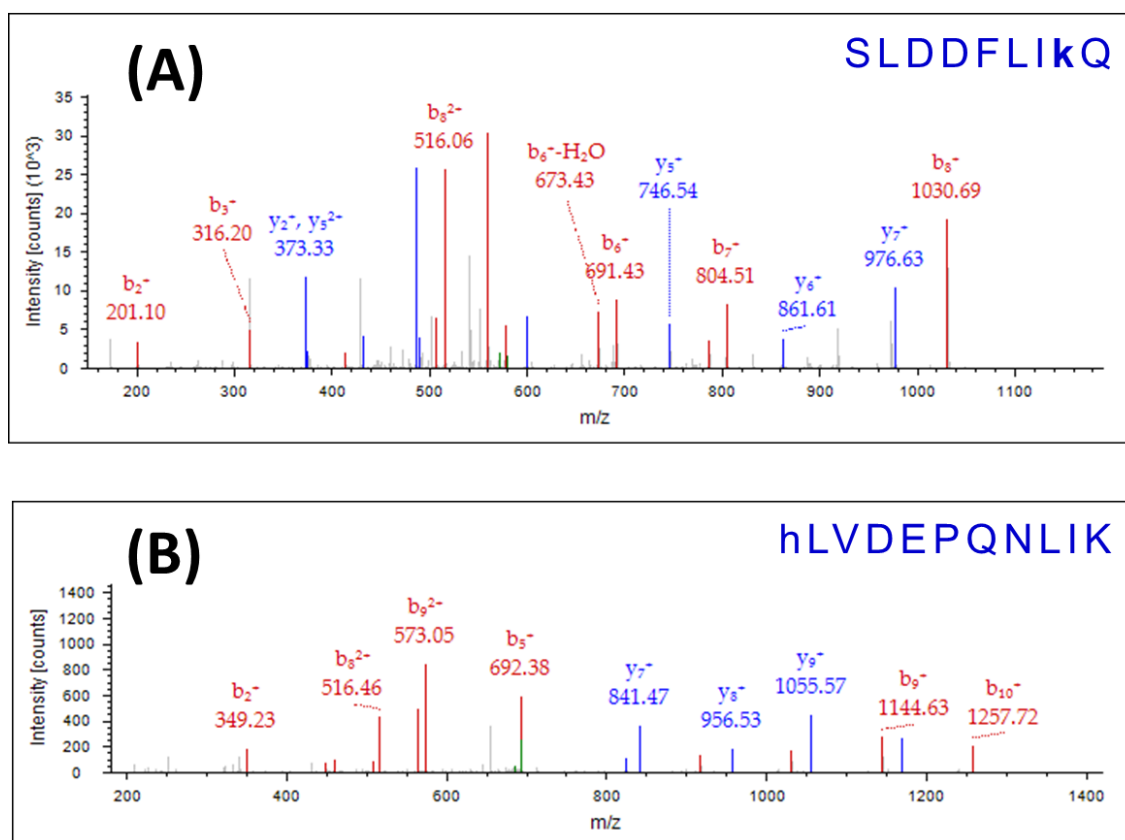


Figure 4.9. MS/MS analysis of tryptic peptides from (a) DNA-binding protein H-NS from *E. coli*, or (b) BSA

Considering the chemical origin of the +98 u artifact, the mass shift implies the involvement of two molecules of acetone, coupled with the loss of water ($2 \times 58 - 18 = 98$ u). Orbitrap MS analysis enabled the determination of exact mass as 98.08 ± 0.01 u. The exact mass suggests addition of the chemical moiety $C_6H_{10}O$, which has 98.07 u exact mass, (error to observed mass = 69 ppm) and is consistent with the proposed reaction. The only other potential chemical adduct within a 100 ppm tolerance would be $C_5H_{10}N_2$, though such a modification cannot be explained through a reaction with acetone. While the addition may occur through direct reaction with acetone in two discrete steps, an alternative hypothesis would be a single step addition onto the protein. Self-aldol condensation of acetone to form diacetone alcohol (**I**), 116 u, can be catalyzed under acidic or basic conditions (Box ‘A’ of Figure 4.10), which subsequently dehydrates to form mesityl oxide (**II**), 98 u. This reaction is also known to be catalyzed over ion exchange resin (141). Given the acid/ base properties and localized charges of a protein, the inclusion of a protein may accelerate the aldol addition reaction of acetone. Moreover, the presence of Fe^{3+} has also been shown to favor the formation of

diacetone alcohol (142). This is consistent with the high level of modification that is observed on heme-containing proteins in this study. To test the hypothesis of a single step addition, cytochrome c was incubated in solutions of 0.1% diacetone alcohol or 0.1% mesityl oxide, prepared in acetonitrile. As seen in Figure 4.11, reactions with each of these compounds resulted in the formation of a +98 u artifact. Together with +98 u artifact, +112 u artifact was observed in sample treated with 0.1% mesityl oxide. However, the origin of this artifact is not known for this time. Proposed reactions between diacetone alcohol or mesityl oxide and protein are summarized in Figure 4.10. Noting the stability of the product in acidic environments, the conjugate addition of mesityl oxide (Box 'B' of Figure 4.10) is viewed as a more favorable product. Formation of the imine through condensation with diacetone alcohol (Box 'C' of Figure 4.10) would be a reversible process, and therefore requires removal of water to stabilize the product. This may be possible given the higher order structure of the protein. One also cannot preclude the dehydration of diacetone to mesityl oxide prior to addition onto the protein. Reactions at other nucleophilic centers of the protein (*eg* cysteine) may also be possible, though to date, such products have yet to be directly observed by MS/MS.

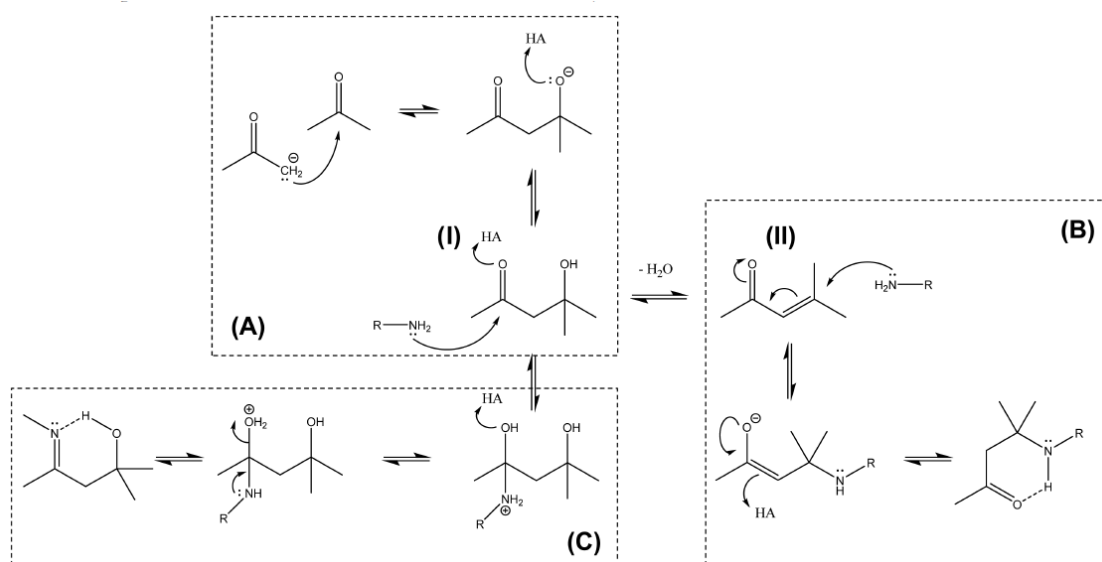


Figure 4.10. Proposed reactions of diacetone alcohol (I) or mesityl oxide (II) with protein

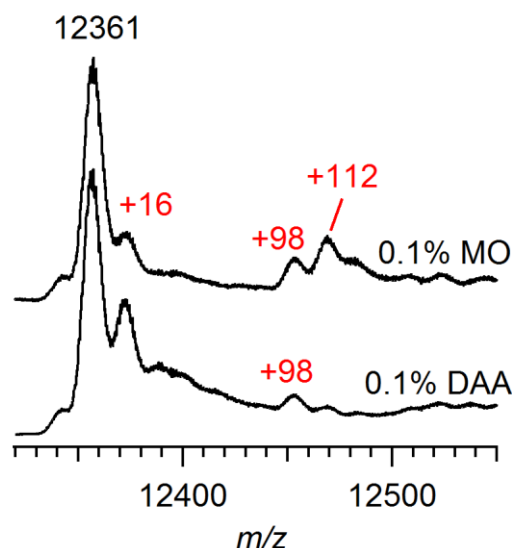


Figure 4.11. Deconvoluted MS spectra of cytochrome c incubated in 80% acetonitrile containing 0.1% mesityl oxide (top) or 0.1% diacetone alcohol (bottom)

4.4. Conclusion

Protein precipitation in organic solvents including acetone has long been used to concentrate and purify proteins. In lowering the dielectric constant of the solution, the addition of organic solvent creates an environment of low solubility, resulting in protein aggregation and precipitation. The method is not only simple and cost effective, but also exceeds the performance of other purification strategies in terms of protein recovery and purity. Despite the benefits of acetone precipitation, this study revealed that the chemistry of acetone precipitation remains poorly understood. Upon incubation in acetone, multiple +98 u artifacts were observed for intact proteins when analyzed by mass spectrometry. Up to date, this study is the first one reporting an acetone induced +98 u modification on intact proteins. Therefore this study conducted the investigation of this novel modification.

The results showed that the formation of modification is dependent on pH of the protein solution, incubation time and temperature. The level of modification increased with longer incubation times, at higher temperatures, and in neutral to basic pH solutions. The modification is suppressed when proteins are incubated with acetone in acidic solutions as well as at incubation temperatures as low as -20 °C and brief incubation periods. The +98 u artifacts were observed on all standard proteins tested in

this study, albeit at varying levels. The modification site was localized to histidine, lysine and arginine residues, though other reaction sites on the protein may exist. The product is speculated to arise following formation of mesityl oxide via aldol condensation of acetone. The exact mass of the artifact was determined as +98.0732 u and the most probable structure was deduced as C₆H₁₀O.

As a result, the researchers who employ acetone precipitation of proteins prior to mass spectrometry analysis should consider the parameters, which effect the formation of modification, reported in this study. In addition, a +98.0732 u dynamic modification at lysine, arginine and histidine residues should be included during MS/MS spectral database searching. The elimination or at least being aware of this and other such modifications would improve the proper identification of proteins by mass spectrometry.

CHAPTER 5

OBSERVATION OF THE SIDE CHAIN METHYLATION OF GLUTAMIC ACID AND ASPARTIC ACID CONTAINING PEPTIDES BY MASS SPECTROMETRY

5.1. Introduction

Mass spectrometry is an unparalleled tool for the identification and characterization of post-translational modifications (PTMs) observed on proteins. As known, PTMs are formed on multiple sites of a protein and result in the generation of a great number of variants of a protein with possible diverse structure and functions. Some of the common PTMs include phosphorylation, methylation, glycosylation, acetylation, ubiquitination, etc. and can readily be identified by tandem mass spectrometry experiments (143, 144). The identification of known and novel PTMs would make important contributions in understanding the biological pathways that induce modification. As well as other PTMs, methylation is a prominent protein modification in eukaryotic cells and catalyzed by a large group of enzymes called methyltransferases (145). Methylation of lysine (K) and arginine (R) residues on proteins has already been documented in literature and have been implicated in several biological processes such as, protein-protein interactions, signal transduction, chromatin remodeling, mRNA splicing and translocation (145, 146). In addition to lysine and arginine, several other amino acid residues on proteins were also found to be methylated. These residues include, aspartic acid (D), glutamic acid (E), histidine (H), asparagine (N), glutamine (Q) and cysteine (C) (145, 147, 148). A very recent report by Wang et al. sorted these amino acids according to the preference for methylation as follows $K \gg R > D > N \sim Q \sim H > E > C$ in *S. cerevisiae* (149). Concerning aspartate methylation, protein-L-isoaspartate-O-methyltransferase (PIMT) is the enzyme that specifically catalyzes the methylation of aspartate residue. Aspartic acid methylation has been implicated in protein repair mechanism (150), though neither glutamic acid methylation nor specific glutamic acid methyltransferases have been reported in

eukaryotes by previous studies (148). Given the importance of aspartic and glutamic acid in protein folding and function, methylation of these residues is likely to have a significant impact on protein structure, interactions and function, due to the neutralization of negative charge on aspartic and glutamic acid side chains by addition of a methyl group.

Methylation of certain amino acid residues is not only observed as a PTM, but also as artifacts of sample handling in proteomic workflows. Haebel et al. reported the first study on the artificial methylation in which glutamic acid side chain and C-terminus of peptide was methylated during staining procedure of polyacrylamide gels (151). The staining solution included trichloroacetic acid (TCA) and methanol (MeOH) and it was shown that γ -carboxyl group of glutamic acid was converted to its corresponding methyl ester form via acid-catalyzed esterification reaction. In addition, the authors reported a slower rate of aspartic acid methylation compared to the glutamic acid methylation. Furthermore, they also demonstrated ethylation of glutamic acid when ethanol (EtOH) was used instead of MeOH as a solvent in staining solution. Likewise, in another study, it was stated that aspartic acid and glutamic acid side chains can potentially undergo methyl esterification when MeOH was used as a solvent in protein extraction protocols and gel-staining/destaining buffers of proteomic studies (148). A short communication by Sumpton and Bienvenut has demonstrated that the use of colloidal Commassie brilliant blue dye dissolved in trace amount of phosphoric acid containing MeOH can induce methylation of side chains of acidic residues (E and D) during gel staining procedures (152). Additionally, it has been shown that MALDI matrix and sample preparation procedures may also cause the methylation of acidic residues (153). Unexpected chemical modifications implicated in proteomic workflows not only increase the complexity of sample, but also impact the success of identification of proteins. Moreover, if one of these modifications, induced by sample handling, result in the introduction of the same chemical moiety to that observed *in vivo*, that would hamper either the qualitative or quantitative data analysis in proteomic experiments. It would be impossible to determine the origin of methylation observed on a protein, if one is searching for methylation as a protein PTM and used acidified methanol solution in the sample preparation procedure.

The reaction chemistry behind methylation is a condensation reaction of carboxylic acids with an excess alcohol produce an ester. The reaction proceeds in the presence of a strong acid, therefore it is also known as an acid-catalyzed esterification

(Fischer esterification). Previous studies showed that various carboxylic acids were converted to their corresponding methyl ester derivative under methanol/trimethylchlorosilane solvent system (154–156). Similarly, glutamic and aspartic residues can be easily converted to their methyl ester form when methanol has been used as a solvent in acidic media (151, 153, 157).

As a solvent system in ESI-MS experiments, a mixture of an equal proportion of polar/volatile organic solvent (e.g. methanol, acetonitrile, or isopropanol) and ultrapure water has been used especially in direct infusions. This solution is acidified through addition of formic acid or acetic acid, within a range of 0.1-1%, for the positive mode MS analysis. The organic solvents are generally added to the liquid mixture to decrease the surface tension, which facilitates formation of gas-phase ions through ESI process. Among the solvents, methanol is one of the most extensively used organic solvent due to its high polarity, volatility, and compatibility. Acidification of organic solvent/water mixture further promotes ionization of analytes.

Following ESI-MS analysis of glutamic and aspartic acid containing model peptides in our laboratory, the excess peptide solutions were stored in the fridge in ESI solution for further experiments. Later on, when these peptides were analyzed again, the resulting peptide spectra contained additional peaks corresponding to $[M + H + 14]^+$ ions. MS/MS experiments of these ions showed that the side chains of glutamic and aspartic acid residues were methylated. The formation of an unintentional methylation of peptides was due to the long-term storage of these peptides in ESI solution, that is an acidified methanol solution. Therefore, in this study, a comprehensive investigation of the extent methylation of glutamic acid and aspartic acid side chains as well as C-terminus carboxyl group was performed. To that end, the level of methylation on model peptides was detected for different incubation temperatures and times, added acid types and its percentages. In addition, two adjacent E or D residue containing model peptides were used in order to probe the extent of multiple methylation reaction. Lastly, the effect of organic solvent on the esterification reaction was studied.

5.2. Experimental Methods

5.2.1. Preparation of Peptide Samples

Model peptides used in this study were purchased from GL Biochem Ltd. (Shanghai, China) as lyophilized powder form. Stock peptide solutions were prepared in a 50:50 (v/v) solution of methanol and MilliQ purified water.

In investigation the effect of temperature and time on methylation reaction, 100 pmol/ μ L peptide solutions were incubated in MeOH/dH₂O/FA (50:50:1, v/v/v) mixture for different temperatures (-20 °C, 4 °C, 22 °C, 37 °C, or 50 °C) and over a range of incubation times (1, 3, 7, 14 and 30 days). The effect acid type present in incubation solution was tested by incubating 100 pmol/ μ L peptide solution at 37 °C over a range of incubation times (1, 3, 7, 14 and 30 days) in MeOH/dH₂O (50:50, v/v) mixture containing either 0.1 or 1% formic acid (FA) or acetic acid (AA). The effect of organic solvent was investigated by incubating 100 pmol/ μ L peptide solutions in 50:50:1 (v/v/v) mixture containing either ethanol (EtOH/dH₂O/FA) or isopropanol (*i*PrOH/dH₂O/FA), at 37 °C over a range of incubation times (1, 3, 7, 14 and 30 days). As control samples, peptides were freshly prepared in solutions specific for each experiment and analyzed immediately.

5.2.2. Mass Spectrometry Analysis of Peptides

Mass spectrometry experiments were conducted on a LTQ XL linear ion-trap mass spectrometer (Thermo Finnigan, San Jose, CA, USA) equipped with an ESI source. Each sample was introduced into the ESI source of the mass spectrometer via a syringe pump at a flow rate of 5 μ L/min. The default peptide working solutions were prepared in MeOH/dH₂O/FA (50:50:1, v/v/v) mixture. The system was operated in the positive mode scanning over the mass range m/z 150-1000. The spray voltage was set to + 5.0 kV and the heated capillary temperature was 300 °C. All ion optics was optimized in order to get maximum ion intensity with an auto-tune function of the ion-trap instrument. Nitrogen was used as a sheath, an auxiliary, and a sweep gas with a flow rate of 10, 1, 1 (in arbitrary unit), respectively, where helium was used as a collision gas

for CID. For the tandem mass spectrometry (MS/MS) experiments, the mass isolation window (m/z) was set at between 1.0 and 1.4 with an activation time of 30 ms and an activation (q) of 0.250. The normalized collision energy was set to 20-26% for the fragmentation of the selected precursor ion *via* MS/MS. Data acquisition was performed with Xcalibur (ver. 2.0) software data system and MS spectra were plotted using Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

5.2.3. HPLC Analysis

For HPLC experiments, 5 μg of peptide sample was injected onto a 10 cm x 2.1 mm Kromasil C₁₈ column coupled to a Dionex Ultimate 3000 Nano/Cap HPLC system (Dionex, Camberley, UK). Solvent A was 0.1% trifluoroacetic acid (TFA) in water and B was 0.1% TFA in acetonitrile. The column temperature was set to 25 °C. Peptides were eluted with a thirty minute gradient program as follows: 0 min, 15% B; 25 min, 15-40% B and 25.01-30 min, 80% B. The UV detection wavelength was set to 220 nm. The solvent flow rate was 200 $\mu\text{L}/\text{min}$ and outlet of detector was directed to the ESI source of LTQ linear ion trap mass spectrometer, maintaining a flow rate of 10 $\mu\text{L}/\text{min}$ via a post-detector split.

5.3. Results and Discussion

In order to investigate the effect of different incubation temperatures and storage times on methylation of aspartic acid (D) and glutamic acid (E) residues, model pentapeptides were used in this part of the study and these peptides were designed to include C-terminus amide group instead of carboxyl group, eliminating the possibility of C-terminus carboxylic acid methylation. In addition to E and D residues, these peptides (EGGFL-NH₂ and DGGFL-NH₂) contained glycine (G), phenylalanine (F) and leucine (L) amino acids, which did not contain carboxylic acid groups in their side chains and are inert to methylation. Each peptide sample were prepared in MeOH/dH₂O/FA (50:50:1, v/v/v) solution at a concentration of 100 pmol/ μL and aliquots of each sample were allowed to incubate at different temperatures such as -20 °C, 4 °C, 22 °C, 37 °C or 50 °C. MS analysis of each sample incubated at different temperatures were performed over a range of incubation times (1, 3, 7, 14 and 30 days)

allowing to investigate the effect of storage time, as well. The recorded mass spectra of EGGFL-NH₂ and DGGFL-NH₂ for each incubation period are given in Appendix Figure C.1-C.4. For ease of comparison, only mass spectra of thirty-day incubation at different temperatures are reflected in Figure 5.1.

