INVESTIGATION OF PROTEIN PRODUCTS UNDER DIFFERENT DIGESTION CONDITIONS

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

BY

DİLA BİLGEN

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

JANUARY 2024

Approval of the thesis:

INVESTIGATION OF PROTEIN DIGESTION PRODUCTS AND THEIR POTENTIAL IMPACT ON PROTEIN ANALYSIS

submitted by **DİLA BİLGEN** in partial fulfillment of the requirements for the degree of **Master of Science** i**n Chemistry, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, **Graduate School of Natural and Applied Sciences** _____

Prof. Dr. Özdemir Doğan
Head of the Department, **Chemistry** _____

Assoc. Prof. Dr. Süreyya Özcan Kabasakal
Supervisor, **Chemistry, METU** _____

Examining Committee Members:

Prof. Dr. Gülay Ertaş
Chemistry, METU _____

Assoc. Prof. Dr. Süreyya Özcan Kabasakal
Chemistry, METU _____

Assoc. Prof. Dr. Özgül Persil Çetinkol
Chemistry, METU _____

Assoc. Prof. Dr. Salih Özçubukçu
Chemistry, METU _____

Prof. Dr. Bekir Salih
Chemistry, Hacettepe University _____

Date: 25.01.2024

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

Name Last name : Dila Bilgen

Signature :

ABSTRACT

INVESTIGATION OF PROTEIN PRODUCTS UNDER DIFFERENT DIGESTION CONDITIONS

Bilgen, Dila
Master of Science, Chemistry
Supervisor: Assoc. Prof. Dr. Süreyya Özcan Kabasakal

January 2024, 168 pages

The proteomics field investigates proteins qualitatively and quantitatively. The peptide-centric approach has become a widely applied technique in clinical biomarker studies based on quantitative mass spectrometry analysis. The fragmentation of proteins into peptides is a crucial part of the peptide-centric approach. Proteins are fragmented into peptides via a digestion process, which involves the steps of dissolving, denaturing, reducing, alkylating, and proteolytically cleaving the protein. Biomarker studies apply these steps to the proteins extracted from several sources, such as human plasma and serum. The literature searched for digestion procedures for human plasma and serum showed that there are numerous methods, including different reagents, reagent concentrations, incubation temperatures, and durations. However, the literature search suggested no implications for the effects of varying digestion methods on protein products. Therefore, this study systematically investigates the effects of different digestion procedures employed in the literature on protein products. Initially, digestion procedures from the scanned articles were filtered according to their frequency of use, employed reagents and their concentrations, incubation temperatures, and durations. Ten methods representing the literature search were chosen accordingly.

Consequently, Alpha-2-Macroglobulin, a protein present in human blood, was digested through ten selected methods. Peptide derivatives obtained from different digestions were analyzed using a high-resolution mass spectrometer. Cleavage efficiency and artificial modification formation in peptide forms were compared to investigate the digestion efficiencies of ten selected methods. In addition, the effects of each step in the digestion procedure were compared. To accomplish this, conditions of a model protein, BSA, were investigated step by step using spectrophotometric methods. This study uncovered that different digestion procedures were not equivalent; their digestion efficiency differed, and they formed different artificial modifications on peptides. In conclusion, digestion procedures were found to be a variation source for clinical biomarker studies.

Keywords: Bottom-Up Proteomics, In-Solution Protein Digestion, Alpha-2-macroglobulin, Spectroscopic Methods, Artificial Modifications

ÖZ

FARKLI SİNDİRİM KOŞULLARINDA PROTEİN ÜRÜNLERİNİN İNCELENMESİ

Bilgen, Dila
Yüksek Lisans, Kimya
Tez Yöneticisi: Doç. Dr. Süreyya Özcan Kabasakal