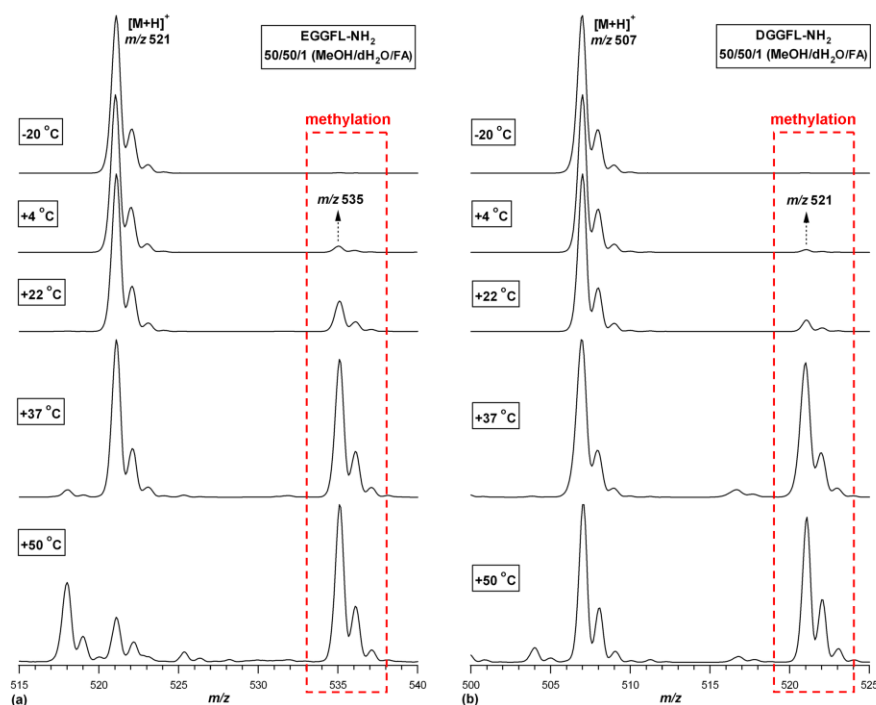


Figure 5.1. ESI-MS spectra of (a) EGGFL-NH₂ and (b) DGGFL-NH₂ following thirty days of incubation in MeOH/dH₂O/FA (50:50:1, v/v/v) at -20 °C, 4 °C, 22 °C, 37 °C and 50 °C

The protonated molecular ion $[M + H]^+$ of EGGFL-NH₂ and DGGFL-NH₂ are m/z 521 and 507, respectively. As a result of incubation of these peptides in MeOH/dH₂O/FA (50:50:1, v/v/v) for thirty days at different temperatures, one can easily note the increase in the intensity of peaks m/z 535 and 521 (highlighted in red box) in EGGFL-NH₂ and DGGFL-NH₂ samples, respectively in Figure 5.1. These peaks are separated by +14 Da from $[M + H]^+$ ions and this mass shift is attributed to the methylation of E and D residues. Previous studies (148, 157) have also reported the methylation of E side chain during storage in ESI solution of MeOH/dH₂O/FA (50:50:1, v/v/v). The product of methylation reaction increased in the abundance in both samples as the peptides were incubated at 4 °C, 22 °C, 37 °C and 50 °C, though none of the samples resulted in the observation of methylated product at -20 °C even for thirty days

of incubation, suggesting the stability of storage solution at -20 °C. In addition, it is apparent that the relative intensity of methylated product in EGGFL-NH₂ sample is higher than that of the DGGFL-NH₂ sample for different incubation temperatures (except for 37 °C). Moreover, at 50 °C, the relative intensity of EGGFL-NH₂ methylated product is higher than the intensity of the [M + H]⁺ ion in EGGFL-NH₂ sample, whereas the intensity of the methylated product has almost the same intensity with the [M + H]⁺ ion in DGGFL-NH₂ sample. Overall, these data demonstrated that methylation of E and D residues within a peptide sequence is a temperature dependent process and the rate of reaction increases as the temperature increases. As figures in Appendix Figure C.1-C.4 illustrate, longer incubation times resulted in the accumulation of more methylated product in each sample at temperatures higher than -20 °C. In addition, E methylation is a more rapid reaction compared to D methylation for all tested temperatures other than +37 °C and when C-terminus amidated model pentapeptides were used. Haebel et al. (151) reported the preference of E methylation over D methylation. However, the results obtained herein differ from their results in terms of D methylation.

As a result of methylation of EGGFL-NH₂, it was suggested that E_{OMe}GGFL-NH₂ peptide product was formed. The MS analysis confirmed the formation of this product by +14 Da mass shift. To further confirm the site of methylation, the ion *m/z* 535 was isolated and allowed to dissociate under low-energy CID conditions. The resulting MS/MS spectrum was compared with the MS/MS spectrum of [M + H]⁺ ion of commercial side-chain methylated pentapeptide, E_{OMe}GGFL-NH₂ and the spectra are reflected in Figure 5.2. From this figure, it is evident that the fragmentation behavior of EGGFL-NH₂ methylation product and [M + H]⁺ ion of commercial E_{OMe}GGFL-NH₂ pentapeptide are virtually identical by having the same fragment ions along with the same relative intensities. The one-to-one similarity of these mass spectra provided a strong evidence for side chain methylation of E during storage in its ESI solution of MeOH/dH₂O/FA (50:50:1, v/v/v). The loss of 32 Da (mass of methanol) from [M + H]⁺, *b*, and/or *a* ions supports the formation of methylated peptide product. In Figure 5.2, water loss (-18 Da) from *b* ions and an ammonia (-17 Da) loss from *a* ions are denoted as *b*^o and *a*^{*}, respectively.

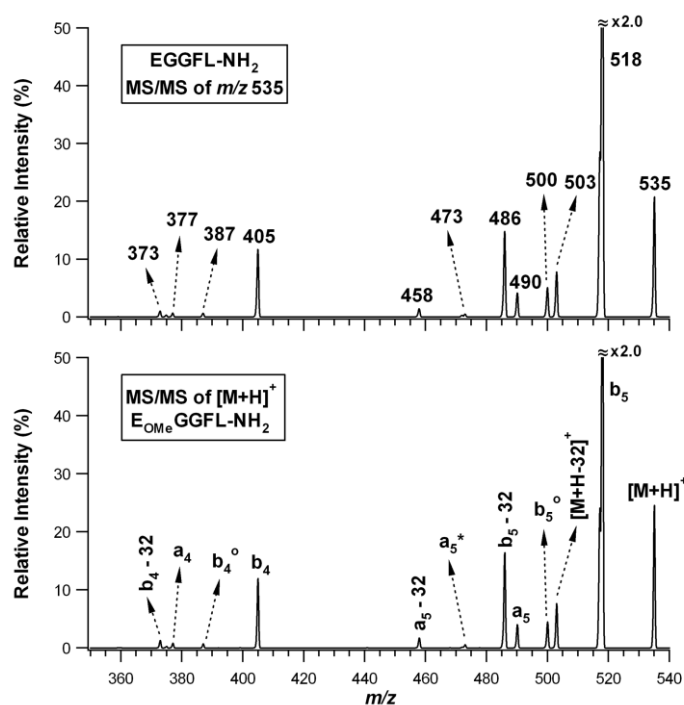


Figure 5.2. Comparison of MS/MS spectrum of m/z 535 ion corresponding to methylated product of EGGFL-NH₂ (top) and MS/MS spectrum of $[M + H]^+$ ion of commercial E_{OMe}GGFL-NH₂ (bottom)

Methylation of E and D residues not only resulted in the observation of +14 Da mass shift in MS analysis, but also a shift in the retention time of the peptide in HPLC analysis. Figure 5.3 highlights the change in the retention time of methylated product E_{OMe}GGFL-NH₂ (obtained by incubation at 37 °C for seven days), compared to freshly prepared EGGFL-NH₂ peptide. The freshly prepared sample of EGGFL-NH₂ was eluted from C18 reversed phase column at a retention time of 7.5 min as a single peak. However, methylated peptide was eluted from the same column under the same chromatographic conditions at a retention time of 9.5 min, confirmed by MS analysis of the peak at 9.5 min in chromatogram. The change in the chromatographic properties of the methylated product was expected due to the introduction of a methyl group to the negatively charged side chain of E residue, thereby increasing nonpolar character of the peptide. In addition to the methylated product peak, another peak eluting at 8.5 min was observed in this chromatogram. MS analysis of this peak resulted in the observation of an ion m/z 503. A literature survey demonstrated the conversion of N-terminus E to pyroglutamic acid (pE) through loss of water from its γ -carboxyl group; especially when its acidic solution was incubated at elevated temperatures (158–160). Therefore,

the ion m/z 503, with -18 Da mass shift and 8.5 min retention time, is attributed to the product of N-terminus pE formation.

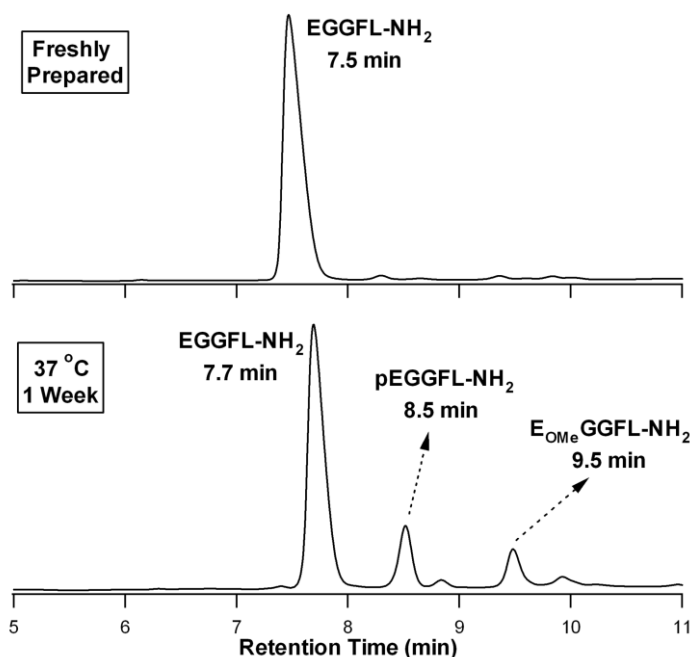


Figure 5.3. LC chromatograms of freshly prepared EGGFL-NH₂ peptide solution (top) and EGGFL-NH₂ peptide solution incubated in MeOH/dH₂O/FA (50:50:1, v/v/v) at 37 °C for a week (bottom)

The results given thus far demonstrated the methylation of acidic amino acid residues within a given peptide sequence. For investigation the methylation of C-terminus carboxyl group of peptides, EGGFL-OH and DGGFL-OH pentapeptides, in which the C-terminus of peptides has free acid functionality, were analyzed. In these peptides, two possible methylation sites exist: the side chains of E and D residues and the C-terminus of peptide. The solution of these peptides were prepared freshly in MeOH/dH₂O/FA (50:50:1, v/v/v) mixture and aliquots of peptides were incubated at 37 °C for 1, 3, 7, 14, and 30 days. For the best representation of the methylation reaction, 37 °C was selected as a standard temperature. Following incubation, MS analyses of the peptides were performed and the results are depicted in Figure 5.4. In this figure, mono- and di-methylated peptide products are noticed by successively spaced peaks with 14 Da mass difference (highlighted in red box). These peaks, which correspond to methylated products, are not present in freshly prepared samples. The peak m/z 536 in EGGFL-OH sample and m/z 552 in DGGFL-OH sample are mono-methylation

products of the peptides, either on the side chain of E/D ($E/D_{OMe}GGFL-OH$) or at the C-terminus of peptides ($E/DGGFL-OCH_3$). Besides, di-methylated peptide products, albeit with a lower relative intensity compared to mono-methylation product, are observed m/z 550 and 536 in EGGFL-OH and DGGFL-OH samples, respectively.

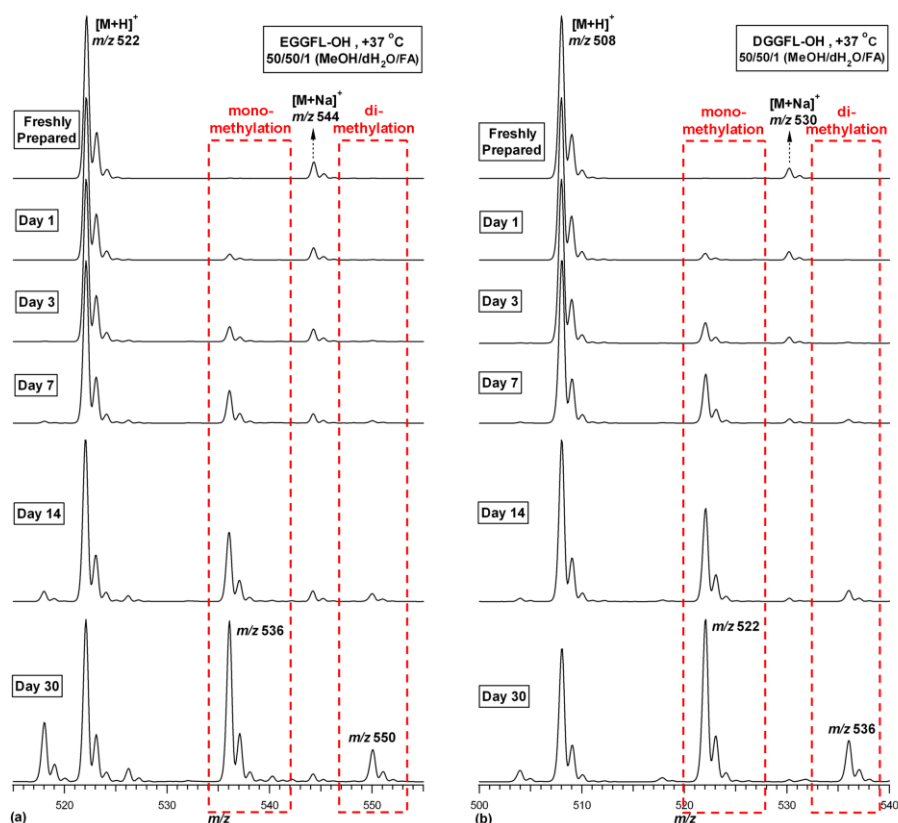


Figure 5.4. ESI-MS spectra of (a) EGGFL-OH and (b) DGGFL-OH recorded following incubation in MeOH/dH₂O/FA (50:50:1, v/v/v) at 37 °C over a range of incubation times

The full scan MS of peptides, which were given in Figure 5.4, did not provide information regarding the site of methylation. To confirm the site of methylation, mono-methylated product (m/z 536) in EGGFL-OH sample, which was incubated at 37 °C for 30 days, was selected and allowed to fragment via CID. The resulting MS/MS spectrum is shown in upper panel of Figure 5.5. The fragment ions originating from $E_{OMe}GGFL-OH$ was highlighted in blue color, whereas EGGFL-OCH₃ fragment ions were highlighted in red. The fragment ions in this spectrum reveal that two mono-methylated peptide products have been formed at the same time during storage in MeOH/dH₂O/FA (50:50:1, v/v/v) solution. Moreover, di-methylated product (m/z 550) in EGGFL-OH

sample, which was incubated at 37 °C for 30 days, was selected for fragmentation and the resulting MS/MS spectrum (lower panel, Figure 5.5) confirmed the di-methylated product (E_{OMe}GGFL-OCH₃) through assigning its sequence specific ions.

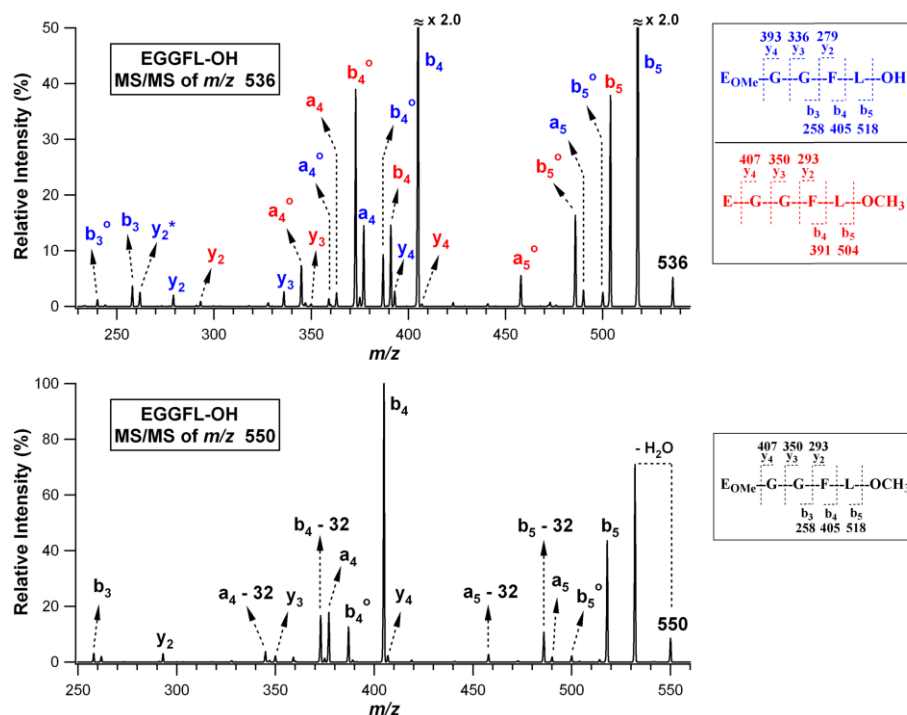


Figure 5.5. MS/MS spectra of mono-methylated (top) and di-methylated products of EGGFL-OH following incubation in MeOH/dH₂O/FA (50:50:1, v/v/v) at 37 °C for 30 days

In order to investigate the extent of methylation of E and D residues within different sequences, AAAXAAA-NH₂ and AAAXAAA-OH (where X is E or D) model heptapeptides were used. Peptide solutions were freshly prepared in MeOH/dH₂O/FA (50:50:1, v/v/v) mixture and incubated at 37 °C for 30 days. The mass spectra of AAAXAAA-NH₂ and AAAXAAA-OH peptides were recorded and given in Figure 5.6 and 5.7, respectively. As observed with XGGFL-NH₂, peptides containing C-terminus amide group resulted in the formation of a single methylation product. The ions *m/z* 587 and 573 in Figure 5.6 correspond to methylation of side chain of E in AAEEAAA-NH₂ and D in AAADAAA-NH₂, respectively. In addition, similar to the results obtained with XGGFL-OH series, peptides with free C-terminus carboxyl group resulted in the observation of two methylation products. The products of mono- and di-methylation reactions were highlighted with a red box in mass spectra of AAEEAAA-OH and

AAADAAA-OH peptides in Figure 5.7. Once again, these data demonstrated the higher methylation rate of E compared to D residue and higher abundance of mono-methylated products compared to di-methylated products. Besides, one can easily note that the methylation of E and D residues is not dependent on specific sequence of peptides.

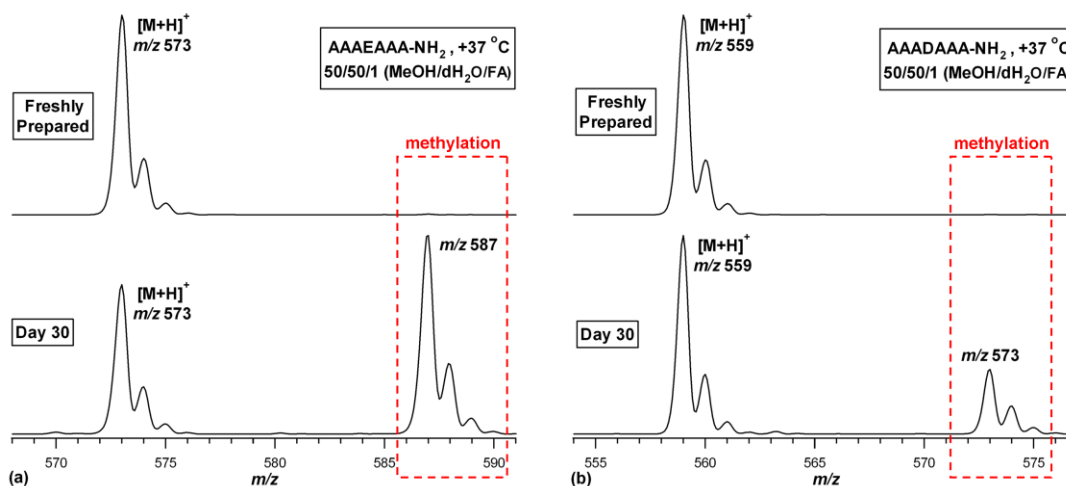


Figure 5.6. ESI-MS spectra of (a) AAAEAAA-NH₂ and (b) AAADAAA-NH₂ following incubation in MeOH/dH₂O/FA (50:50:1, v/v/v) at 37 °C for 30 days

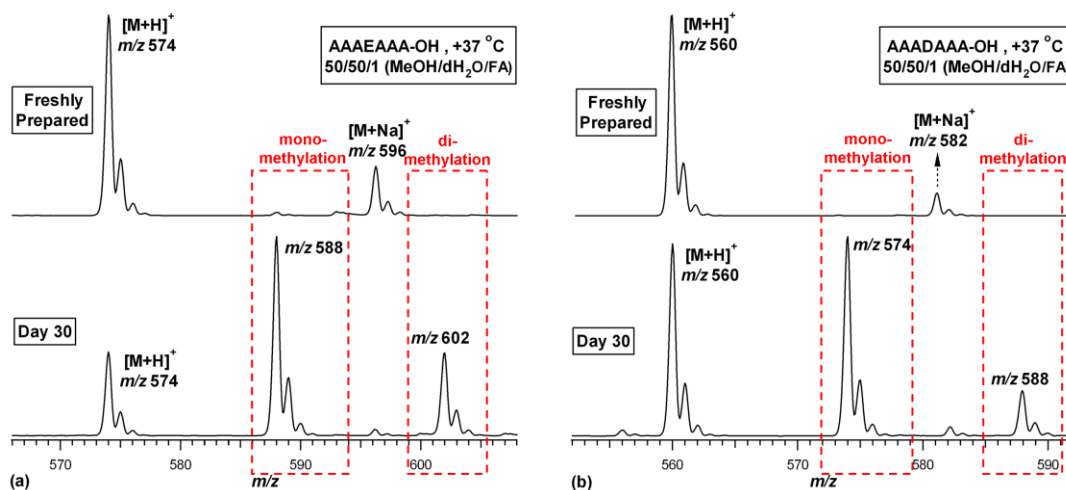


Figure 5.7. ESI-MS spectra of (a) AAAEAAA-OH and (b) AAADAAA-OH following incubation in MeOH/dH₂O/FA (50:50:1, v/v/v) at 37 °C for 30 days

The methylation of multiple acidic residues was investigated in this study, as well. The synthetic model peptides utilized for this purpose, AAAAAAEE-NH₂, AAAAAAEE-OH and AAAAAADD-OH, includes more than one acidic site available for methylation. The C-terminus amide group containing peptide has two potential methylation sites (7th and 8th acidic residues); whereas C-terminus free acid forms include three potential methylation sites (two adjacent acidic residues positioned at 7th and 8th and C-terminus of peptide). The acquired mass spectra of peptides, following incubation in MeOH/dH₂O/FA (50:50:1, v/v/v) mixture at 37 °C for 30 days, are given in Figure 5.8. For AAAAAAEE-NH₂ sample, two isomeric mono-methylated peptide products (AAAAAAE_{OMe}E-NH₂ or AAAAAAEE_{OMe}-NH₂) were observed at *m/z* 716. In addition, di-methylated product (AAAAAAE_{OMe}E_{OMe}-NH₂) was detected at *m/z* 730. For peptides with free C-terminus acid group, AAAAAAEE-OH and AAAAAADD-OH, mono-, di- and tri-methylated peptide products were observed and these products were highlighted in Figure 5.8b and Figure 5.8c. Concerning AAAAAAEE-OH as an example, three isomeric mono-methylated peptide derivatives, AAAAAAE_{OMe}E-OH, AAAAAAEE_{OMe}-OH and AAAAAAEE-OCH₃ were formed in solution. In addition, three di-methylated products AAAAAAE_{OMe}E_{OMe}-OH, AAAAAAEE_{OMe}-OCH₃, AAAAAAE_{OMe}E-OCH₃ and a tri-methylated product, AAAAAAE_{OMe}E_{OMe}-OCH₃, were observed. These results revealed that, mono-methylation rate is higher than di-methylation rate in AAAAAAEE-NH₂ and AAAAAADD-OH, and the level of tri-methylation is lower than mono- and di-methylation levels in peptides containing C-terminus carboxyl group. Moreover, overall rate of methylation (sum of mono-, di- and tri-methylation levels) of E residue is significantly higher than D, confirmed by the lower relative intensities of methylated products in AAAAAADD-OH peptide.

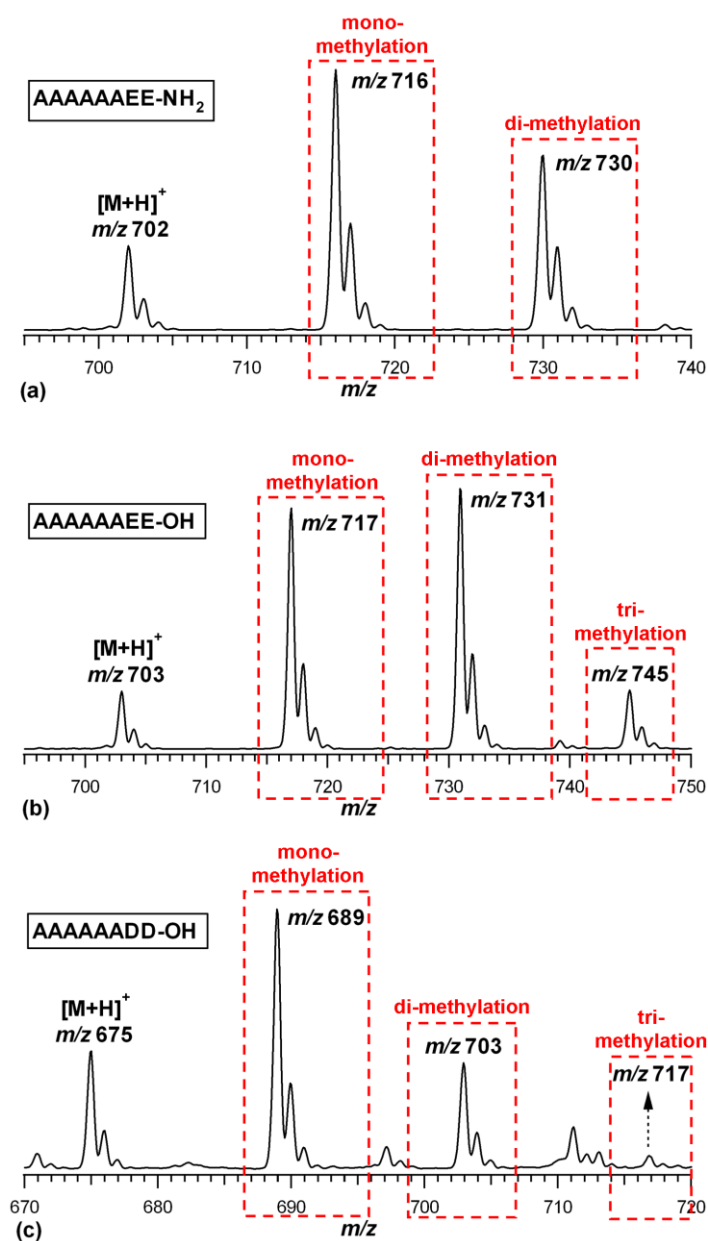


Figure 5.8. ESI-MS spectra of (a) AAAAAAEE-NH₂, (b) AAAAAAEE-OH and (c) AAAAAADD-OH following incubation in MeOH/dH₂O/FA (50:50:1, v/v/v) mixture at 37 °C for 30 days

Concerning the effect of acid type and its percentage included in MeOH/dH₂O (50:50, v/v) mixture on the formation of methylation of E and D residues, 0.1 or 1.0% concentrations of formic acid or acetic acid were included in MeOH/dH₂O mixture and the peptides were prepared using these solutions. Again, EGGFL-NH₂, DGGFL-NH₂, EGGFL-OH and DGGFL-OH synthetic model pentapeptides were used and 37 °C was selected as incubation temperature. The pH of MeOH/dH₂O/FA and MeOH/dH₂O/AA (50:50:1, v/v/v) mixtures were measured and their pK_a values were calculated as 4.71

and 5.76, respectively. The mass spectra of each sample were acquired over time periods of 1, 3, 7, 14 and 30 days and given in Appendix C (Figure C.5-C.8). For ease of comparison, the mass spectra of EGGFL-NH₂ and DGGFL-NH₂ samples of thirty-day incubation were reflected in Figure 5.9. This figure demonstrates that 0.1% concentration of acid, either FA or AA, resulted in a decrease in the level of methylated products in both peptide samples, compared to 1.0%. Moreover, it is evident that the methylation reaction is dependent on the type of acid used, having higher reaction rates in FA compared to AA. As an exception, only methylation level of E in EGGFL-NH₂ peptide incubated in solution containing 0.1% FA was almost the same with that of incubated in solution containing 0.1% AA. Nevertheless, on the basis of the given acidities (pKa) of FA as 3.75 and AA as 4.76 at 25 °C, it is likely that the strong acidic behavior of FA results in higher methylation rates. Additionally, EGGFL-OH and DGGFL-OH peptides, which contain two potential methylation sites, were used to investigate the effect of acid type and concentration on the methylation of multiple acidic sites. The results of these experiments were given in Appendix C (Figure C.9-C.12) and showed that the rate of mono- and di-methylation was reduced in both peptide samples when 0.1% concentration of both types of acid were used instead of 1.0%. Besides, decreased levels of mono- and di-methylated products were observed when AA was used in the incubation solution. Moreover, the rate of methylation of DGGFL-OH was higher than EGGFL-OH, especially in solutions containing FA. Overall, these results clearly demonstrated that the level of methylation of E/D residues and C-terminus carboxyl group was significantly reduced when ESI solution includes lower concentrations of any type of acid, as well as in the presence of acetic acid.

Lastly, the effect of different organic solvents on the esterification reaction was investigated. To that end, incubation solution was prepared with either ethanol (EtOH) or isopropanol (*i*PrOH), instead of methanol. These solvents were used to compare the effects of primary (MeOH or EtOH) and secondary alcohol (*i*PrOH) on the esterification reaction of E/D residues and C-terminus of peptide. The peptides, EGGFL-NH₂, DGGFL-NH₂, EGGFL-OH and DGGFL-OH, were freshly prepared in EtOH/dH₂O/FA or *i*PrOH/dH₂O/FA (50:50:1, v/v/v) mixture and incubated at 37 °C for 1, 3, 7, 14 and 30 days. Following incubation, full scan MS were recorded.

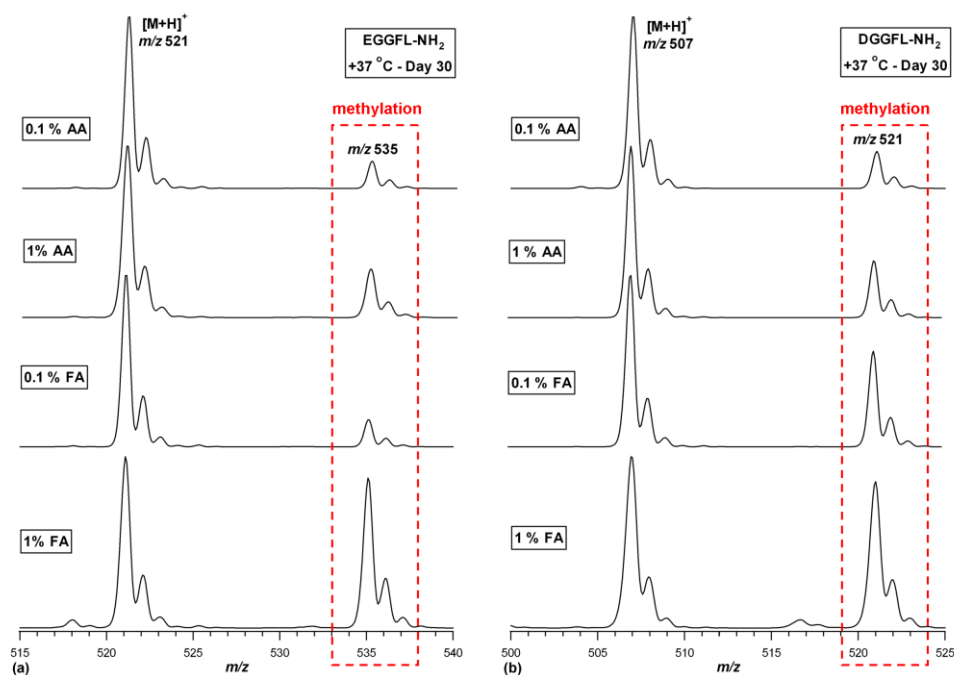


Figure 5.9. ESI-MS spectra of (a) EGGFL-NH₂ and (b) DGGFL-NH₂ following incubation in MeOH/dH₂O (50:50, v/v) mixture containing 0.1% or 1.0% of acetic acid or formic acid, at 37 °C for 30 days

The mass spectra of EGGFL-NH₂, DGGFL-NH₂ incubated in ethanol solution for 30 days are given in 5.10a and b. The recorded spectra of peptides for 1, 3, 7 and 14 days of incubation are shown in Appendix C (Figure C.13). The peaks at m/z 549 in EGGFL-NH₂ and m/z 535 in DGGFL-NH₂ were separated from their corresponding $[M + H]^+$ ions by +28 Da mass shift. These ions were not observed in freshly prepared samples and can be attributed to the ethylation of E and D residues. In terms of ethylation reaction rate, D residue is slightly higher than E residue. Similarly, previous studies reported the higher ethylation rate of D residues (151, 153, 161). The results for EGGFL-OH and DGGFL-OH peptides (Appendix Figure C.14) highlighted the mono- and di-ethylation of peptides by +28 Da and +56 Da mass shift, respectively. Mono-ethylation rate is higher than di-ethylation rate for both peptides and the mono-ethylation rate is slightly higher in DGGFL-OH compared to EGGFL-OH. On the other hand, following thirty-day of incubation at 37 °C in solution containing isopropanol, observation of +42 Da mass shift by ions m/z 563 in EGGFL-NH₂ and m/z 549 in DGGFL-NH₂, pointed the isopropylation of E and D residues (Figure 5.10 c and d). The recorded spectra of peptides for 1, 3, 7 and 14 days of incubation are shown in Appendix C (Figure C.15). Similar to ethylation reaction, the rate of isopropylation of D

residue is higher than E. Haebel et al. have also used *i*PrOH as an organic solvent in esterification reaction of E/D residues. However, they did not observe isopropylation of these residues (152). In investigating isopropylation of multiple sites, the results for the incubation of EGGFL-OH and DGGFL-OH peptides in *i*PrOH containing solutions (Appendix Figure C.16) showed that, both peptides only resulted in the observation of mono-isopropylation, though peaks corresponding to di-isopropylation were not detected as a result of thirty-day incubation period.

All in all, these results clearly demonstrated the impact of different solvents on esterification reaction. The level of isopropylation of E and D residues is significantly lower than the level of methylation and ethylation. Regarding the structures of these three molecules, observation of slower isopropylation reaction rates might be due the steric hinderance of *i*PrOH compared to MeOH and EtOH.

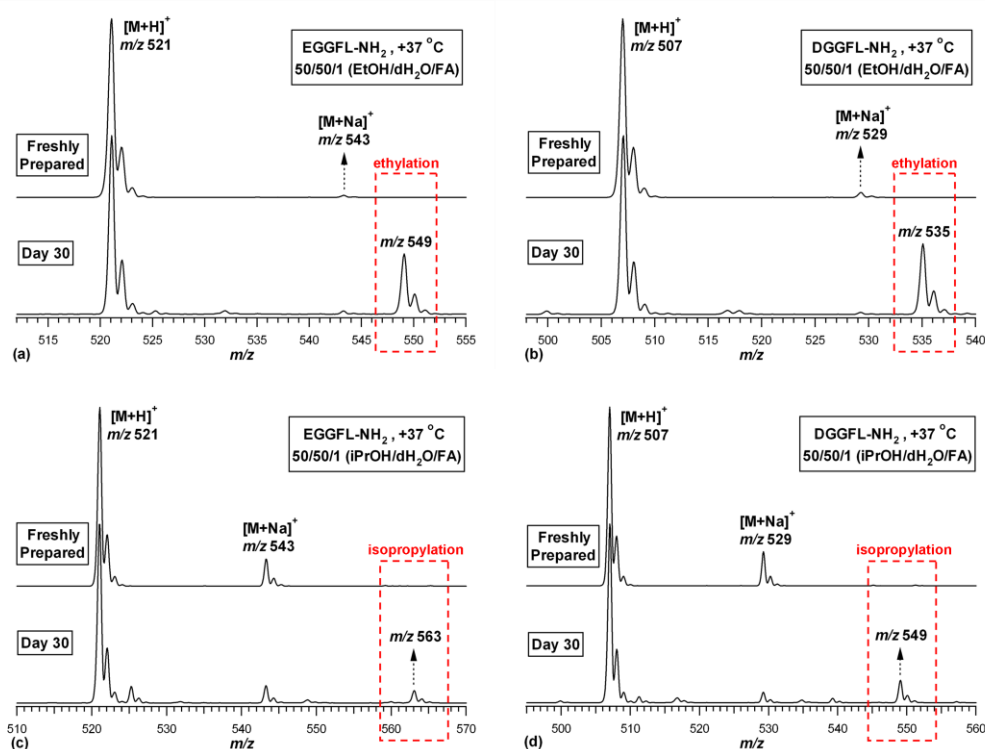


Figure 5.10. ESI-MS spectra of (a) EGGFL-NH₂ and (b) DGGFL-NH₂ following incubation in EtOH/dH₂O/FA (50:50:1, v/v/v) mixture, (c) EGGFL-NH₂ and (d) DGGFL-NH₂ following incubation in *i*PrOH/dH₂O/FA (50:50:1, v/v/v) mixture, at 37 °C for 30 days

5.4. Conclusion

Methylation of proteins is an important post-translational modification implicated in several biological processes. Besides, it is also an important *in vitro* modification that is observed in proteomic workflows. The utilization of acidic solutions of methanol during sample handling have been reported to result in the methylation of acidic amino acid residues on proteins, as well as C-terminus carboxyl groups. The observation of a novel peak spaced by +14 Da from the protonated molecular ion peak of peptides, which have been previously prepared in a common ESI working solution of MeOH/dH₂O/FA (50:50:1, v/v/v) mixture for MS analysis and stored in this solution for a long period of time, pointed the methylation of E and D residues of the peptides in ESI solution. Therefore, in this study, a comprehensive investigation of the extent of methylation of acidic amino acid residues (E and D) as well as C-terminus carboxyl group of model peptides in ESI solution was performed, with a motivation to demonstrate important parameters that effect the formation of this modification.

According to the data obtained in this study, methylation of side chains of E/D residues and C-terminus carboxyl group is temperature and time dependent reaction. As the incubation time and temperature increased, the rate of reaction increased significantly. The solution is stable at -20 °C, confirmed by the absence of methylated products in peptide samples incubated at this temperature for a month. The type and concentration of acid included in solution has a high impact on the rate of the reaction, such that solutions containing 1.0% formic acid resulted in the observation of higher reaction rates when compared to solutions containing 0.1% formic acid or acetic acid at the same concentration. Minimum amount of methylated products were observed in ESI solutions containing 0.1% acetic acid, therefore ESI solutions acidified with 0.1% acetic acid are recommended to avoid the methylation of peptides upon long-term storage. Moreover, it was noted that methylation of E/D residues does not require specific amino acid sequences and not effected by the length of peptide, position of the residue and the presence of multiple methylation sites within a peptide. When possible sites are available (C-terminus carboxyl group, multiple E/D residues), mono-, di- and tri-methylated peptides were detected, albeit at varying levels. Lastly, it was shown that ethylation and isopropylation reactions were also possible, when these organic solvents

were used instead of methanol in ESI solution. However, the rate of isopropylation is the lowest compared to other two organic solvents, due to the bulky nature of isopropanol, which directly effects the esterification reaction.

As a conclusion, storage of peptides prepared in acidified MeOH/dH₂O solutions at temperatures other than -20 °C, induces formation of methylated peptide products. Likewise, acidified solutions of methanol used in several steps of proteomic workflows can induce the methylation of certain amino acids on proteins. Therefore, researchers who use these solutions prior to peptide or protein analysis are cautioned on the factors reported in this study, if they want to avoid misidentification of endogenous methylation of peptides and proteins. Moreover, if acidified solutions of methanol are used during sample handling and identification of *in-vivo* methylation is not an issue, inclusion of methylation as a dynamic modification during MS/MS database searching may help to improve results. On the other hand, this study provides a simple and cost-effective method for methylation of E/D residues and C-terminus end of peptides.

CHAPTER 6

CONCLUSION

This PhD dissertation aimed the use of mass spectrometry for identification of proteins isolated from different organisms and identification of modifications on proteins and peptides observed as artifacts of sample preparation in proteomic workflows. These studies were investigated in 4 parts in the thesis.

In chapter 2, proteomic analyses of antimony resistant and sensitive *L. tropica* were carried out, in order to understand mechanism of antimony resistance in *in vitro* conditions and host organism. The proteome of antimony sensitive *L. tropica* isolates were compared with their corresponding *in vitro* generated antimony resistant *L. tropica* lines and with antimony resistant clinical isolate. In addition, the proteome of *in vitro* resistant lines were compared with resistant clinical isolate. Identification of differentially expressed proteins by mass spectrometry revealed that majority of the metabolic enzymes were down regulated, antioxidant defence mechanism and stress proteins were up regulated in *in vitro* resistant lines when compared with sensitive isolates; whereas stress proteins were down regulated, metabolic enzymes and antioxidant defense mechanism were upregulated in resistant clinical isolate when compared with sensitive isolate and *in vitro* resistant lines. These results indicated that the generation of antimony resistance in *Leishmania* is associated with altered abundance of proteins with different cellular functions. Moreover, the mechanism of generation of antimony resistance in *Leishmania* follow different strategies in host cell and *in vitro* conditions. Given the importance of drug resistance in treatment of Leishmaniasis, the results of this study may help to understand differences in drug resistance generated in different conditions and can be used in the further studies for the development of efficient therapeutics against Leishmaniasis.

In chapter 3, proteomic analysis of Bence Jones proteins isolated from urine of multiple myeloma patients was carried out so as to unravel the differences in renal insufficiency observed among patients. Among the 6 samples investigated, 1D-GE and MS analysis showed that light chains in samples 1, 2 and 3 tend to exist as monomers and light chains in samples 4, 5 and 6 tend to exist as dimers in their native state.

MS/MS analysis revealed that only sample 4 contained a lambda type of light chain, whereas others were identified as kappa type of light chain. The factors which may induce the formation of dimer structures of light chains on the basis of differences in amino acid sequence could not be deduced but it can be concluded that the nephrotoxicity caused by BJPs is complex and multifactorial. Regarding the results of previous studies using the same samples, formation of multimeric complexes of BJPs, ability of BJPs to inhibit DNA synthesis and tendency to induce production of inflammatory cytokines can all together give insights into the nephrotoxicity of BJPs.

In chapters 4 and 5, two different modifications of proteins and peptides were identified, which were introduced to samples as chemical artifacts during preparation of samples for proteomic studies. In chapter 4, it was shown that acetone, a commonly used organic solvent for precipitation and concentration of proteins, induced the formation of +98 u chemical artifacts on proteins, observed by MS spectra of intact proteins. The factors affecting the formation of modification was investigated and it was shown that the modification is suppressed when proteins are incubated in acidic solutions, at brief incubation times and low incubation temperatures. The modification site was localized to histidine, arginine and lysine residues. The product of modification was speculated to arise following formation of mesityl oxide via aldol condensation of acetone. In chapter 5, MS analysis revealed the methylation of aspartic acid, glutamic acid residues and C-terminus of peptides upon incubation of peptides in ESI working solution, composed of methanol and formic acid. The results have shown that the level of methylation reaction was increased at prolonged incubation times and higher incubation temperatures, in ESI solutions containing higher percentage of acid. The added acid type also influenced the level of modification, such that solutions containing formic acid resulted in the formation of higher level of methylation reaction when compared with solutions containing acetic acid. The formation of unintentional chemical modifications in proteomic workflows increase sample complexity and hamper the proper identification of proteins by MS. Therefore, if the factors that induce the formation of modifications are known, steps should be taken to minimize the level of modification or at least the mass shift for the modification should be included in MS/MS spectral database searching.