Ocak 2024, 168 sayfa

Proteomik, proteinleri niteliksel ve niceliksel olarak araştırır. Aşağıdan yukarıya proteomik yaklaşımı klinik biyobelirteç çalışmalarında yaygın olarak kullanılmaktadır. Bu yaklaşım, proteinlerin çözündürülmesi, denatüre edilmesi, indirgenmesi, alkillenmesi ve proteolitik olarak parçalanması yoluyla proteinlerin peptitler halinde sindirilmesini içerir. Biyobelirteç çalışmaları bu adımları insan plazması ve serumu gibi çeşitli kaynaklardan elde edilen proteinlere uygular. Literatürde değişen reaktifler, reaktif konsantrasyonları, inkübasyon sıcaklıkları ve süreleri dahil olmak üzere birçok insan serumu ve plazma sindirim prosedürü vardır. Ancak değişen sindirim yöntemlerinin protein ürünleri üzerindeki etkilerine ilişkin bilgiler sınırlıydı. Bu çalışma, literatürdeki farklı sindirim prosedürlerinin özgün peptit seviyeleri üzerindeki etkilerini sistematik olarak araştırmaktadır. Başlangıçta, sindirim prosedürleri literatür aracılığıyla tarandı ve kullanım sıklıkları, reaktifler ve kullanılan konsantrasyonların yanı sıra inkübasyon sıcaklıkları ve enzimatik sindirim sürelerine göre filtrelendi. Literatür araştırmasını temsil eden on yöntem seçildi. Öncelikle parçalama prosedürünün etkileri spektrofotometrik yöntemler kullanılarak adım adım araştırıldı. Bu adımlar çözünürlük ve denatürasyon verimliliklerini, enzim aktivitelerini ve toplam peptit verimini içeriyordu. Daha

sonra, bir insan kan proteini olan Alfa-2-Makroglobulin, seçilen on yöntem kullanılarak sindirildi. Özgün peptitler ve bunların türevleri, yüksek çözünürlüklü bir kütle spektrometresi kullanılarak analiz edildi. Seçilen on yöntemin sindirim verimliliklerini araştırmak için peptit formlarında bölünme verimliliği ve yapay modifikasyon oluşumu karşılaştırıldı. Bu çalışma, farklı sindirim prosedürlerinin protein ürünleri üzerinde farklı verimlilik ve etkiler gösterdiğini ortaya çıkardı.

Anahtar Kelimeler: Aşağıdan Yukarıya Proteomik, Çözelti İçi Protein Parçalanması, Alfa-2-makroglobulin, Spektroskopik Yöntemler, Yapay Modifikasyonlar

to my Mother and

my angel little sister Feryal Özge…

ACKNOWLEDGMENTS

I would like to thank my supervisor, Assoc. Prof. Süreyya Özcan Kabasakal, for accepting me to her group when I needed it the most and reminding me that I love what I do. Also, for her support throughout my journey and for encouraging me to do better at every step while teaching me a new branch of science.

I would also like to thank my analytical chemistry professors, Prof. Dr. Gülay Ertaş and Assoc. Prof. Dr. Ezel Boyacı for their help and support.

I am fortunate to be a member of a group that is so benevolent and attentive as the analytical chemistry group. Also, I would like to thank my research assistant friends for their support.

I would like to thank all of my lab mates, my project partner Ceren Kul, my experiment buddy Lara Su Kartav, and software genius Enver Öztürk. I am very lucky to have as good friends as Ayşe Efran Boz, Aslı Nur Gökpınar, and Berkay Göçtürk. They have supported me whenever I needed them, listened to me even when I told the same story over and over again, and always made me feel welcome in their lives. My journey would have lacked color without them.

Last but not least, I would like to thank my lovely mother for raising me to become the strong and independent woman I am today and for always being there for me. I am very lucky to have my grandmother, Aysın Bilgen, and my aunts and my uncle, also my babies Herkül and Lady.

TABLE OF CONTENTS

ABSTRACT v

ÖZ vii

ACKNOWLEDGMENTS x

TABLE OF CONTENTS xi

LIST OF TABLES xv

LIST OF FIGURES xvii

LIST OF ABBREVIATIONS xxiii

CHAPTERS

1 INTRODUCTION 1

1.1 Proteins 1

1.1.1 Protein Modifications 1

1.2 Proteomics 4

1.2.1 Clinical Proteomics & Protein Biomarkers 5

1.2.2 Plasma & Serum 6

1.3 Clinical Proteomics-Based Protein Analysis 7

1.3.1 Protein Analysis Techniques in Clinical Proteomics 7

1.3.1.1 Absorption Spectroscopy Analysis in Proteomics 9

1.3.1.2 Circular Dichroism Spectroscopy Analysis in Proteomics 10

1.3.2 Colorimetric Assay Analyses in Proteomics 12

1.3.3 Mass Spectrometry Analysis in Proteomics 14

1.3.4 Database Search and Protein Identification 17
1.3.5 Peptide Centric Approach 18
1.3.6 Sample Preparation for Bottom-Up Proteomics 19
1.3.6.1 Denaturation Process 20
1.3.6.2 Reduction Process 21
1.3.6.3 Alkylation Process 22
1.3.6.4 Proteolytic Digestion Process 23
1.4 Model Proteins 25
1.4.1 Bovine Serum Albumin (BSA) 25
1.4.2 Alpha-2-Macroglobulin (A2MG) 27
1.5 Motivation 28
2 METHODS AND MATERIALS 29
2.1 Materials 29
2.2 Methods 30
2.2.1 Literature Search 30
2.2.2 Digestion Procedures 30
2.2.2.1 Method A 31
2.2.2.2 Method B 32
2.2.2.3 Method C 32
2.2.2.4 Method D 33
2.2.2.5 Method E 34
2.2.2.6 Method F 34
2.2.2.7 Method G 35

2.2.2.8 Method H 36
2.2.2.9 Method I 36
2.2.2.10 Method J 37
2.2.3 Absorption Spectroscopy Analysis 38
2.2.4 Quality Control Samples for BCA Assay & Peptide Assay 38
2.2.5 BCA Assay Procedure 40
2.2.6 Circular Dichroism Spectra Procedure 41
2.2.7 Enzyme Activity Assay Procedure 41
2.2.8 Peptide Assay Procedure 42
2.2.9 Mass Spectrometry based Protein/Peptide Analysis 42
2.2.10 Data Analysis 44
3 RESULTS AND DISCUSSION 45
3.1 Workflow 45
3.2 Literature Search and Method Selection 48
3.3 Investigation of Digestion Method 59
3.3.1 Native Protein Analysis/ Solubility Comparison 59
3.3.1.1 Solubility Comparison via Absorption Spectroscopy 59
3.3.1.2 Solubility Comparison via BCA Analysis 64
3.3.1.2.1 QC Sets for BCA Assay 64
3.3.1.2.1.1 QC-BCA Set 64
3.3.1.2.1.2 QC-Δ Set 65
3.3.1.2.2 Quantitative Solubility Comparison 68
3.3.2 Denaturation Efficiency Comparison 71

3.3.3 Enzyme Activity Analysis 76
3.3.4 Total Peptide Analysis 79
3.3.4.1 QC Sets for Peptide Assay 79
3.3.4.1.1 QC-PA Set 79
3.3.4.1.2 QC-Reverse Set 81
3.3.4.2 Quantitative Total Peptide Concentration Comparison 83
3.4 Investigation of Digestion Products 88
3.4.1 Protein Groups 88
3.4.2 Peptide Groups 90
3.4.2.1 Unique Peptide Distribution and Cleavage Efficiencies 91
3.4.2.2 Artificial Modifications 96
4 CONCLUSION 117
REFERENCES 121
APPENDICES 153
A. Calibration Curves for BCA Assay 153
B. Total Ion Chromatogram of A2MG Sample Digests 157

LIST OF TABLES

TABLES

Table 1.1. Artificial Modification Types and the Amino Acid Sites. 3
Table 1.2. Examples of Protein Biomarkers generated from List of Qualified Biomarkers Table in U.S. Food and Drug Administration (FDA) 5
Table 2.1. MaxQuant Parameters 44
Table 2.2. Proteome Discoverer Parameters 44
Table 3.1 Literature Search for In-Solution Digestion Products with Human Blood/ Plasma /Serum. 50
Table 3.2. Selected Methods from Literature. Ten methods were coded Method A to J for ease of reference. 56
Table 3.3. $\%T_{max}$ and $\%T_{min}$ values of six BSA replicates for each buffer and deionized water mediums at 277 nm. Difference between these values was given under $\%T_{max-min}$. $\%T_{average}$ values are calculated by averaging the percent transmittance values of each replicate in respective buffer medium. 62
Table 3.4. Quality Control Set for Investigation of BCA Assay (QC-BCA) 65
Table 3.5. Quality Control Set for Investigating BCA Assay Results Before and After Thermal Incubation (QC- Δ) 66
Table 3.6. Normalized Values for QC- Δ. 67
Table 3.7. LOD and LOQ values of calibration curves in buffers of selected methods and DI 69
Table 3.8. Concentrations of experimental replicates of BSA proteins 70
Table 3.9. Quality Control Set for Investigation of Peptide Assay (QC-PA) 80
Table 3.10. Quality Control Set for Investigating Peptide Assay Results in a Complete Protein Environment (QC-Reverse). 81
Table 3.11. Normalization values of QC-Reverse Samples 82
Table 3.12. Yielded Peptide Concentrations of Different Digestion Methods Obtained from Equivalent Proteins 84