REFERENCES

- (1) Yates, J. R. III. A century of mass spectrometry: from atoms to proteomes. *Nat. Methods* **2011**, *8*, 633-637.
- (2) Griffiths, J. A brief history of mass spectrometry. *Anal. Chem.* **2008**, *80*, 5678-5683.
- (3) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and polymer analyses up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.
- (4) Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 Da. *Anal. Chem.* **1988**, *60*, 2299-2301.
- (5) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray ionization mass spectrometry of large molecules. *Science* **1989**, *246*, 64-71.
- (6) Wilkins, M. R.; Pasquali, C.; Appel, R. D.; Ou, K.; Golaz, O.; Sanchez, J. C.; Yan, J. X., Hughes, G.; Humphery-Smith, I.; Williams, K. L.; Hochstrasser, D. F. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology* **1996**, *14*, 61-65.
- (7) Gevaert, K.; Vandekerckhove, J. Protein identification methods in proteomics. *Electrophoresis* **2000**, *21*, 1145-1154.
- (8) Kinter, M.; Sherman, N. E. The primary structure of proteins and a historical overview of protein sequencing In *Protein sequencing and identification using tandem mass spectrometry*, 5; Desiderio, D. M., Nibbering, N. M. M., Eds.; John Wiley and Sons: New York, 2000; pp 10-15.
- (9) Lin, D.; Tabb, D. L.; Yates, J. R. III. Large-scale protein identification using mass spectrometry. *Biochim. Biophys. Acta* **2003**, *1646*, 1-10.
- (10) Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **2009**, *11*, 49-79.
- (11) Jonsson, A. P. Mass spectrometry for protein and peptide characterization. *Cell. Mol. Life Sci.* **2001**, *58*, 868-884.

- (12) Lane, C. S. Mass spectrometry-based proteomics in the life sciences. *Cell. Mol. Life Sci.* **2005**, *62*, 848-869.
- (13) El-Aneed, A.; Cohen, A.; Joseph, B. Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Appl. Spectrosc. Rev.* **2009**, *44*, 210-230.
- (14) Han, X.; Aslanian, A.; Yates, J. R. III. Mass spectrometry for proteomics. *Curr. Opin. Chem. Biol.* **2009**, *12*, 483-490.
- (15) Chen, C. H. Review of a current role of mass spectrometry for proteome research. *Anal. Chim. Acta* **2008**, *624*, 16-36.
- (16) Roepstorff, P.; Fohlmann, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
- (17) Biemann, K. Contributions of mass spectrometry to peptide and protein structure. *Biomed. Environ. Mass Spectrom.* **1988**, *16*, 99-111.
- (18) Harrison, A. G. To *b* or not to *b*: the ongoing saga of peptide *b* ions. *Mass Spectrom. Rev.* **2009**, *28*, 640-654.
- (19) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
- (20) Atik, A. E. Studies of gas-phase fragmentation mechanisms of peptide *b* ions by mass spectrometry. Ph. D. Dissertation, Izmir Institute of Technology, Izmir, 2013.
- (21) Cottrell, J. S. Protein identification using MS/MS data. *J. Proteomics* **2011**, *74*, 1842-1851.
- (22) Hale, J. E.; Gelfanova, V.; Ludwig, J. R.; Knierman, M. D. Application of proteomics for discovery of protein biomarkers. *Brief Funct. Genomic. Proteomic* **2003**, *2*, 185-193.
- (23) He, Q.; Chiu, J. Proteomics in biomarker discovery and drug development. *J. Cell. Biochem.* **2003**, *89*, 868-886.

- (24) Leifso, K.; Cohen-Freue, G.; Dogra, N.; Murray, A.; McMaster, W.R. Genomic and proteomic expression analysis of *Leishmania* promastigote and amastigote life stages: The *Leishmania* genome is constitutively expressed. *Mol. Biochem. Parasitol.* **2007**, 152, 35-46.
- (25) Lynn, M. A.; Marr, A.K.; McMaster, W.R. Differential quantitative proteomic profiling of *Leishmania infantum* and *Leishmania mexicana* density gradient separated membranous fractions. *J. Proteomics* **2013**, 82, 179-192.
- (26) Control of leishmaniasis: report of a meeting of the WHO expert committee on the control of leishmaniasis Geneva, Switzerland, March 22-26, 2010.
- (27) Pawar, H.; Sahasrabudde, N.A.; Renuse, S., Keerthikumar, S.; Sharma, J.; Kumar, G.S.; Venugopal, A.; Sekhar, N. R.; Kelkar, D.S.; Nemade, H.; Khobragade, S.N.; Muthusamy, B.; Kandasamy, K.; Harsha, H. C.; Chaerkady, R.; Patole M. S.; Pandey, A. A proteogenomic approach to map the proteome of an unsequenced pathogen *Leishmania donovani*. *Proteomics* **2012**, 12, 832-844.
- (28) Ravel, C.; Cortes, S.; Pratlong, F.; Morio, F.; Dedet, J.; Campino, L. First report of genetic hybrids between two very divergent *Leishmania* species: *Leishmania infantum* and *Leishmania major*. *Int. J. Parasitol.* **2006**, 36, 1383-1388.
- (29) Alvar, J.; Velez, I.D.; Bern, C.; Herrero, M.; Desjeux, P.; Cano, J.; Jannin, J.; Boer, M.; the WHO Leishmaniasis Control Team. Leishmaniasis worldwide and global estimates of its incidence. *PloS One* **2012**, 7, e35671.
- (30) Culha, G.; Akyar, I.; Yıldız Zeyrek, F.; Kurt, Ö.; Gündüz, C.; Özensoy Töz, S.; Östan, I.; Cavus, I.; Gulkan, B.; Kocagöz, T.; Özbel, Y.; Özbilgin, A. Leishmaniasis in Turkey: Determination of *Leishmania* species by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Iran. J. Parasitol.* **2014**, 9, 239-248.
- (31) Walker, J.; Vasquez, J.J.; Gomez, M. A.; Drummel-Smith, J.; Burchmore, R.; Girard, I.; Quелlette, M. Identification of developmentally-regulated proteins in *Leishmania panamensis* by proteome profiling of promastigote and axenic amastigotes. *Mol. Biochem. Parasitol.* **2006**, 147, 64-73.
- (32) Brotherton, M.C., Bourassa, S., Leprohon, P., Legare, D., Poirier, G.G., Droit, A., Quелlette, M. Proteomic and genomic analyses of antimony resistant *Leishmania infantum* mutant. *PloS One* **2013**, 8, e81899.
- (33) Ashutosh; Sundar, S.; Goyal, N. Molecular mechanisms of antimony resistance in *Leishmania*. *J. Med. Microbiol.* **2007**, 56, 143-153.

- (34) Jeddi, F.; Piarroux, R.; Mary, C. Antimony resistance in Leishmania, focusing on experimental research. *J. Trop. Med.* **2011**, Article ID 695382, 1-15.
- (35) Mishan-Shaked, P.; Ulrich, N.; Ephros, M.; Zilberstein, D. Novel intracellular SbV reducing activity correlates with antimony susceptibility in *Leishmania donovani*. *J. Biol. Chem.* **2001**, 276, 3971-3976.
- (36) Sereno, D.; Cavaleyra, M.; Zemzoumi, K.; Maquaire, S.; Quassi, A.; Lemesre, J. L. Axenically grown amastigotes of *Leishmania infantum* used as an in vitro model to investigate the pentavalent antimony mode of action. *Antimicrob. Agents Chemother.* **1998**, 42, 3097-3102.
- (37) Denton, H.; McGregor, J. C.; Coombs, H. Reduction of anti-leishmanial pentavalent antimonial drugs by a parasite-specific-thiol-dependent reductase, TDR1. *Biochem. J.* **2004**, 381, 405-412.
- (38) Zhou, Y.; Messier, N.; Quellette, M.; Rosen, B. P.; Mukhopadhyay, R. *Leishmania major* LmACR2 is a pentavalent antimony reductase that confers sensitivity to the drug pentostam. *J. Biol. Chem.* **2004**, 36, 37445-37451.
- (39) Frezard, F.; Demicheli, C.; Ferreira, C. S.; Costa, M. A. P. Glutathione-induced conversion of pentavalent antimony to trivalent antimony in meglumine antimoniate. *Antimicrob. Agents Chemother.* **2001**, 45, 913-916.
- (40) Cunnungham, M.L.; Fairlamb, A. H. Trypanothione reductase from *Leishmania donovani* purification, characterization and inhibition by trivalent antimonials. *Eur. J. Biochem.* **1995**, 230, 460-468.
- (41) Wyllie, S.; Fairlamb, A. H. Differential toxicity of antimonial compounds and their effects on glutathione homeostasis in a human leukaemia monocyte cell line. *Biochem. Pharmacol.* **2006**, 71, 257-267.
- (42) Wyllie, S.; Cunnungham, M. L.; Fairlamb, A. H. Dual action of antimonial drugs on thiol redox metabolism in the human pathogen *Leishmania donovani*. *J. Biol. Chem.* **2004**, 279, 39925-39932.
- (43) Rosen, B. P. Transport and detoxification systems for transition metals, heavy metals and metalloids in eukaryotic and prokaryotic microbes. *Comp. Biochem. Physiol. A* **2002**, 133, 689-693.
- (44) Gourbal, B.; Sonuc, N.; Bhattacharjee, H.; Lagare, D.; Sundar, S.; Quellette, M.; Rosen, B. P.; Mukhopadhyay, R. Drug uptake and modulation of drug resistance in *Leishmania* by an aquaglyceroporin. *J. Biol. Chem.* **2004**, 279, 31010-31017.

- (45) Decuypere, S.; Rijal, S.; Yardley, V.; Doncker, S.; Laurent, T.; Khannal, B.; Chappuis, F.; Dujardin, J. C. Gene expression analysis of the mechanism of natural Sb(V) resistance in *Leishmania donovani* isolates from Nepal. *Antimicrob. Agents Chemother.* **2005**, 49, 4616-4621.
- (46) Legare, D.; Richard, D.; Mukhopadhyay, R.; Stierhof, Y. D.; Rosen, B. P.; Haimeur, A.; Papadopoulou, B.; Quellette, M. The *Leishmania* ATP-binding cassette protein PGPA is an intracellular metal-thiol transporter ATPase. *J. Biol. Chem.* **2001**, 276, 26301-26307.
- (47) Callahan, H. L.; Roberts, W. L.; Rainey, P. M.; Beverley, S. M. The PGPA gene of *Leishmania major* mediates antimony (SbIII) resistance by decreasing influx and not by increasing efflux. *Mol. Biochem. Parasitol.* **1994**, 64, 145-149.
- (48) Coelho, A. C.; Beverley, S. M.; Cotrim, P. C. Functional genetic identification of PRP1, an ABC transporter superfamily member conferring pentamidine resistance in *Leishmania major*. *Mol. Biochem. Parasitol.* **2003**, 130, 83-90.
- (49) Meister, A.; Anderson, M. E. Glutathione. *Annu Rev Biochem* **1983**, 52, 711-760.
- (50) Mukhopadhyay, R.; Dey, S.; Xu, N.; Gage, D.; Lightbody, J.; Quellette, M.; Rosen, B. P. Trypanothione overproduction and resistance to antimonials and arsenicals in *Leishmania*. *Proc. Natl. Acad. Sci.* **1996**, 93, 10383-10387.
- (51) Fairlamb, A. H.; Cerami, A. Metabolism and functions of trypanothione in the Kinetoplastida. *Annu. Rev. Microbiol.* **1992**, 46, 695-729.
- (52) Haimeur, A.; Guimond, C.; Pilote, S.; Mukhopadhyay, R.; Rosen, B. P.; Poulin, R.; Quellette, M. Elevated levels of polyamines and trypanothione resulting from overexpression of the ornithine decarboxylase gene in arsenite-resistant *Leishmania*. *Mol. Microbiol.* **1999**, 34, 726-735.
- (53) Grondin, K.; Haimeur, A.; Mukhopadhyay, R.; Rosen, B. P.; Quellette, M. Co-amplification of the gamma-glutamylcysteine synthetase gene *gsh1* and of the ABC transporter gene *pgpA* in arsenite-resistant *Leishmania tarantolae*. *EMBO J.* **1997**, 11, 3057-3065.
- (54) Arana, F. E.; Perez-Victoria, J. M.; Repetto, Y.; Morello, A.; Castanys, S.; Gamarro, F. Involvement of thiol metabolism in resistance to glucantime in *Leishmania tropica*. *Biochem. Pharmacol.* **1998**, 56, 1201-1208.

- (55) Quellette, M.; Legare, D.; Papadopoulou, B. Multidrug resistance and ABC transporters in parasitic protozoa. *J. Mol. Microbiol. Biotechnol.* **2001**, 3, 201-206.
- (56) Vergnes, B.; Gourbal, B.; Girard, I.; Sundar, S.; Drummelsmith, J.; Quellette, M. A proteomics screen implicates HSP83 and a small Kinetoplastid calpain-related protein in drug resistance in *Leishmania donovani* clinical field isolates by modulating drug-induced programmed cell death. *Mol. Cell. Proteomics* **2007**, 6, 88-101.
- (57) Moreira, D. S.; Pescher, P.; Laurent C.; Lenormand, P.; Spath, G. F.; Murta, S. M. F. Phosphoproteomic analysis of wild-type and antimony-resistant *Leishmania braziliensis* lines by 2D-DIGE technology. *Proteomics* **2015**, 15, 2999-3019.
- (58) Matrangolo, F. S. V.; Liarte, D. B.; Andrade, L. C.; Melo, M. F.; Andrade, J. M.; Ferreira, R. F.; Santiago, A. S.; Pirovani, C. P.; Silva-Pereira, R. A.; Murta, S. M. F. Comparative proteomic analysis of antimony-resistant and –susceptible *Leishmania braziliensis* and *Leishmania infantum chagasi* lines. *Mol. Biochem. Parasitol.* **2013**, 190, 63-75.
- (59) Fadili, K. E.; Drummelsmith, J.; Roy, G.; Jardim, A.; Quellette, M. Down regulation of KMP-11 in *Leishmania infantum* axenic antimony resistant amastigotes as revealed by a proteomic screen. *Exp. Parasitol.* **2009**, 123, 51-57.
- (60) Kumar, A.; Sisodia, B.; Misra, P.; Sundar, S.; Shasany, A. K.; Dube, A. Proteome mapping of overexpressed membrane-enriched and cytosolic proteins in sodium antimony gluconate (SAG) resistant clinical isolate of *Leishmania donovani*. *Br. J. Pharmacol.* **2010**, 70, 609-617.
- (61) Biyani, N.; Singh, A. K.; Mandal, S.; Chawla, B.; Madhubala, R.; Differential expression of proteins in antimony-susceptible and –resistant isolates of *Leishmania donovani*. *Mol. Biochem. Parasitol.* **2011**, 179, 91-99.
- (62) Hajjaran, H.; Azarian, B.; Mohebbi, M.; Hadighi, R.; Assareh, A.; Vaziri, B. Comparative proteomics study on meglumine antimoniate sensitive and resistant *Leishmania tropica* isolated from Iranian anthroponotic cutaneous leishmaniasis patients. *East. Mediterr. Health J.* **2012**, 18, 165-171.
- (63) Walker, J.; Gongora, R.; Vasquez, J. J.; Drummelsmith, J.; Burchmore, R.; Roy, G.; Quellette, M.; Gomez, M. A.; Saravia, N. G. Discovery of factors linked to antimony resistance in *Leishmania panamensis* through differential proteome analysis. *Mol. Biochem. Parasitol.* **2012**, 183, 166-176.

- (64) Portman, N.; Gull, K. The paraflagellar rod of kinetoplastid parasites: From structure to components and function. *Int. J. Parasitol.* **2010**, *40*, 135-148.
- (65) Requena, J. M.; Mantalvo, A. M.; Fraga, J. Molecular chaperons of Leishmania: Central players in many stress-related and –unrelated physiological processes. *BioMed. Res. Int.* **2015**, Article ID 301326, 1-21.
- (66) Arya, R.; Mallik, M.; Lakhotia, S. C. Heat shock genes-integrating cell survival and death. *J. Biosci.* **2007**, *3*, 595-610.
- (67) Barratt, J.; Topham, P. Urine proteomics: the present and future of measuring urinary protein components in disease. *CMAJ.* **2007**, *177*, 361-368.
- (68) Hu, S.; Loo, J. A.; Wong, D. T. Human body fluid proteome analysis. *Proteomics* **2006**, *6*, 6329-6353.
- (69) Niwa, T. Biomarker discovery for kidney diseases by mass spectrometry. *J. Chromat. B.* **2008**, *870*, 148-153.
- (70) Wu, J.; Chen, Y. D.; Gu, W. Urinary proteomics as a novel tool for biomarker discovery in kidney diseases. *J. Zhejiang. Univ. Sci. B.* **2010**, *4*, 227-237.
- (71) Kocabaş, R.N.; Başol, G. Proteinüri ve laboratuvar değerlendirmesi. *Türk Klinik Biyokimya Derg.* **2006**, *4*, 133-145.
- (72) Chailyan, A.; Marcatili, P.; Cirillo, D.; Tramontano, A. Structural repertoire of immunoglobulin λ light chains. *Proteins* **2011**, *79*, 1513-1524.
- (73) Kaplan, B.; Livneh, A.; Sela, B. Immunoglobulin free light chain dimers in human diseases. *Scientific. World. J.* **2011**, *11*, 726-735.
- (74) Batuman, V. Proximal tubular injury in myeloma. In *The kidney in plasma cell dyscrasias: a current view and a look at the future*; Herrera, G. A., Eds.; Karger: Basel, 2007; 153, pp 87-104.
- (75) Beck, A.; Sanglier-Cianferani, S.; Dorselaer, A. V. Biosimilar, biobetter and next generation antibody characterization by mass spectrometry. *Anal. Chem.* **2012**, *84*, 4637-4646.
- (76) Williams, K.; Williams, J.; Marshall, T. Analysis of Bence Jones proteinuria by high resolution two-dimensional electrophoresis. *Electrophoresis* **1998**, *19*, 1828-1835.

- (77) (Sikkink, L. A.; Ramirez-Alvarado, M. Biochemical and aggregation analysis of Bence Jones proteins from different light chain diseases. *Amyloid* **2008**, *15*, 29-39.
- (78) Marshall, T.; Williams K. M. Electrophoretic analysis of Bence Jones proteinuria. *Electrophoresis* **1999**, *20*, 1307-1324.
- (79) Jones, H. B. On a new substance occurring in the urine of a patient with mollities ossium. *Phil. Trs. Roy. Soc. London* **1848**, *138*, 55.
- (80) Edelman, G. M.; Gally, J. A. The nature of Bence Jones proteins. *J. Exp. Med.* **1962**, *116*, 207-227.
- (81) Makino, D. L.; Henschen-Edman, A. H.; Larson, S. B.; McPherson, A. Bence Jones KWR protein structures determined by X-ray crystallography. *Acta Cryst.* **2007**, *D63*, 780-792.
- (82) Poshusta, T. L.; Sikkink, L. A.; Leung, N.; Clark, R. J.; Dispenzieri, A.; Ramirez-Alvarado, M. Mutations in specific structural regions of immunoglobulin light chains are associated with free light chain levels in patients with AL amyloidosis. *PloS One* **2009**, *4*, e5169.
- (83) Abraham, R. S.; Charlesworth, M. C.; Oweb, B. L.; Benson, L. M.; Katzmman, J. A.; Reeder, C. B.; Kyle, R. A. Trimolecular complexes of λ light chain dimers in serum of a patient with multiple myeloma. *Clin. Chem.* **2002**, *48*, 1805-1811.
- (84) Comenzo, R. L.; Zhang, Y.; Martinez, C.; Osman, K.; Herrera, G. A. The tropism of organ involvement in primary systemic amyloidosis: contributions of Ig V (L) germ line use and clonal plasma cell burden. *Blood* **2001**, *98*, 714-720.
- (85) Markowitz, G. S.; Flis, R. S.; Kambham, N.; D'Agati, V. D. Fanconi syndrome with free kappa light chains in the urine. *Am. J. Kidney Dis.* **2000**, *35*, 777-781.
- (86) Batuman, V.; Sastrasinh, M.; Sastrasinh, S. Light chain effects on alanine and glucose uptake by renal brush border membranes. *Kidney Int.* **1986**, *30*, 662-655.
- (87) Batuman, V.; Guan, S.; O'Donovan, R.; Puschett, J.B. Effect of myeloma light chains on phosphate and glucose transport in renal proximal tubule cells. *Ren. Physiol. Biochem.* **1994**, *17*, 294-300.

- (88) Guan, S.; el-Dahr, S.; Dipp, S.; Batuman, V. Inhibition of Na-K-ATPase activity and gene expression by a myeloma light chain in proximal tubule cells. *J. Investig. Med.* **1999**, *47*, 496-501.
- (89) Pote, A.; Zwizinski, C.; Simon, E. E.; Meleg-Smith, S.; Batuman, V. Cytotoxicity of myeloma light chains in cultured human kidney proximal tubule cells. *Am. J. Kidney Dis.* **2000**, *36*, 735-744.
- (90) Sengul, S.; Zwizinski, C.; Simon, E. E.; Kapasi, A.; Singhal, P. C.; Batuman, V. Endocytosis of light chains induces cytokines through activation of NF- κ B in human proximal tubule cells. *Kidney Int.* **2002**, *62*, 1977-1988.
- (91) Sengul, S.; Zwizinski, C.; Batuman, V. Role of MAPK pathways in light chain-induced cytokine production in human proximal tubule cells. *Am. J. Physiol. Renal Physiol.* **2003**, *284*, F1245-1254.
- (92) Kishida F.; Azuma, T.; Hamaguchi, K. A type κ Bence Jones protein containing cysteinyl residue in the variable region. *J. Biochem.* **1975**, *77*, 481-491.
- (93) Khamlichi, A. A.; Aucouturier, P.; Silvain, C.; Bauwens, M.; Touchard, G.; Preud'Homme, J. L.; Nau, F.; Cogne, M. Primary structure of a monoclonal κ chain in myeloma with light chain deposition disease. *Clin. Exp. Immunol.* **1992**, *87*, 122-126.
- (94) Klafki, H. W.; Kratzin, H. D.; Pick, A. I.; Eckart, K.; Karas, M.; Hilschmann, N. Complete amino acid sequence determinations demonstrate identity of the urinary Bence Jones protein (BJP-DIA) and the amyloid fibril protein (AL-DIA) in a case of AL-amyloidosis. *Biochemistry* **1992**, *31*, 3265-3272.
- (95) Lim, A.; Wally, J.; Walsh, M.; Skinner, M.; Costello, C. E. Identification and location of a cysteinyl posttranslational modification in an amyloidogenic κ 1 light chain protein by electrospray ionization and matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Biochem.* **2001**, *295*, 45-56.
- (96) Barnidge, D. R.; Dispenzieri, A.; Merlini, G.; Katzmann, J. A.; Murray, D. L. Monitoring free light chains in serum using mass spectrometry. *Clin. Chem. Lab. Med.* **2016**, *54*, 1073-1083.
- (97) VanDuijn, M. M.; Jacobs, J. F.; Wevers, R. A.; Engelke, U. F.; Joosten, I.; Luider, T. M. Quantitative measurement of immunoglobulins and free light chains using mass spectrometry. *Anal. Chem.* **2015**, *87*, 8268-8274.

- (98) Barnidge, D. R.; Dasari, S.; Ramirez-Alvarado, M.; Fontan, A.; Willrich, M. A.; Tschumper, R. C.; Jelinek, D. F.; Snyder, M. R.; Dispenzieri, A.; Katzmann, J. A.; Murray, D. L. Phenotyping polyclonal kappa and lambda light chain molecular mass distributions in patient serum using mass spectrometry. *J. Proteome Res.* **2014**, *13*, 5198-5205.
- (99) Bergen, H. R.; Abraham, R. S.; Johnson, K. L.; Bradwell, A. R.; Naylor, S. Characterization of amyloidogenic immunoglobulin light chains directly from serum by on-line immunoaffinity isolation. *Biomed. Chromatogr.* **2004**, *18*, 191-201.
- (100) Vrana, J. A.; Gamez, J. D.; Madden, B. J.; Theis, J. D.; Bergen, H. R.; Dogan, A. Classification of amyloidosis by laser microdissection and mass spectrometry-based proteomic analysis in clinical biopsy specimens. *Blood* **2009**, *114*, 4957-4959.
- (101) Clyne, D. H.; Pesce, A. J.; Thompson, R. E. Nephrotoxicity of Bence Jones proteins in the rat: importance of protein isoelectric point. *Kidney Int.* **1979**, *16*, 345-352.
- (102) Coward, R. A.; Delamore, I. W.; Mallick, N. P.; Robinson, E. L. The importance of urinary immunoglobulin light chain isoelectric point (pI) in nephrotoxicity in multiple myeloma. *Clin. Sci. (Lond)*. **1984**, *66*, 229-232.
- (103) Johns, E. A.; Turner, R.; Cooper, E. H.; MacLennan, I. C. Isoelectric points of urinary light chains in myelomatosis: analysis in relation to nephrotoxicity. *J. Clin. Pathol* **1986**, *39*, 833-837.
- (104) Norden, A. G.; Flynn, F. V.; Fulcher, L. M.; Richards, J. D. Renal impairment in myeloma: negative association with isoelectric point of excreted Bence-Jones protein. *J. Clin. Pathol.* **1989**, *42*, 59-62.
- (105) Tracy, R. P.; Currie, R. M.; Kyle, R. A.; Young, D. S. Two-dimensional gel electrophoresis of serum specimens from patients with monoclonal gammopathies. *Clin. Chem.* **1982**, *28*, 900-907.
- (106) Harrison, H. H. Fine structure of "light chain ladders" in urinary immunofixation studies revealed by ISO-DALT two-dimensional gel electrophoresis. *Clin. Chem.* **1990**, *36*, 1526-1527.
- (107) Harrison, H. H. The "Ladder Light Chain" or "Pseudo-Oligoclonal" pattern in urinary immunofixation electrophoresis (IFE) studies: a distinctive IFE pattern and an explanatory hypothesis relating it to free polyclonal light chains. *Clin. Chem.* **1991**, *37*, 1559-1564.

- (108) Marshall, T.; Williams K. M. High resolution two-dimensional gel electrophoresis of human urinary proteins. *Anal. Chim. Acta.* **1998**, 372, 147-160.
- (109) Batuman, V.; Guan, S. Receptor mediated endocytosis of immunoglobulin light chains by renal proximal tubule cells. *Am. J. Physiol.* **1997**, 272, F521-530.
- (110) Batuman, V.; Verroust, P. J.; Navar, G. L. Myeloma light chains are ligands for cubulin (gp280). *Am. J. Physiol.* **1998**, 275, F246-254.
- (111) Lawrance, T. The nuclear factor NF- κ B pathway in inflammation. *Cold Spring Harb. Perspect. Biol.* **2009**, 1, a001651.
- (112) Dinarello, C. A. Proinflammatory cytokines. *Chest* **2000**, 118, 503-508.
- (113) Jiang, L.; He, L.; Fountoulakis, M. Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *J. Chromatogr. A* **2004**, 1023, 317-320.
- (114) Burgess, R. R. Protein precipitation techniques. In *Methods in Enzymology*, 2nd ed.; Burgess, R. R.; Deutscher, M. P.; Eds.; Elsevier Inc: San Diego; 2009; pp 331-342.
- (115) Green, A. A. The amphoteric properties of certain globulin fractions of normal horse serum. *J. Am. Chem. Soc.* **1938**, 60, 1108–1115.
- (116) Chick, B. Y. H.; Martin, C. J. The precipitation of egg-albumin by ammonium sulphate. A contribution to the theory of the “salting-out” of proteins. *Biochem. J.* **1913**, 7, 380–398.
- (117) Zeppezauer, M.; Brishammar, S. Protein precipitation by uncharged water-soluble polymers. *Biochim. Biophys. Acta. Biophys. Incl. Photosynth.* **1965**, 94, 581-583.
- (118) Scopes, R. K. Protein purification: Principles and practice; Springer: New York, NY, 1994; pp 238-269.
- (119) Crowell, A. M. J.; Wall, M. J.; Doucette, A. A. Maximizing recovery of water-soluble proteins through acetone precipitation. *Anal. Chim. Acta* **2013**, 796, 48–54.

- (120) Polson, C.; Sarkar, P.; Incledon, B.; Raguvaran, V.; Grant, R. Optimization of protein precipitation based upon effectiveness of protein removal and ionization effect in liquid chromatography-tandem mass spectrometry. *J. Chromatogr. B* **2003**, *785*, 263-275.
- (121) Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6*, 359-362.
- (122) Puchades, M.; Westman, A.; Blennow, K.; Davidsson, P. Removal of sodium dodecyl sulfate from protein samples prior to matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 344-349.
- (123) Catherman, A. D.; Durbin, K. R.; Ahlf, D. R.; Early, B. P.; Fellers, R. T.; Tran, J. C.; Thomas, P. M.; Kelleher, N. L. Large-scale top-down proteomics of the human proteome: membrane proteins, mitochondria, and senescence. *Mol. Cell. Proteomics* **2013**, *12*, 3465-3473.
- (124) Wu, C.; Tran, J. C.; Zamdborg, L.; Durbin, K. R.; Li, M.; Ahlf, D. R.; Early, B. P.; Thomas, P. M.; Sweedler, J. V.; Kelleher, N. L. A protease for “middle-down” proteomics. *Nat. Methods* **2012**, *9*, 822-824.
- (125) Kachuk, C.; Stephen, K.; Doucette, A. Comparison of sodium dodecyl sulfate depletion techniques for proteome analysis by mass spectrometry. *J. Chromatogr. A* **2015**, *1418*, 158-166.
- (126) An, B.; Zhang, M.; Johnson, R. W.; Qu, J. Surfactant-aided precipitation/on-pellet-digestion (SOD) procedure provides robust and rapid sample preparation for reproducible, accurate and sensitive LC/MS quantification of therapeutic protein in plasma and tissues. *Anal. Chem.* **2015**, *87*, 4023-4029.
- (127) Jeannot, M. A.; Zheng, J.; Li, L. Observation of gel-induced protein modifications in sodium dodecylsulfate polyacrylamide gel electrophoresis and its implications for accurate molecular weight determination of gel-separated proteins by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 512-520.
- (128) Sprung, R.; Chen, Y.; Zhang, K.; Cheng, D.; Zhang, T.; Peng, J.; Zhao, Y. Identification and validation of eukaryotic aspartate and glutamate methylation in proteins. *J. Proteome Res.* **2008**, *7*, 1001-1006.
- (129) Manning, M. C.; Chou, D. K.; Murphy, B. M.; Payne, R. W.; Katayama, D. S. Stability of protein pharmaceuticals: an update. *Pharm. Res.* **2010**, *27*, 544-575.

- (130) Xing, G.; Zhang, J.; Chen, Y.; Zhao, Y. Identification of four novel types of in vitro protein modifications. *J. Proteome Res.* **2008**, *7*, 4603-4608.
- (131) Stark, G. R.; Stein, W. H.; Moore, S. Reactions of the cyanate present in aqueous urea with amino acids and proteins. *J. Biol. Chem.* **1960**, *235*, 3177–3181.
- (132) Zheng, S.; Doucette, A. A. Preventing N- and O- formylation of proteins when incubated in concentrated formic acid. *Proteomics* **2016**, *16*, 1059-1068.
- (133) Narita, K. Reaction of anhydrous formic acid with proteins. *J. Am. Chem. Soc.* **1959**, *81*, 1751-1756.
- (134) Hine, J.; Cholod, M. S.; Chess, W. K. Kinetics of the formation of imines from acetone and primary amines. Evidence for internal acid-catalyzed dehydration of certain intermediate carbinolamines. *J. Am. Chem. Soc.* **1973**, *95*, 4270–4276.
- (135) Layer, R. W. The chemistry of imines. *Chem. Rev.* **1963**, *63*, 489–510.
- (136) Simpson, D. M.; Beynon, R. J. Acetone precipitation of proteins and the modification of peptides. *J. Proteome Res.* **2010**, *9*, 444–450.
- (137) Conant, J. B.; Tuttle, N. Mesityl oxide. *Org. Synth.* **1921**, *1*, 53.
- (138) Fisher, W. R.; Taniuchi, H.; Anfinsen, C. B. On the role of heme in the formation of the structure of cytochrome c. *J. Biol. Chem.* **1973**, 3188–3195.
- (139) Orton, D. J.; Wall, M. J.; Doucette, A. A. Dual LC-MS platform for high-throughput proteome analysis. *J. Proteome Res.* **2013**, *12*, 5963–5970.
- (140) Chowdhury, S. K.; Katta, V.; Beavis, R. C.; Chait, B. T. Origin and removal of adduct (molecular mass = 98u) attached to peptide and protein ions in electrospray mass spectra. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 382-388.
- (141) Klein, F. G.; Banchero, J. T. Condensation of acetone to mesityl oxide: sulfonated polystyrene-divinylbenzene resin as ion exchange resin catalyst. *Ind. Eng. Chem.* **1956**, *48*, 1278–1286.
- (142) Cooke, B. K.; Western, N. M. Diacetone alcohol, an artefact of acetone extraction of soil. *Analyst* **1980**, *105*, 4–7.

- (143) Mann, M.; Jensen, O. N. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **2003**, *21*, 255-261.
- (144) Witze, E. S.; Old, W. M.; Resing, K. A.; Ahn, N. G. Mapping protein post-translational modifications with mass spectrometry. *Nat. Methods* **2007**, *4*, 798-806.
- (145) Paik, W. K.; Paik, D. C.; Kim, S. Historical review: the field of protein methylation. *Trends Biochem. Sci.* **2007**, *32*, 146-152.
- (146) Bedford, M. T. Arginine methylation at a glance. *J. Cell Sci.* **2007**, *120*, 4243-4246.
- (147) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto Jr, G. J. Protein posttranslational modifications: the chemistry of proteome diversifications. *Agrew. Chem. Int. Ed. Engl.* **2005**, *44*, 7342-7372.
- (148) Sprung, R.; Chen, Y.; Zhang, K.; Cheng, D.; Zhang, T.; Peng, J.; Zhao, Y. Identification and validation of eukaryotic aspartate and glutamate methylation in proteins. *J. Proteome Res.* **2008**, *7*, 1001-1006.
- (149) Wang, K.; Zhou, Y.; Liu, H.; Cheng, K.; Mao, J.; Wang, F.; Liu, W.; Ye, M.; Zhao, Z.K.; Zou, H. Proteomic analysis of protein methylation in yeast *Saccharomyces cerevisiae*. *J. Proteomics* **2015**, *114*, 226-233.
- (150) Biterge, B.; Richter, F.; Mittler, G.; Schneider, R. Methylation of histone H4 at aspartate 24 by Protein L-isoaspartate O-methyltransferase (PCMT1) links histone modifications with protein homeostasis. *Sci Rep.* **2014**, *4*, 6674.
- (151) Haebel, S.; Albrecht, T.; Sparbier, K.; Walden, P.; Korner, R.; Steup, M. Electrophoresis-related protein modification: alkylation of carboxy residues revealed by mass spectrometry. *Electrophoresis* **1998**, *19*, 679-686.
- (152) Sumpton, D.; Bienvenut, W. Coomassie stains: are they really mass spectrometry compatible? *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1525-1529.
- (153) Chen, G.; Liu, H.; Wang, X.; Li, Z. In vitro methylation by methanol: proteomic screening and prevalence investigation. *Anal. Chim. Acta* **2010**, *661*, 67-75.
- (154) Nakao, R.; Oka, K.; Fukumoto, T. A simple method for the esterification of carboxylic acids using chlorosilanes. *Bull. Chem. Soc. Japan.* **1981**, *54*, 1267-1268.

- (155) Lee, A. S.-Y.; Yang, H.-C.; Su, F.-Y. An unprecedented and highly chemoselective esterification method. *Tetrahedron Lett.* **2001**, *42*, 301–303.
- (156) Chen, M.-Y.; Lee, A. S.-Y. A simple and efficient esterification method. *J. Chin. Chem. Soc.* **2003**, *50*, 103–108.
- (157) Jung, S. Y.; Li, Y.; Wang, Y.; Chen, Y., Zhao, Y.; Qin, J. Complications in the assignment of 14 and 28 Da mass shift detected by mass spectrometry as in vivo methylation from endogenous proteins. *Anal. Chem.* **2008**, *80*, 1721–1729.
- (158) Dookeran, N. N.; Yalcin, T.; Harrison, A. G. Fragmentation reactions of protonated α -amino acids. *J. Mass Spectrom.* **1996**, *31*, 500–508.
- (159) Harrison, A. G. Ion chemistry of protonated glutamic acid derivatives. *Int. J. Mass Spectrom.* **2001**, *210/211*, 361–370.
- (160) Chelius, D.; Jing, K.; Lueras, A.; Rehder, D. S.; Dillon, T. M.; Vizel, A.; Rajan, R. S.; Li, T.; Treuheit, M. J.; Bondarenko, P. V. Formation of pyroglutamic acid from N-terminus glutamic acid in immunoglobulin gamma antibodies. *Anal. Chem.* **2006**, *78*, 2370–2376.
- (161) Xing, G.; Zhang, J.; Chen, Y.; Zhao, Y. Identification of four novel types of in vitro protein modifications. *J. Proteome Res.* **2008**, *7*, 4603–4608.

APPENDIX A

BUFFERS AND GEL PREPARATION FOR SDS-PAGE (LAEMMLI BUFFER SYSTEM)

Stock Solutions

A. Acrylamide/Bisacrylamide Mixture (30% T, 2.67% C)

- 29.2 g acrylamide
- 0.8 g N'N'-bis-methylene-acrylamide

Dissolve acrylamide and bisacrylamide in ~80 mL distilled water and make up to 100 mL with distilled water. Filter and store at 4°C in the dark.

B. 1.5 M Tris-HCl, pH 8.8

- 18.15 g Tris Base
- ~80 mL distilled water

Dissolve Tris base in distilled water, adjust to pH 8.8 with HCl. Make up to 100 mL with distilled water and store at 4°C.

C. 0.5 M Tris-HCl, pH 6.8

- 6 g Tris Base
- ~80 mL distilled water

Dissolve Tris base in distilled water, adjust to pH 6.8 with HCl. Make up to 100 mL with distilled water and store at 4°C

D. 10% SDS

Dissolve 10 g SDS in 90 mL distilled water with gentle stirring and bring to 100 mL with distilled water.

E. Reducing Sample Buffer

- 3.8 mL distilled water
- 1.0 mL 0.5 M Tris-HCl, pH 6.8
- 0.8 mL Glycerol
- 1.6 mL 10% (w/v) SDS
- 0.4 mL 2-mercaptoethanol
- 0.4 mL 1% (w/v) bromophenol blue

F. 5X Running Buffer

- 15 g Tris Base
- 72 g Glycine
- 5 g SDS

Dissolve Tris base, glycine and SDS in ~800 mL distilled water and make up to 1 L with water. Store at 4°C. For electrophoretic run, dilute 5X stock solution to 1X with distilled water.

G. 10% Ammonium persulfate (APS)

Dissolve 0.1 g APS in 1 mL distilled water. This solution should be prepared fresh daily.

H. Colloidal Coomassie Staining Solution

Dissolve 50 g ammonium sulfate in ~300 mL distilled water, add 11.76 mL 85% o-phosphoric acid and 0.5 g Coomassie Brilliant Blue G-250. Make up to 400 mL with distilled water and finally add 100 mL methanol. Store at 4°C.

J. Destaining Solution

Prepare 25% (v/v) methanol solution in distilled water.

Preparation of Separating and Stacking Gel

- Separating Gel

Table A.1. Preparation of 12% and 15% separating gel (for 10 mL)

Buffer/Reagent	Volume (12%)	Volume (15%)
Distilled water	3.35 mL	2.35 mL
1.5 M Tris-HCl, pH 8.8	2.5 mL	2.5 mL
10% SDS	100 μ L	100 μ L
Acrylamide Stock	4 mL	5 mL
10% APS	50 μ L	50 μ L
TEMED	5 μ L	5 μ L

- Stacking Gel

Table A.2. Preparation of 4% stacking gel (for 10 mL)

Buffer/Reagent	Volume
Distilled water	6.1 mL
0.5 M Tris-HCl, pH 6.8	2.5 mL
10% SDS	100 μ L
Acrylamide Stock	1.33 mL
10% APS	50 μ L
TEMED	10 μ L

APPENDIX B

HIGH RESOLUTION MS SPECTRUM OF CYTOCHROME C AND LIST OF ACETONE MODIFIED *E. COLI* PEPTIDES

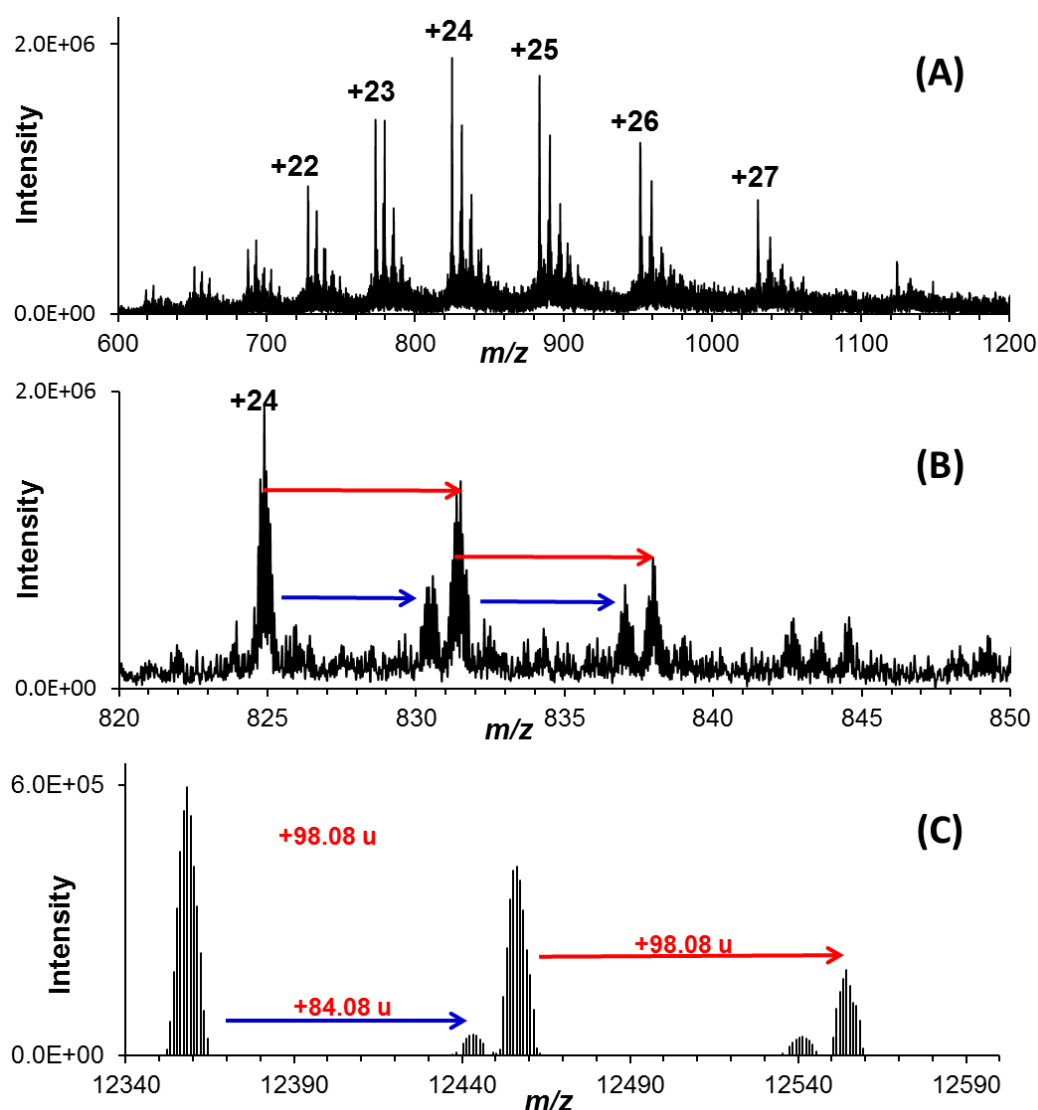


Figure B.1. (a) ESI MS spectrum of cytochrome c following overnight incubation in acetone (0 °C), as observed on an Orbitrap Velos Pro. (b) A zoomed in region of the +24 charge state ion shows formation of multiple adducts. (c) The deconvoluted MS spectrum shows the primary adduct (red arrow) to be spaced at 98.08 ± 0.01 u. The smaller adduct (blue arrow) is spaced at 84.06 ± 0.01 u above from unmodified protein