Table 3.13. Protein groups results of A2MG samples digested with Methods A to J.....89
Table 3.14. Unique Peptide List of A2MG Protein obtained from UniProt BLAST database. 91
Table 3.15. Monoisotopic mass values in Da for unmodified DMYSFLEDMGLK peptide sequence.....100
Table 3.16. Monoisotopic mass values in Da for DMYSFLEDMGLK peptide which was dethiometylated on M2.....103
Table 3.17. Monoisotopic mass values in Da for DMYSFLEDMGLK peptide which was carbamylated on N-terminal and K12 106
Table 3.18. Monoisotopic mass values in Da for DMYSFLEDMGLK peptide which was oxidized on M9.....109
Table 3.19. Monoisotopic mass values in Da for DMYSFLEDMGLK peptide which was carbamidomethylated on N-terminal, dethiometylated on M2 and M9 112

LIST OF FIGURES

FIGURES

Figure 1.1. Branches of the Proteomics Field[17] 4
Figure 1.2. Common Proteins in Human Plasma and Their Percentage Values in Human Plasma[25] 7
Figure 1.3. Amino Acids with Aromatic Side Chains 9
Figure 1.4. Absorption Spectra of Aromatic Amino Acids. 10
Figure 1.5. Circular Dichroism Spectrum of Secondary Structures of Proteins[61,63] 11
Figure 1.6. Reduction of cupric ion by protein to cuprous ion and oxidation of protein as a result of this reduction also known as biuret reaction. 12
Figure 1.7. Chelation Reaction of Bicinchoninic Acid Disodium Salt and Cuprous Ion in BCA Assay[73] 13
Figure 1.8. Chelation Reaction of Bathocuproine Disulfonate (BCS) Salt and Cuprous Ion in BCA Assay[76] 14
Figure 1.9. General Block Diagram of a LC/MS Instrument 15
Figure 1.10. MS Analysis vs MS/MS Analysis. 16
Figure 1.11. Software Match of *in silico* MS/MS Spectra with Experimental Results to Yield Qualitative and Quantitative Results 17
Figure 1.12. Schemes of Top-Down and Bottom-Up Approaches in Proteomics.. 18
Figure 1.13. In-Solution Digestion Procedure Steps 20
Figure 1.14. Reduction of Disulfide Bonds with DTT[111] 21
Figure 1.15. Reduction of Disulfide Bonds with TCEP[110] 22
Figure 1.16. Alkylation of reduced Cys Residues with IAA or CAA[110] 22
Figure 1.17. Trypsin Enzyme Structure with Catalytic Triad (Asp102+His57+Ser195) Highlighted[116] 23
Figure 1.18. Reaction Mechanism of Cleavage Site of the Trypsin[114,118] 24
Figure 1.19. 3D Structure of BSA Protein 26

Figure 1.20. (a) CD Spectrum and (b) Absorption Spectrum of pure BSA. 26
Figure 1.21. 3D Structure of A2MG Protein 27
Figure 2.1. Digestion steps each QC represents. 39
Figure 3.1. Workflow of this Study 47
Figure 3.2. Workflow of this Study 59
Figure 3.3. Absorption Spectrum of six experimental replicates of native BSA 60
Figure 3.4. Average absorbance values of experimental replicates were calculated to compare BSA protein solubilities in selected digestion buffers. 61
Figure 3.5. Equation to calculate transmittance percent (%T) from absorbance (A) 62
Figure 3.6. Student's t-test for calculating differences in means of two independent groups[212] 67
Figure 3.7. Limit of detection (LOD) and Limit of linearity (LOL) equations 68
Figure 3.8. Concentrations of BSA Protein obtained from the Native Protein Analysis 69
Figure 3.9. Native and denaturated & reduced protein sample peaks in CD spectra according to Methods A to E. 72
Figure 3.10. Native and denaturated & reduced protein sample peaks in CD spectra according to Methods F to J. 73
Figure 3.11. Comparison of CD Spectra of Denaturated and Reduced BSA Proteins according to methods and Native BCA Protein 74
Figure 3.12. Calibration Curve of Enzyme Activity Kit measured at 405 nm 76
Figure 3.13. Trypsin Activity in samples prepared by Methods A-J 77
Figure 3.14. Final p-NA Concentration 78
Figure 3.15. Peptide Assay Calibration Curve measured at 480 nm 83
Figure 3.16. Yielded Peptide Concentrations as the result of Different Digestion Methods Obtained from Equivalent Proteins 85

Figure 3.17. Varying Native Protein Concentrations due to Solubility Difference in different buffers associated with different Methods and Yielded Peptide Concentrations 85
Figure 3.18. Normalized Native Protein and Yielded Peptide Concentrations 86
Figure 3.19. Unique Peptide Ratios Every Method Yield and Their Miscleaved Forms (Given in The Outer Ring of The Diagram) 92
Figure 3.20. Frequency of use of unique peptides in studies on human plasma/ serum 93
Figure 3.21. Chromatogram and mass spectrum for unmodified DMYSFLEDMGLK peptide 98
Figure 3.22. MS/MS spectrum for unmodified DMYSFLEDMGLK peptide..... 99
Figure 3.23. Chromatogram and mass spectrum of DMYSFLEDMGLK peptide which was dethiometylated on M2 101
Figure 3.24. MS/MS spectrum of DMYSFLEDMGLK peptide which was dethiometylated on M2 102
Figure 3.25. Chromatogram and mass spectrum of DMYSFLEDMGLK peptide which was carbamylated on N-terminal and K12 104
Figure 3.26. MS/MS spectrum of DMYSFLEDMGLK peptide which was carbamylated on N-terminal and K12 105
Figure 3.27. Chromatogram and mass spectrum of DMYSFLEDMGLK peptide which was oxidized on M9 107
Figure 3.28. MS/MS spectrum of DMYSFLEDMGLK peptide which was oxidized on M9 108
Figure 3.29. Chromatogram and mass spectrum of DMYSFLEDMGLK peptide which was carbamidomethylated on N-terminal, dethiometylated on M2 and M9 110
Figure 3.30. MS/MS spectrum of DMYSFLEDMGLK peptide which was carbamidomethylated on N-terminal, dethiometylated on M2 and M9 111

Figure 3.31. Modification Types and Ratios Form on VGFYESDVMGR, DMYSFLEDMGLK, LLIYAVLPTGDVIGDSAK, LVHVEEPHTETVR, and QFSFPLSSEPFQGSYK 114
Figure A.1. BCA Assay Calibration Curve (QC-DI-BCA) for Native BSA in Deionized Water. Absorption values are measured at 562 nm. 153
Figure A.2. BCA Assay Calibration Curve for Method A in 50 mM ABC Buffer Absorption values are measured at 562 nm. 153
Figure A.3. BCA Assay Calibration Curve for Method B, Method C, and Method J in100 mM Tris Buffer. Absorption values are measured at 562 nm. 154
Figure A.4. BCA Assay Calibration Curve for Method D and Method G in 25 mM ABC Buffer. Absorption values are measured at 562 nm. 154
Figure A.5. BCA Assay Calibration Curve for Method E in 100 mM ABC Buffer Absorption values are measured at 562 nm. 155
Figure A.6. BCA Assay Calibration Curve for Method F with 10 mM Tris Buffer. Absorption values are measured at 562 nm. 155
Figure A.7. BCA Assay Calibration Curve for Method H with 300 mM Tris Buffer. Absorption values are measured at 562 nm. 156
Figure A.8. BCA Assay Calibration Curve for Method I with 200 mM Tris Buffer. Absorption values are measured at 562 nm. 156
Figure B.9. Total Ion Chromatography of A2MG Protein Digested with Method A – Replicate 1 157
Figure B.10. Total Ion Chromatography of A2MG Protein Digested with Method A – Replicate 2 157
Figure B.11. Base Peak Spectra of A2MG Protein Digested with Method A – Replicate 1 158
Figure B.12. Base Peak Spectra of A2MG Protein Digested with Method A – Replicate 2 158