Table B.1. List of *E. coli* peptides identified with a +98.073 Da acetone modification

Peptide Sequence	Protein Accession	Modification	XCorr	Charge	ΔM [ppm]	RT [min]
FhTLSSGGKPVVEGAEDYTDSDD	P0ADN2	H2(Acetone)	3.46	3	-0.73	73.65
KhQKPVPALNQPGGIVEK	P60624	H2(Acetone)	3.40	3	0.34	57.03
IhhLTDDSFDTDVLK	P0AA25	H3(Acetone)	3.95	3	-0.65	96.48
AHhVGEWASLR	P0AB14	H3(Acetone)	3.19	3	0.74	63.66
TVThMQDEAANFPDPVDR	P0ABS1	H4(Acetone)	3.38	3	-1.10	80.52
LMEIAQQQhAQQQTAGADASANNAK... DDDVVDAEFEEVK	P0A6Y8	H9(Acetone)	6.60	4	1.10	86.20
LMEIAQQQhAQQQTAGADASANNAK	P0A6Y8	H9(Acetone)	4.49	3	-0.27	62.89
kiIGEQLGVK	P0A6A8	K1(Acetone)	3.88	3	-2.19	66.34
SGDTLSAISKQVYGNANLYNK	P0ADE6	K10(Acetone)	3.78	3	0.64	93.11
DAGFQAFADkVLDAAVAGK	P0A6P1	K10(Acetone)	3.42	3	-1.22	100.52
DkPEDA VLDVQGLATVTPAIVQAcTQDK	P0AES9	K2(Acetone); C24(Carbamido)	4.65	3	-0.01	98.23
EGSVkGYAGDTATTSEIK	P0AFH8	K5(Acetone)	3.31	3	-0.75	61.10
EQWGkLTDDDMTIEGK	P68206	K5(Acetone)	3.30	3	-2.15	96.06
LTVAkLNIDQNPGTAPK	P0AA25	K5(Acetone)	3.15	3	-0.99	76.96
WFNESkGFGFITPADGSK	P0A9Y6	K6(Acetone)	3.02	3	-2.11	96.92
AMDEQgkSLDDFLIKQ	P0ACF8	K7(Acetone)	3.38	3	-0.92	94.42
SLDDFLIKQ	P0ACF8	K8(Acetone)	2.92	2	-0.89	98.15

APPENDIX C

ESI-MS SPECTRA OF METHYLATED PEPTIDES

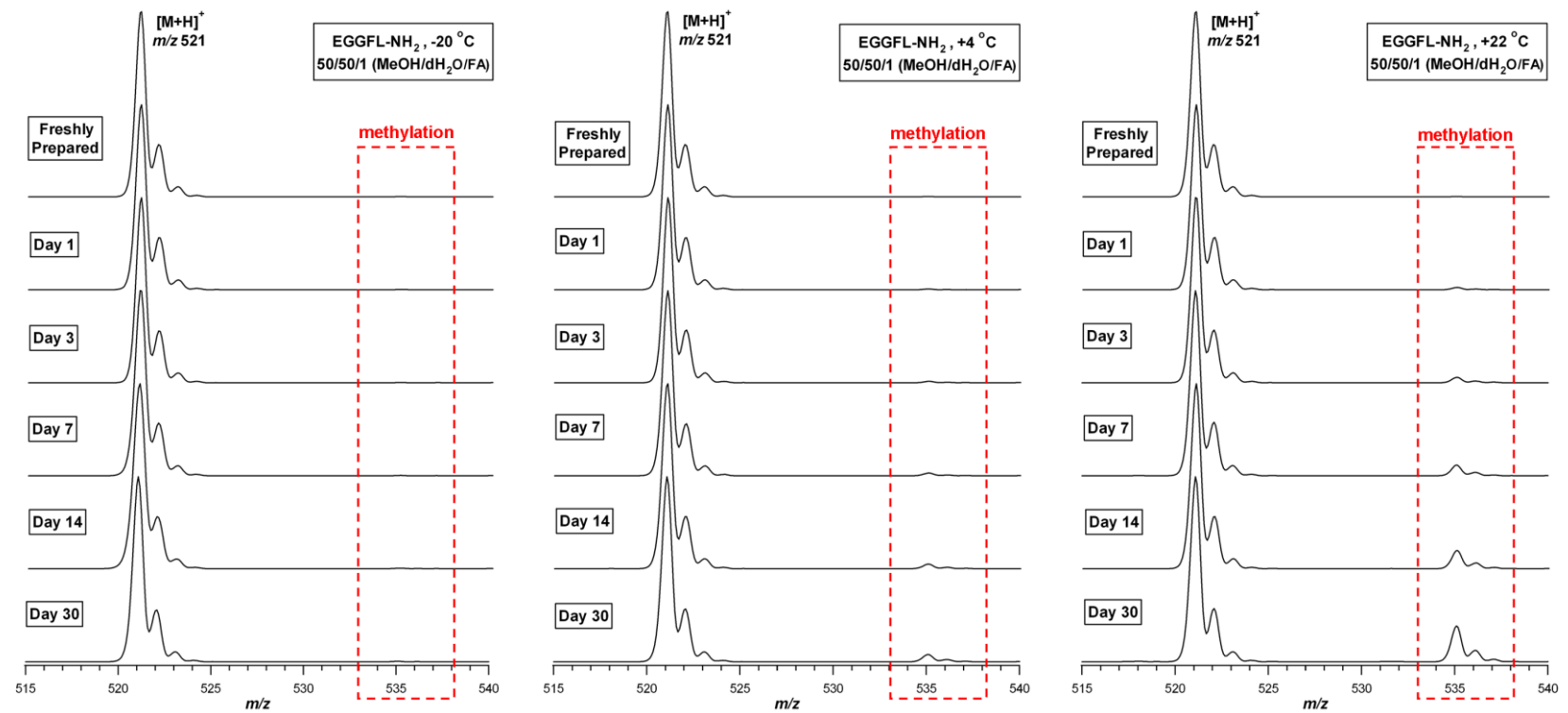


Figure C.1. ESI-MS spectra of EGGFL-NH₂ following incubation at -20 °C, 4 °C, or 22 °C over a range of incubation times. The peptide solutions were prepared in 50:50:1 (v/v/v) MeOH/dH₂O/FA

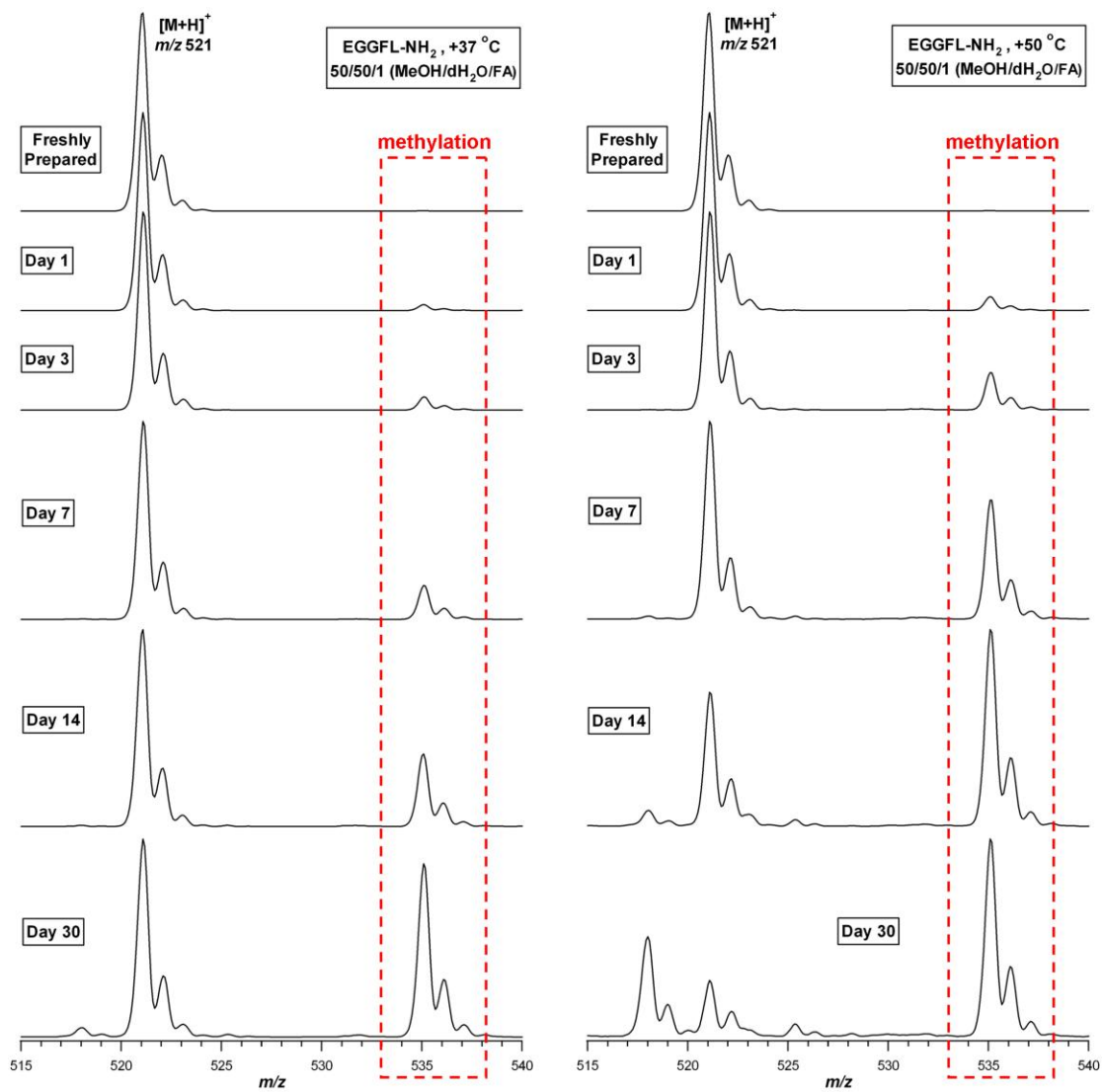


Figure C.2. ESI-MS spectra of EGGFL-NH₂ following incubation at 37 °C or 50 °C over a range of incubation times. The peptide solutions were prepared in 50:50:1 (v/v/v) MeOH/dH₂O/FA

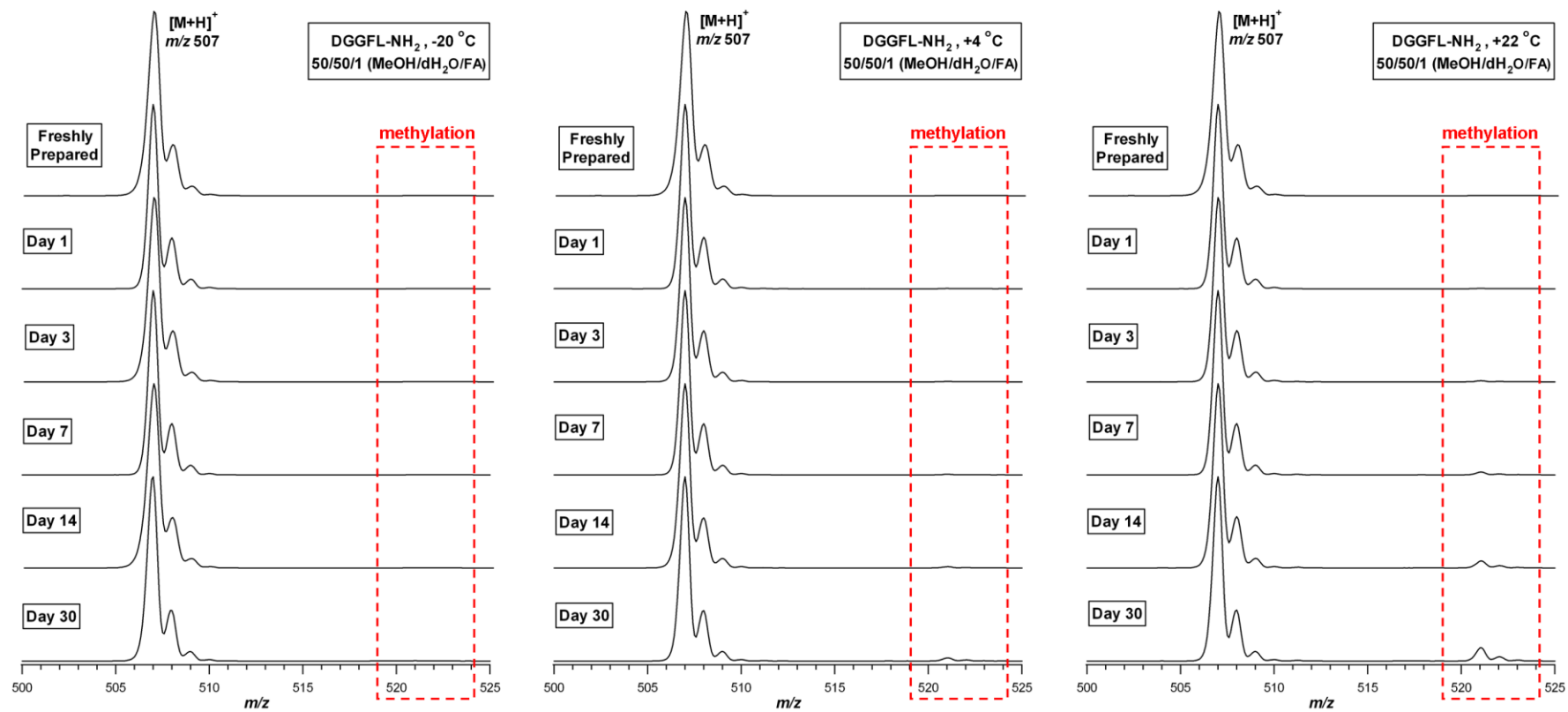


Figure C.3. ESI-MS mass spectra of DGGFL-NH₂ following incubation at -20 °C, 4 °C, or 22 °C over a range of incubation times. The peptide solutions were prepared in 50:50:1 (v/v/v) MeOH/dH₂O/FA

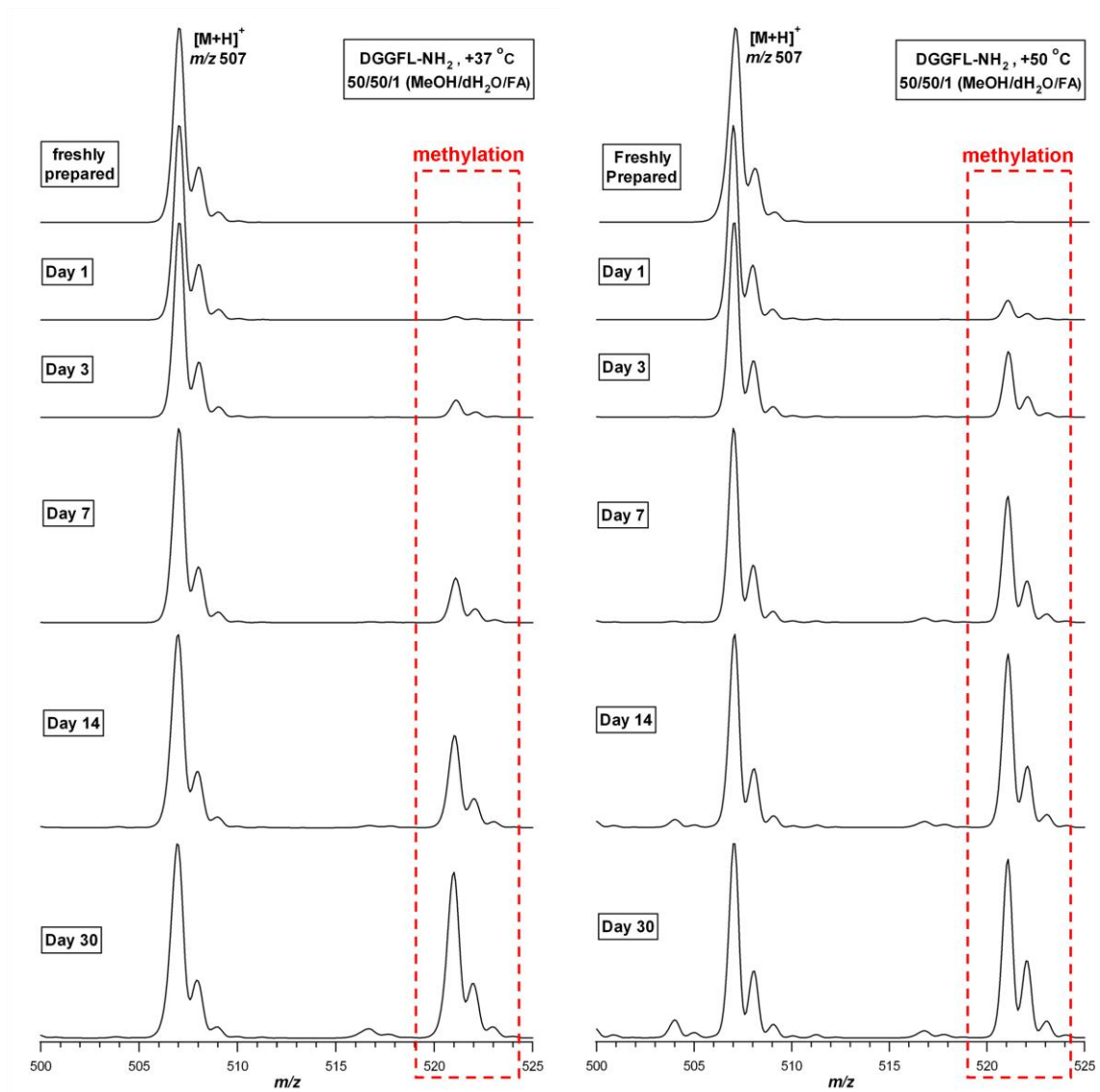


Figure C.4. ESI-MS mass spectra of DGGFL-NH₂ following incubation at 37 °C or 50 °C over a range of incubation times. The peptide solutions were prepared in 50:50:1 (v/v/v) MeOH/dH₂O/FA

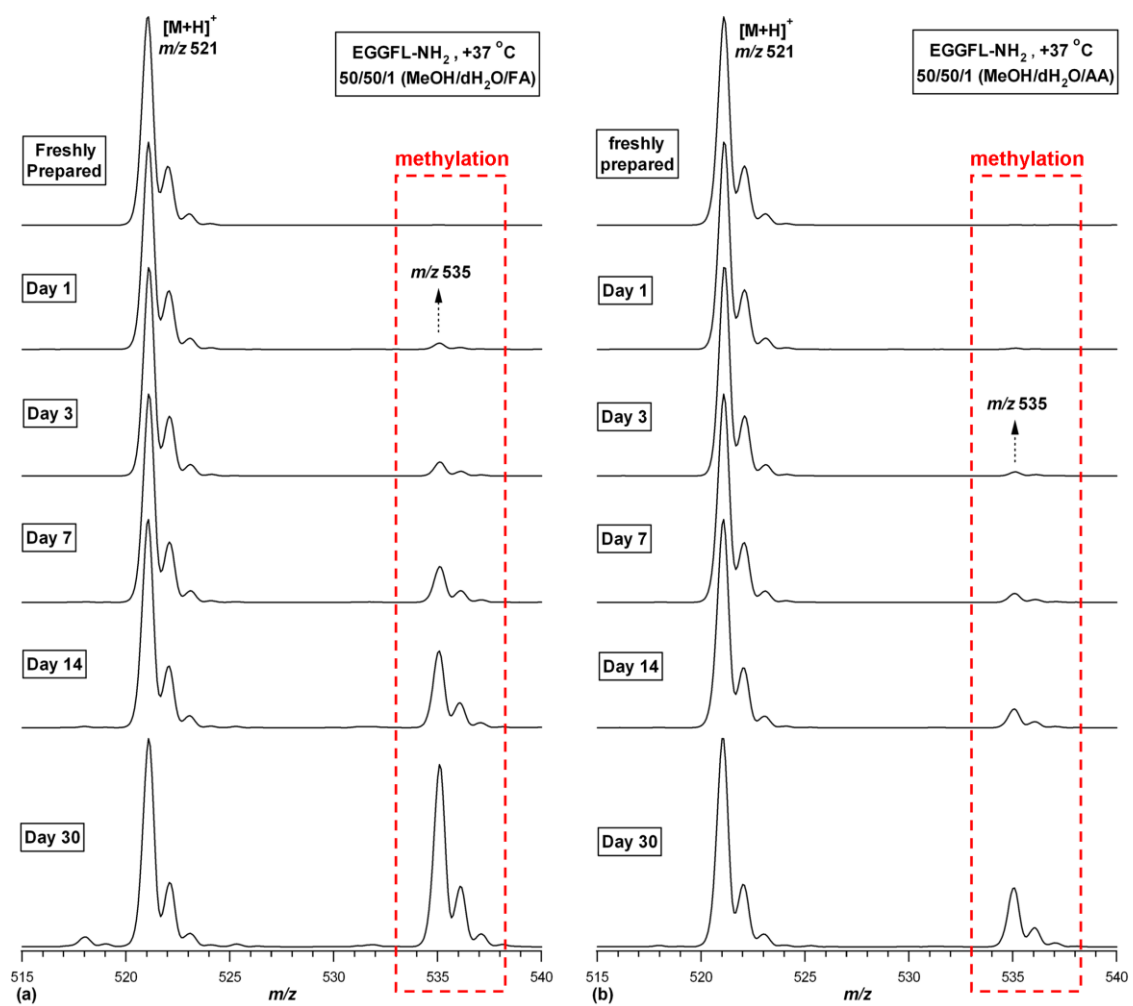


Figure C.5. ESI-MS spectra of EGGFL-NH₂ prepared in (a) 50:50:1 (v/v/v) MeOH/dH₂O/FA or (b) 50:50:1 (v/v/v) MeOH/dH₂O/AA, incubated at 37 °C over a range of incubation times