Figure B.13. Total Ion Chromatography of A2MG Protein Digested with Method B – Replicate 1 159
Figure B.14. Total Ion Chromatography of A2MG Protein Digested with Method B – Replicate 2 159
Figure B.15. Base Peak Spectra of A2MG Protein Digested with Method B – Replicate 1 160
Figure B.16. Base Peak Spectra of A2MG Protein Digested with Method B – Replicate 2 160
Figure B.17. Total Ion Chromatography of A2MG Protein Digested with Method C – Replicate 1 161
Figure B.18. Total Ion Chromatography of A2MG Protein Digested with Method C – Replicate 2 161
Figure B.19. Base Peak Spectra of A2MG Protein Digested with Method C – Replicate 1 162
Figure B.20. Base Peak Spectra of A2MG Protein Digested with Method C – Replicate 2 162
Figure B.21. Total Ion Chromatography of A2MG Protein Digested with Method F – Replicate 1 163
Figure B.22. Total Ion Chromatography of A2MG Protein Digested with Method F – Replicate 2 163
Figure B.23. Base Peak Spectra of A2MG Protein Digested with Method F – Replicate 1 164
Figure B.24. Base Peak Spectra of A2MG Protein Digested with Method F – Replicate 2 164
Figure B.25. Total Ion Chromatography of A2MG Protein Digested with Method G – Replicate 1 165
Figure B.26. Total Ion Chromatography of A2MG Protein Digested with Method G – Replicate 2 165

Figure B.27. Base Peak Spectra of A2MG Protein Digested with Method G – Replicate 1 166
Figure B.28. Base Peak Spectra of A2MG Protein Digested with Method G – Replicate 2 166
Figure B.29. Total Ion Chromatography of A2MG Protein Digested with Method J – Replicate 1 167
Figure B.30. Total Ion Chromatography of A2MG Protein Digested with Method J – Replicate 2 167
Figure B.31. Base Peak Spectra of A2MG Protein Digested with Method J – Replicate 1 168
Figure B.32. Base Peak Spectra of A2MG Protein Digested with Method J – Replicate 2 168

LIST OF ABBREVIATIONS

ABBREVIATIONS	
A2MG	Alpha-2-macroglobulin
ABC	Ammonium bicarbonate
ACN	Acetonitrile
BAPNA	Nα-Benzoyl-DL-arginine 4-nitroanilide Hydrochloride
BCA	Bicinchoninic Acid
BCS	Bathocuproine Disulfonate
BSA	Bovine Serum Albumin
CAA	Chloroacetamide
CV	Coefficient of Variation
CD	Circular Dichroism Spectroscopy
DI	Deionized Water
DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
FASP	Filter Aided Sample Preparation
IAA	Iodoacetamide
LC	Liquid chromatography
LC-MS	Liquid chromatography-mass spectrometer
LOD	Limit of detection

LOQ	Limit of quantification
LTQ	Linear ion trap
m/z	Mass-to-charge ratio
MS	Mass spectrometry
MS/MS	Tandem Mass Spectrometry
PD	Proteome Discoverer
p-NA	p-Nitroaniline
QC	Quality Control
QQQ-MS	Triple Quadruple Mass Spectrometer
TCEP	Tris(2-carboxyethyl)phosphine Hydrochloride
TRIS	Tris(hydroxymethyl)aminomethane
UV-Vis	Ultraviolet-Visible

CHAPTER 1

INTRODUCTION

1.1 Proteins

Proteins are biopolymeric macromolecules composed of amino acids. Their structural information is categorized as primary, secondary, tertiary, and quaternary.[1] Proteins play essential roles in the function of the cells, biological pathways, regulatory processes, and disease mechanisms. They are also one of the main targets for therapeutic drugs.[2] The entirety of the proteins that are expressed by an organism is called proteome. Proteome also refers to the set of proteins an organism produces at a particular time. It is an expression of the genome. Genome is the complete gene set of an organism. The finalization of the human genome project showed that the genome itself was not enough to fully understand organisms.[3] This was due to stability of it throughout lifetime. Rather, dynamic information sources are needed in addition to stable information source that is genome. Proteome was found to be a great source for understanding organisms due to its great variation depending on the development stages of the organism, environmental exposures, and disease processes. The expression variation during therapeutic treatments was also considered to be a major benefit for clinical studies.[4]