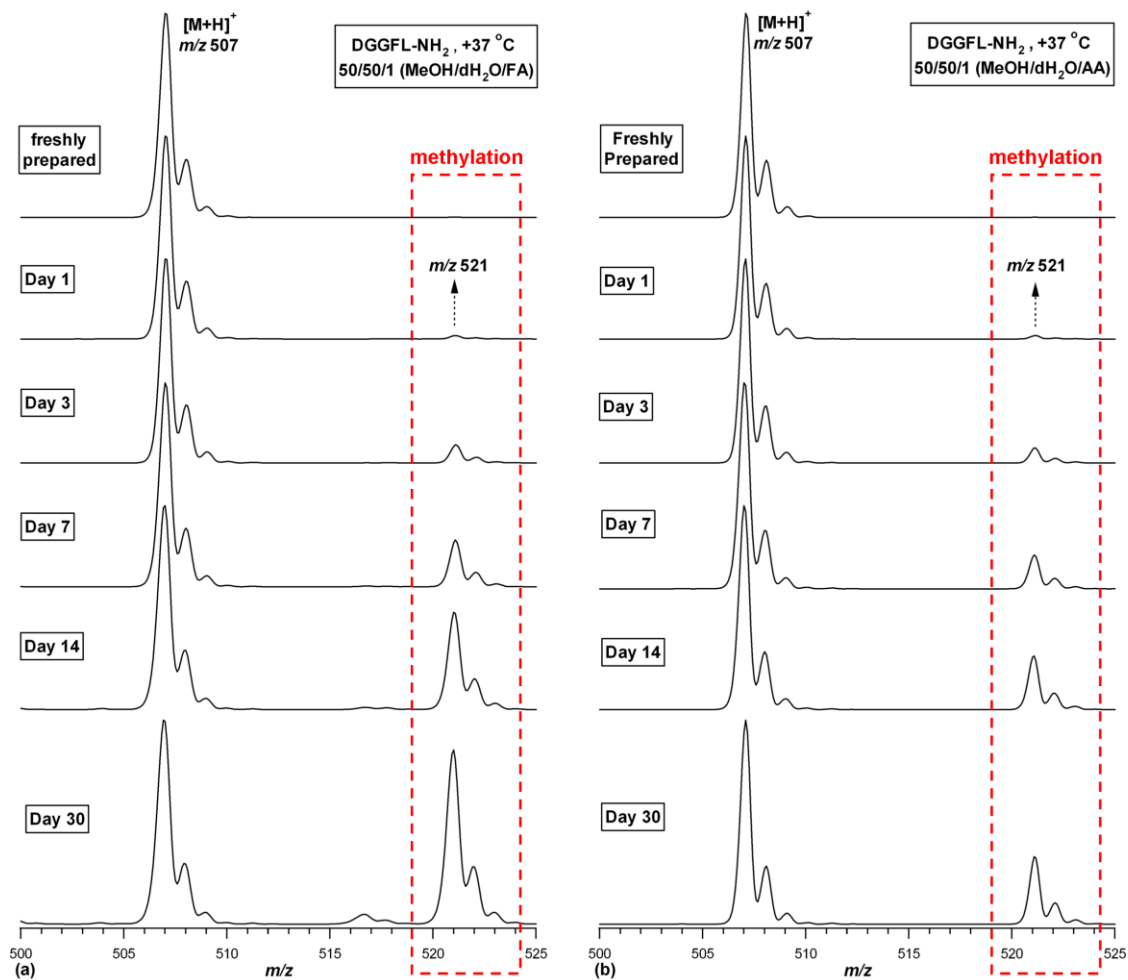


Figure C.6. ESI-MS spectra of DGGFL-NH₂ prepared in (a) 50:50:1 (v/v/v) MeOH/dH₂O/FA or (b) 50:50:1 (v/v/v) MeOH/dH₂O/AA, incubated at 37 °C over a range of incubation times

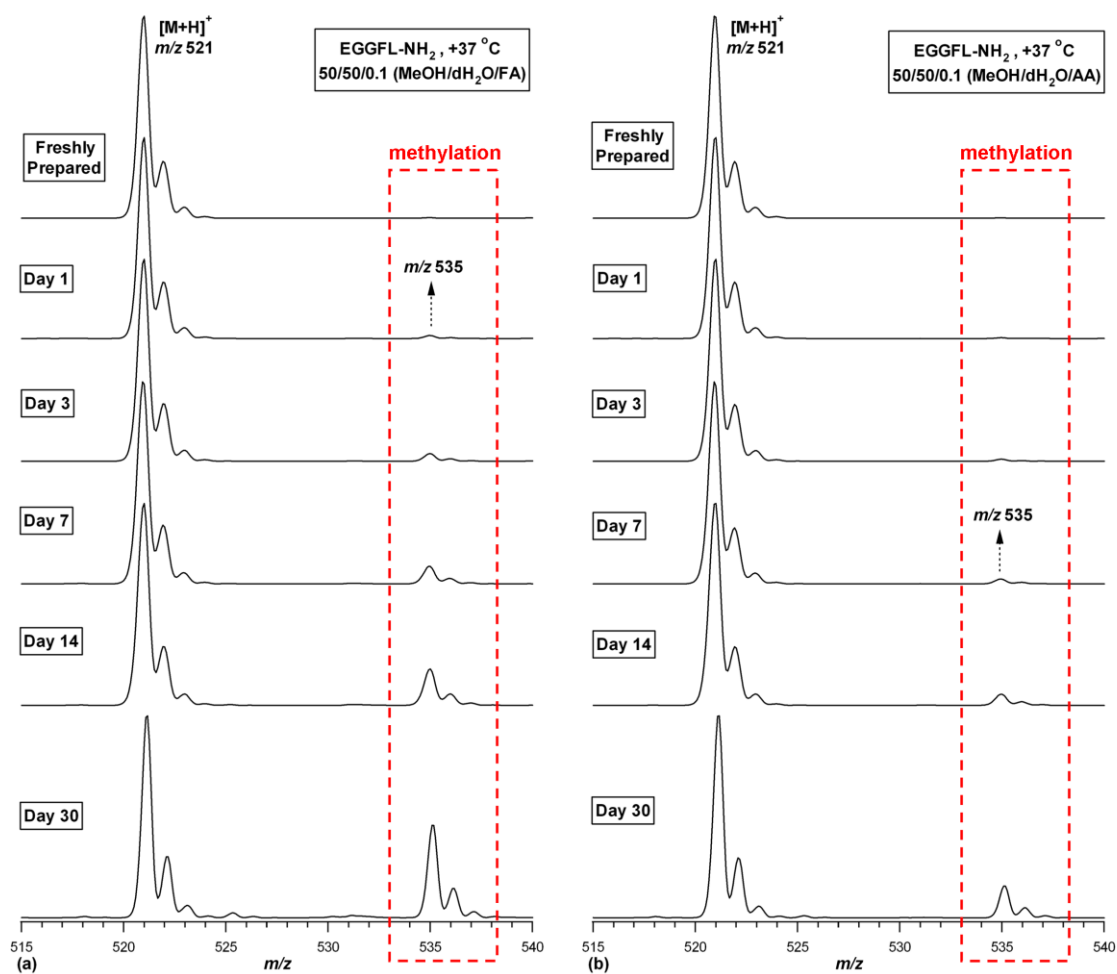


Figure C.7. ESI-MS spectra of EGGFL-NH₂ prepared in (a) 50:50:0.1 (v/v/v) MeOH/dH₂O/FA or (b) 50:50:0.1 (v/v/v) MeOH/dH₂O/AA, incubated at 37 °C over a range of incubation times

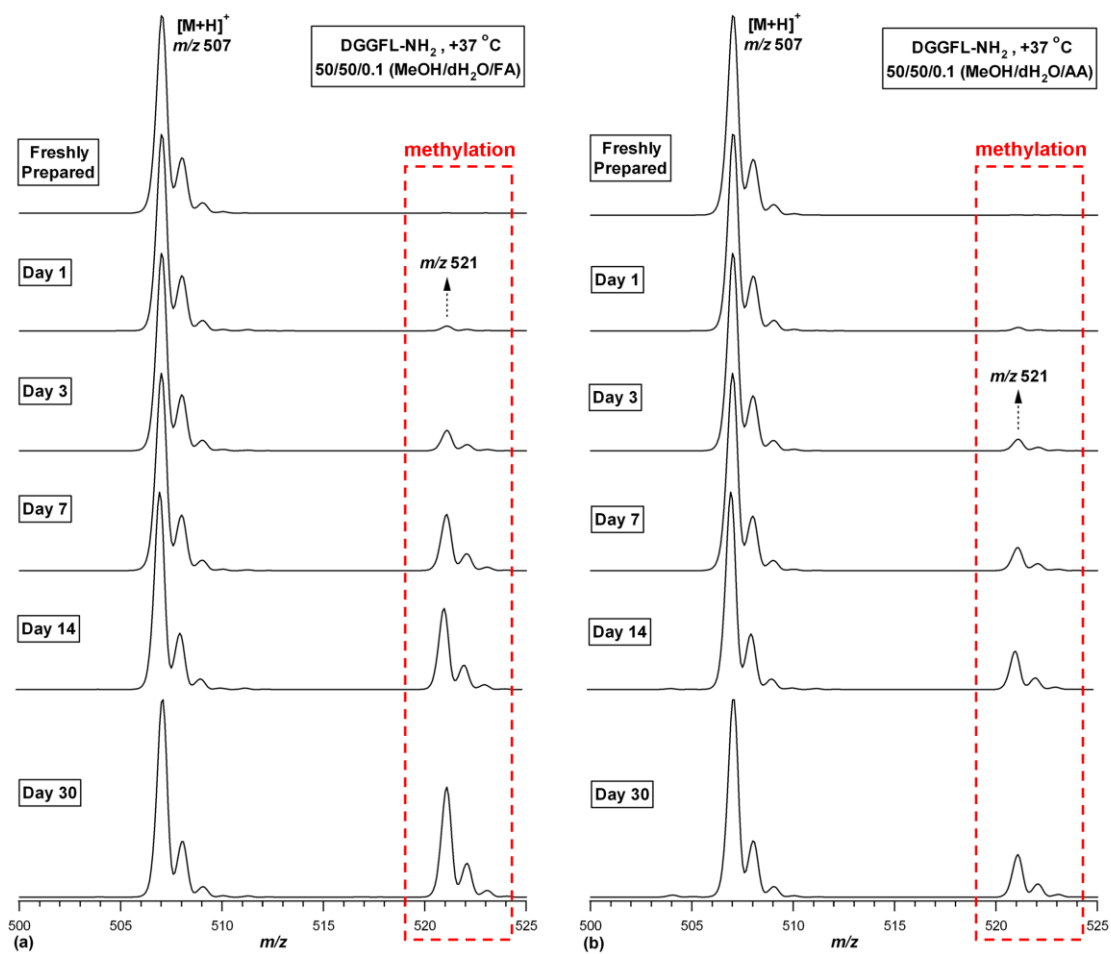


Figure C.8. ESI-MS spectra of DGGFL-NH₂ prepared in (a) 50:50:0.1 (v/v/v) MeOH/dH₂O/FA or (b) 50:50:0.1 (v/v/v) MeOH/dH₂O/AA, incubated at 37 °C over a range of incubation times

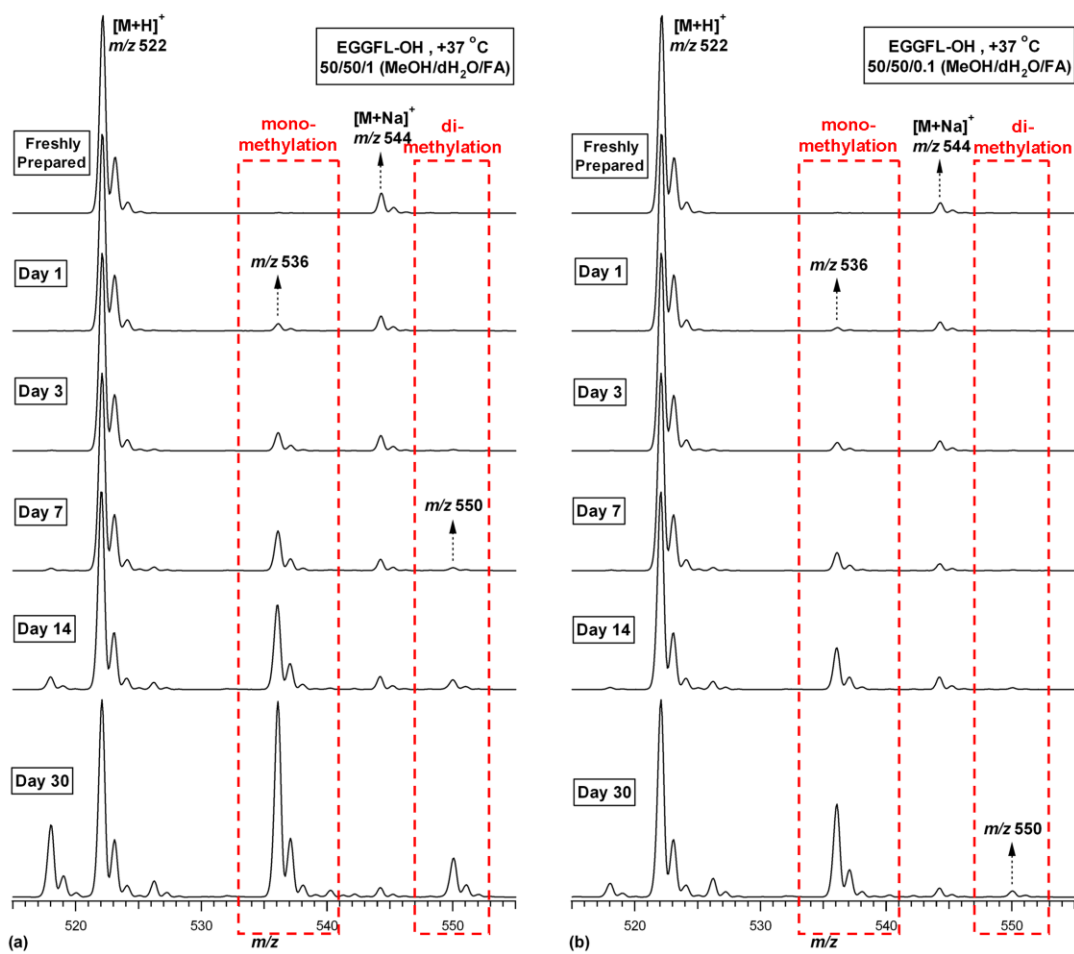


Figure C.9. ESI-MS mass spectra of EGGFL-OH prepared in (a) 50:50:1 (v/v/v) MeOH/dH₂O/FA or (b) 50:50:0.1 (v/v/v) MeOH/dH₂O/FA, incubated at 37 °C over a range of incubation times

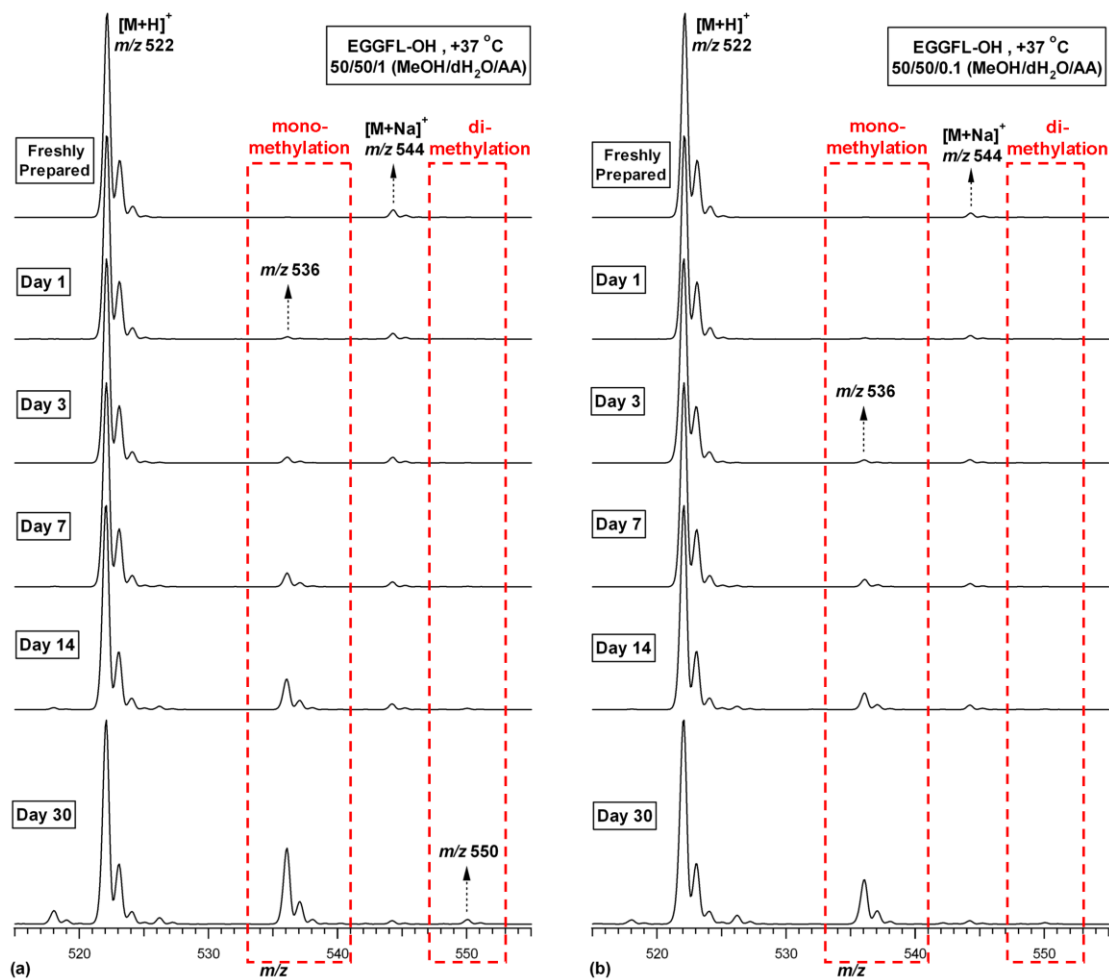


Figure C.10. ESI-MS spectra of EGGFL-OH prepared in (a) 50:50:1 (v/v/v) MeOH/dH₂O/AA or (b) 50:50:0.1 (v/v/v) MeOH/dH₂O/AA, incubated at 37 °C over a range of incubation times

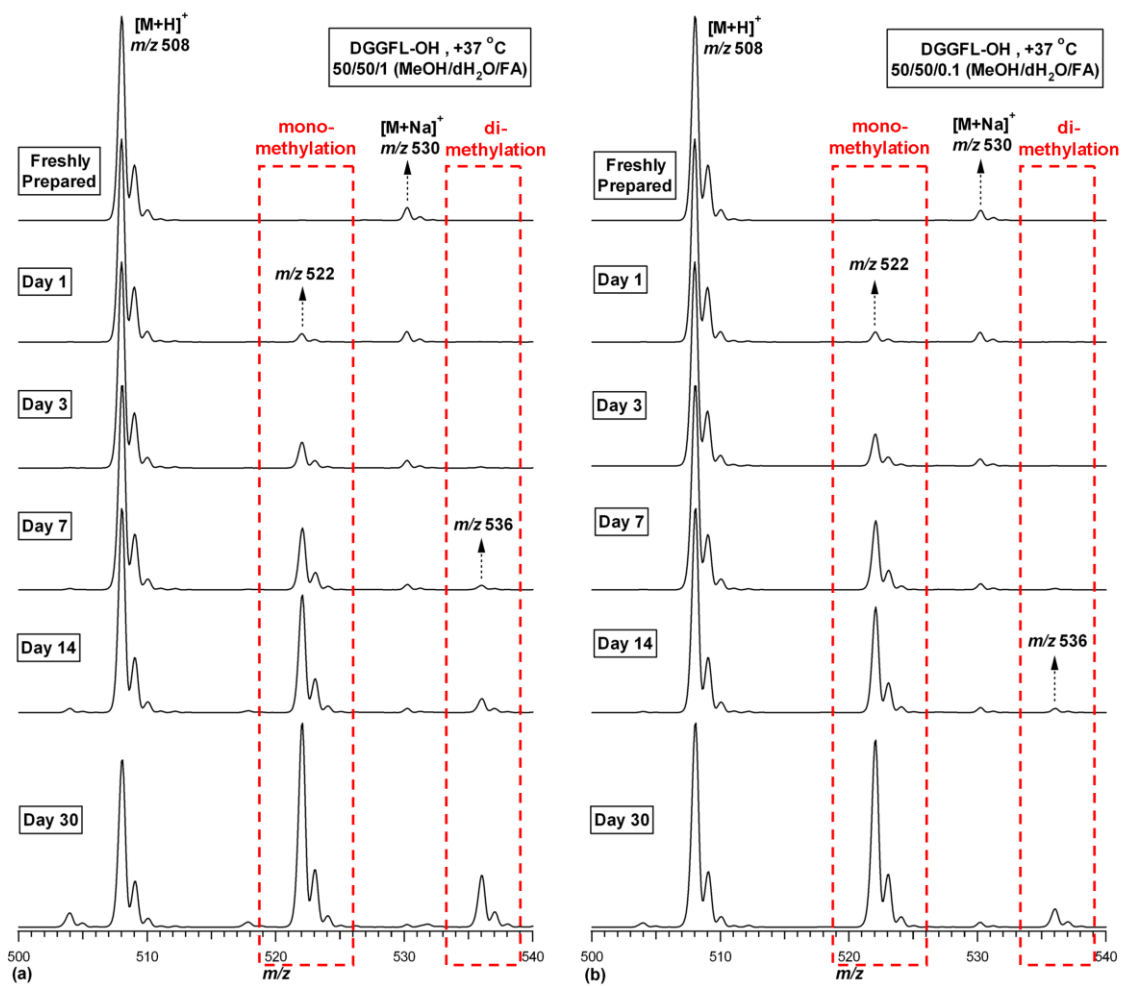


Figure C.11. ESI-MS spectra of DGGFL-OH prepared in (a) 50:50:1 (v/v/v) MeOH/dH₂O/FA or (b) 50:50:0.1 (v/v/v) MeOH/dH₂O/FA, incubated at 37 °C over a range of incubation times