1.1.1 Protein Modifications

Proteins undergo alterations upon translation according to their functions and activities by the organisms.[5] These alterations are called post-translational modifications (PTMs) and they give rise to a vast array of proteoforms, posing a challenge in the majority of proteomics studies.[6] Also, reagents used during sample

processing/digestion and storage processes may create additional modifications on residues, called artifact modifications. It is given in the literature that different reagents may cause varying artifact modifications.[7] Examples to artificial modifications are given in Table 1.1. Table shows which artificial modifications are occurring which residues or terminals. "UniMod: Protein Modifications for Mass Spectrometry" database has been referenced.[8] It is seen that artificial modifications are not seen on the residues with aliphatic side chains which are G, A, I, L, P, and V. Other residues are observed to undergo one or multiple artificial modifications. Most reactive sites were observed to be M, K and N-term. Artificial modifications occur as a result of the reagent used, temperature and pH employed in the digestion process. The extent of chemical alterations is correlated with the incubation time in the reduction and alkylation buffer. Additionally, exposure of the residues to the buffer affects the rate of these artifact modifications.[9] It is known in the literature that urea causes carbamylation of N-Term Lys, and Arg[10,11] which is an artifact modification itself. Further, it is seen that digestion processes using ammonium bicarbonate (ABC) buffer rather than tris buffer showed less carbamylation modification.[12] These modifications affect the quantification of the peptides in bottom-up proteomics since they create mass shifts during the MS and MS/MS analyses. However, including all of the modifications during data analysis is not a solution since it would greatly increase analysis time. Most importantly, it would increase the false-positive identifications.[13]

Table 1.1. Artificial Modification Types and the Amino Acid Sites.
The Table displays artificial modifications and residues or terminals they occur on with "+" sign. "-" sign means that given artificial modification is not seen on corresponding residue/ terminal site. The "Unimod: Protein Modifications for Mass Spectrometry" database has been referenced. Unimod Accesion Numbers and Mass Shift values has been provided for each modification. Artificial modifications are absent on residues with aliphatic side chains, namely G, A, I, L, P, and V. Additional residues are observed to undergo one or multiple synthetic alterations. The most reactive sites that were observed were the Lysine residue (K) and the N-terminal (N-term).

Artificial Modifications	Unimod Accession #	Mass Shift (Da)	Residue/ Terminal Site															
			C	D	E	F	H	K	M	N	Q	R	S	T	W	Y	C-Term	N-Term
Acetylation	1	42.010565	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	+
Amidation	2	-0.984016	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
Carbamidomethylation	4	57.021464	-	+	+	-	+	+	-	-	-	-	+	+	-	+	-	+
Carbamylation	5	43.024700	+	-	-	-	-	+	+	-	-	+	-	-	-	-	-	+
Carboxymethylation	6	58.005479	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+
Deamidation	7	0.984016	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
Cation:Na	30	21.981943	-	+	+	-	-	-	-	-	-	-	-	-	-	-	+	-
Oxidation	35	15.994915	-	-	-	+	+	-	+	-	-	-	-	-	+	-	-	-
Methylthiolation	39	45.987721	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+
Formylation	122	27.994915	-	-	-	-	-	+	-	-	-	-	+	+	-	-	-	+
Iodination	129	125.896648	-	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-
di-Iodination	130	251.793296	-	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-
tri-Iodination	131	377.689944	-	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-
Methylamine	337	13.031634	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-
Trp->Hydroxykynurenin	350	19.989829	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
Trp->Kynurenin	351	3.994915	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
Dethiomethylation	526	-48.003371	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
Cation:K	530	37.955882	-	+	+	-	-	-	-	-	-	-	-	-	-	-	+	-

1.2 Proteomics

The research area that investigates the proteome is called proteomics, and the term is derived from the combination of "proteome" and "omics." The term omics refers to the utilization of informatics and technologies for obtaining and processing extensive biological data sets.[14] Proteomics term was first introduced by Marc Wilkins in 1994.[15] This field refers to the utilization of technological advancements to detect and quantify the total number of proteins present in a given organism, cell, bodily fluids, or tissue.[16] Branches of proteomics are given in Figure 1.1. It can be seen that there are multiple subcategories and one of them is biomarker detection. Biomarkers are crucial in the identification of diseases and monitoring the progress of treatment. It is crucial to acknowledge that diseases in the initial phase are usually treated with a high chance of success.