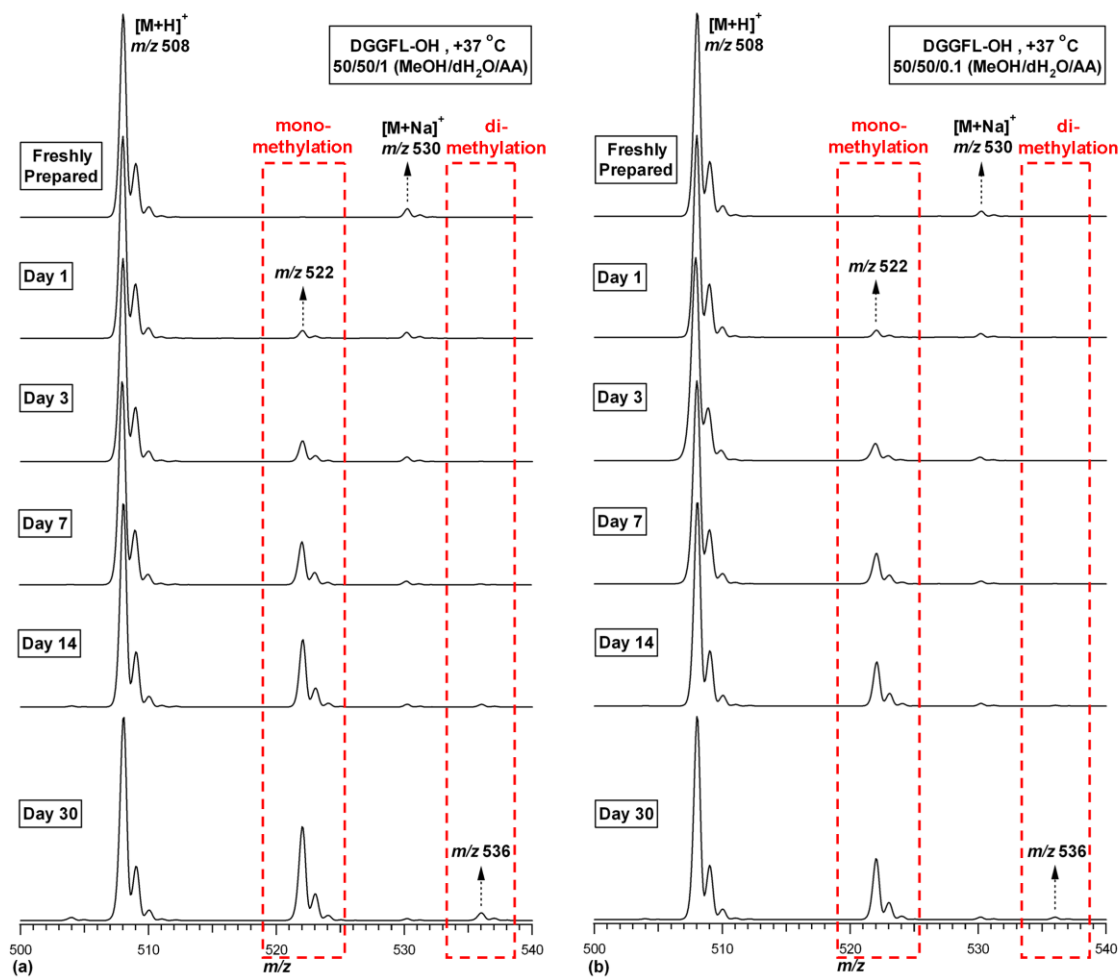


Figure C.12. ESI-MS spectra of DGGFL-OH prepared in (a) 50:50:1 (v/v/v) MeOH/dH₂O/AA or (b) 50:50:0.1 (v/v/v) MeOH/dH₂O/AA, incubated at 37 °C over a range of incubation times

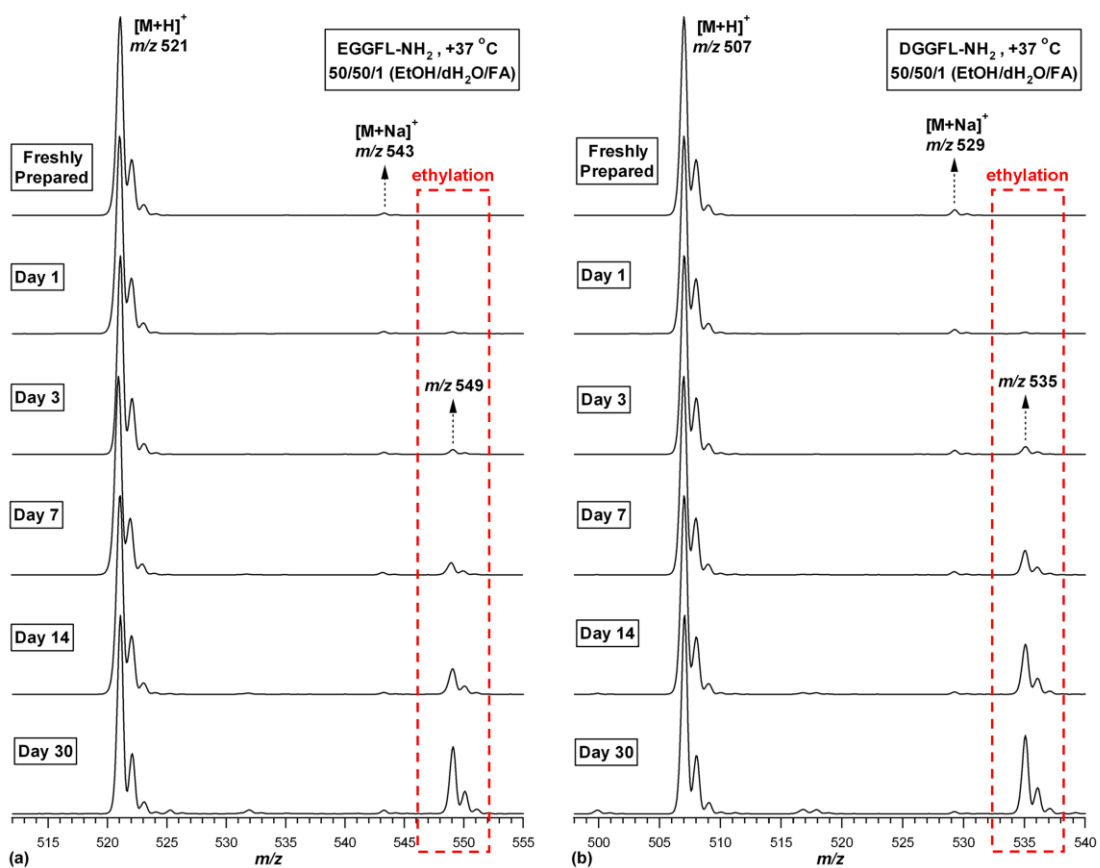


Figure C.13. ESI-MS spectra of (a) EGGFL-NH₂ and (b) DGGFL-NH₂ incubated at 37 °C over a range of incubation times. The peptide solutions were prepared in 50:50:1 (v/v/v) EtOH/dH₂O/FA

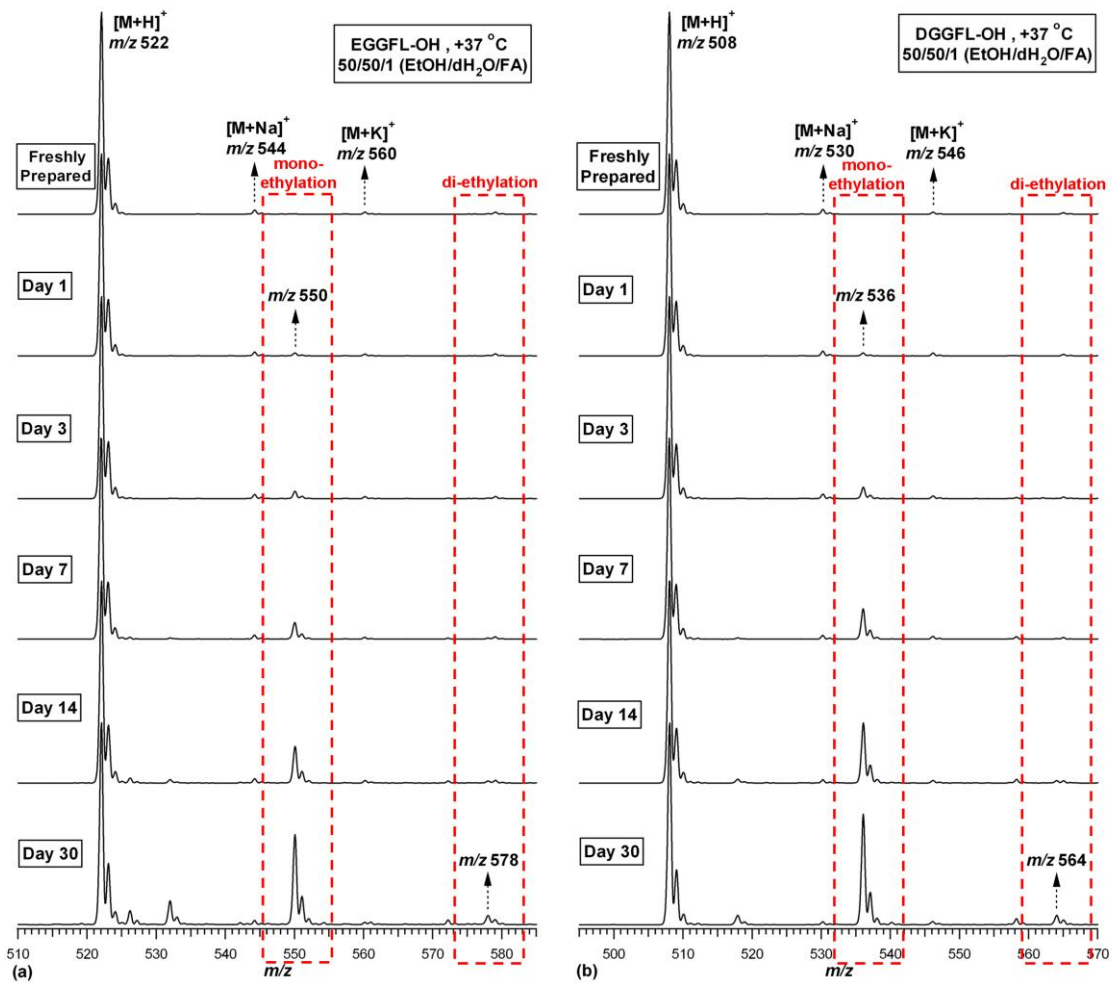


Figure C.14. ESI-MS spectra of (a) EGGFL-OH and (b) DGGFL-OH solutions incubated at 37 °C over a range of incubation times. The peptide solutions were prepared in 50:50:1 (v/v/v) EtOH/dH₂O/FA

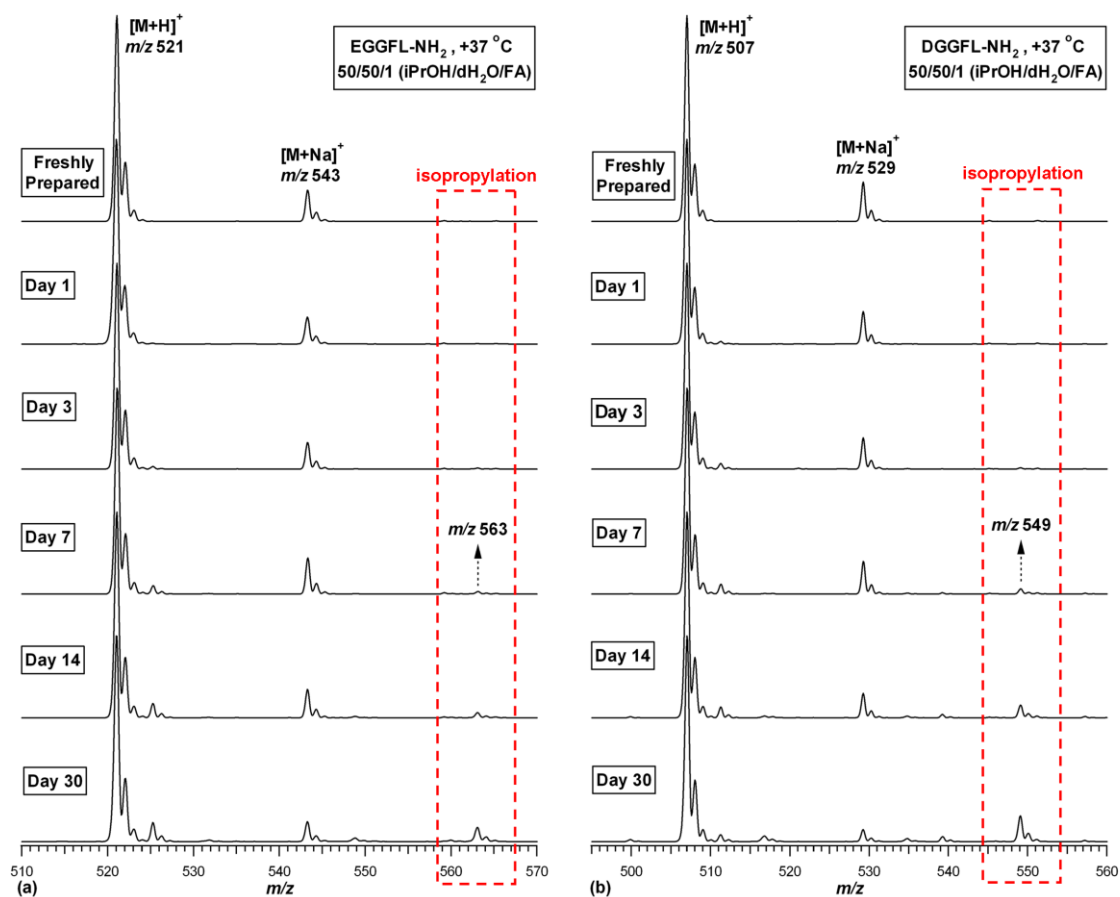


Figure C.15. ESI-MS spectra of (a) EGGFL-NH₂ and (b) DGGFL-NH₂ incubated at 37 °C over a range of incubation times. The peptide solutions were prepared in 50:50:1 (v/v/v) *i*PrOH/dH₂O/FA

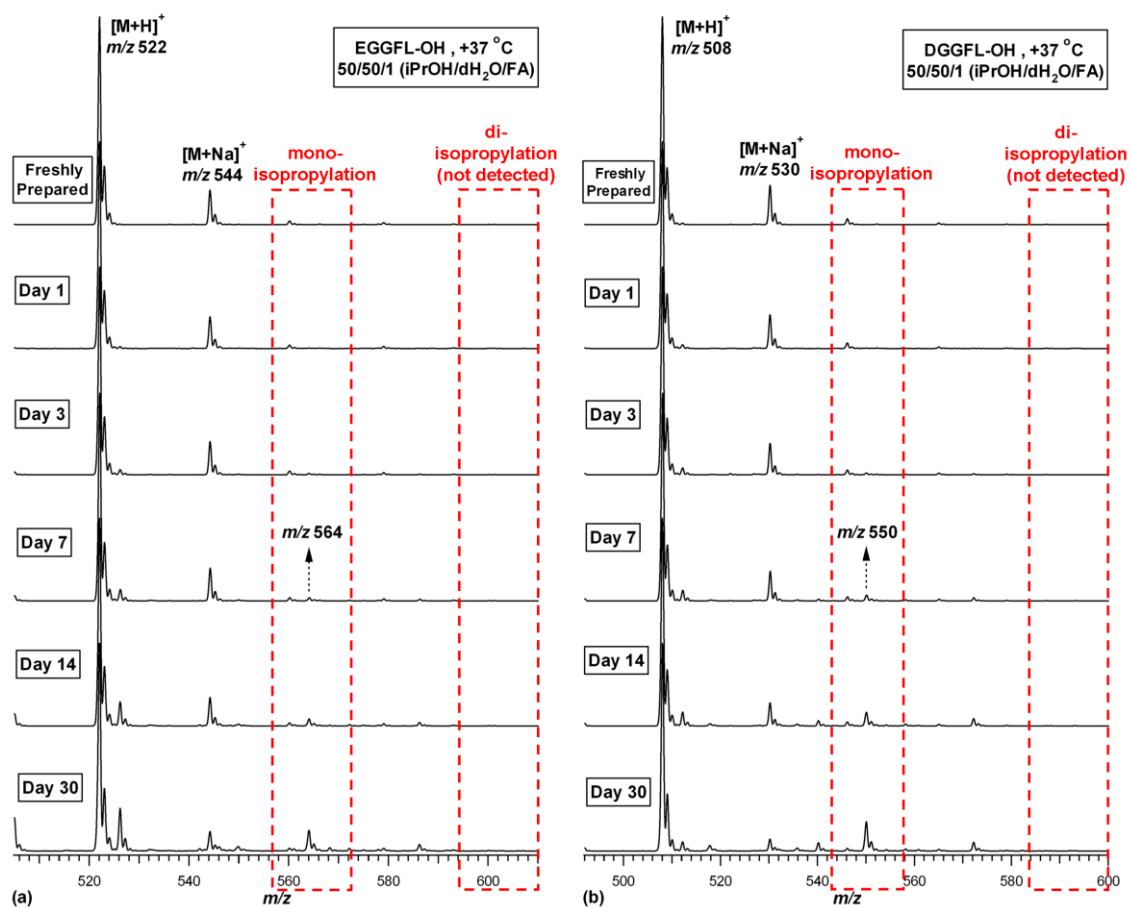


Figure C.16. ESI-MS spectra of (a) EGGFL-OH and (b) DGGFL-OH solutions incubated at 37 °C over a range of incubation times. The peptide solutions were prepared in 50:50:1 (v/v/v) *i*PrOH/dH₂O/FA

VITA

MELDA ZEYNEP GÜRAY

PERSONAL INFORMATION

Date of Birth: May 31, 1985
Nationality: Turkish
e-mail: mldguray@gmail.com

EDUCATION

- 2010-2017** **Doctor of Philosophy (PhD)**
İzmir Institute of Technology, Department of Chemistry
Dissertation Title: Proteomic Studies and It's Application to Biological Samples Using Mass Spectrometry-GPA: 3.79/4.00
Supervisor: Prof. Dr. Talat YALÇIN
- 2009-2010** **Scientific Competence**
İzmir Institute of Technology, Department of Chemistry
- 2007-2009** **Master of Science (MSc)**
İzmir Institute of Technology, Biotechnology and Bioengineering
Department Dissertation Title: Partial Purification and Characterization of Polyphenol Oxidase from Thermophilic *Bacillus* sp.-GPA: 3.79/4.00
Supervisor: Assoc. Prof. Gülşah ŞANLI-MOHAMED
- 2003-2007** **Bachelor of Science (BSc)**
Ege University, Department of Biochemistry
GPA: 71.9/100

WORK AND RESEARCH EXPERIENCES

- Biochemistry Specialist:** İzmir International Biomedicine and Genome Institute, İzmir-Turkey, April 2017-Present
- Visiting Scientist:** Dalhousie University, Department of Chemistry, Halifax-Canada, Feb 2016-June 2016
- Research and Teaching Assistant:** İzmir Institute of Technology, Department of Chemistry, İzmir-Turkey, Dec 2010-Sept 2016

PUBLICATIONS

3. **Guray, M.Z.**; Zheng, S.; Doucette, A.A. Mass Spectrometry of Intact Proteins Reveals +98 u Chemical Artifacts Following Precipitation in Acetone. *Journal of Proteome Research*. **2017**, 16 (2), 889-897.
2. Atik, A. E., **Guray, M.Z.**, Yalcin, T. Observation of the Side Chain *O*-Methylation of Glutamic Acid or Aspartic Acid Containing Model Peptides by Electrospray Ionization-Mass Spectrometry. *Journal of Chromatography B*. **2017**, 1047, 75-83.
1. **Guray, M. Z.**; Sanli-Mohamed, G. A New Thermophilic Polyphenol Oxidase from *Bacillus* sp.: Partial Purification and Biochemical Characterization. *Journal of Proteins and Proteomics*. **2013**, 4 (1), 11-20.