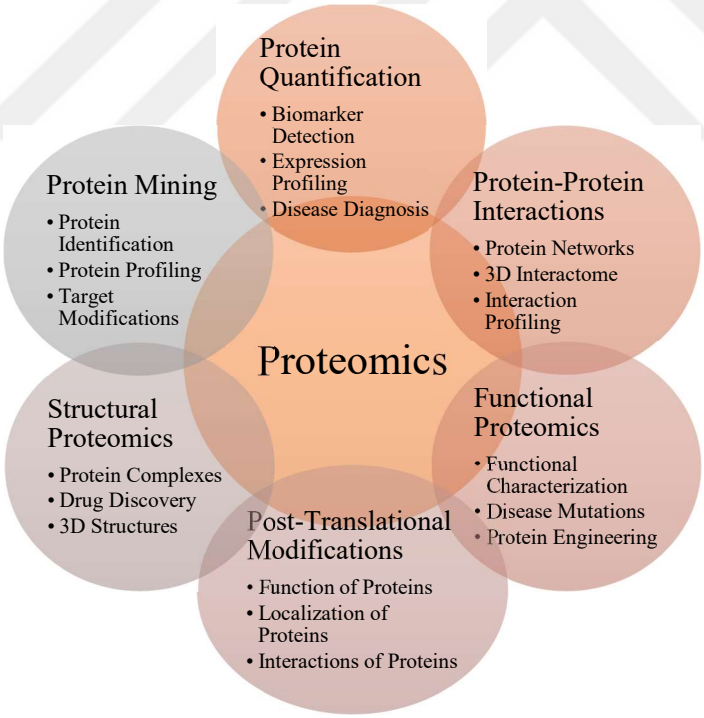


Figure 1.1. Branches of the Proteomics Field[17]

1.2.1 Clinical Proteomics & Protein Biomarkers

Successful treatment of a wide number of diseases requires understanding of the dynamic protein pathways involved in both healthy and diseased states, as well as their response to medical therapy.[18] Clinical proteomics studies diligently utilize proteomic technologies and bioinformatics to identify protein patterns associated with diseases. These disease-associated proteins are examples of biomarkers and some of them are given in Table 1.2.[19]

Table 1.2. Examples of Protein Biomarkers generated from List of Qualified Biomarkers Table in U.S. Food and Drug Administration (FDA)

Biomarker	Related Diseases
Albumin	Liver or Kidney Disease, Malnutrition, or Inflammation
β2- Microglobulin	Chronic Lymphocytic Leukemia, Multiple Myeloma, or Kidney Disease
Insulin	Diabetes
C-reactive protein (CRP)	Inflammatory Conditions
Clusterin	Cancers, Neurodegenerative Diseases
Prostate-specific antigen (PSA)	Prostate Cancer
Cystatin C	Impaired Kidney Function
Trefoil factor-3	Inflammatory Bowel Disease and Colorectal Cancer
Human Epidermal Growth Factor Receptor 2 (HER2)	Breast Cancer
Alpha-2-macroglobulin (A2MG)	Alzheimer's Disease, Parkinson's Disease, Liver and Kidney Diseases, Inflammatory Conditions, Cancer
Fibrinogen	Chronic Obstructive Pulmonary Disease
N-acetyl-beta-D-glucosaminidase (NAG)	Kidney Function and Renal Damage
Neutrophil Gelatinase-Associated Lipocalin (NGAL)	Acute Kidney Injury, Cardiovascular Disease and Sepsis
Osteopontin (OPN)	Neurological Disorders, Autoimmune Diseases, Cancer

In general, biomarker is a biological entity that can be measured as indicator of biological and pathogenic activities in a given therapeutic process.[20] Biomarkers should exhibit minimal to no variability in their measurements, possess a considerable signal-to-noise ratio, and demonstrate rapid and dependable changes in response to modifications in the disease or its treatment.[21] Proteins are the most abundantly influenced biological entities from diseases, treatment, and recovery processes.[22] Hence, clinical proteomics studies show great potential to discover biomarkers.[23] These studies also enable the selection of personalized therapies and monitoring treatment response with samples collected from bodily fluids like blood or urine.[24]

1.2.2 Plasma & Serum

Blood is in direct relation to most of our body hence making it dynamic and comprehensive in terms of information collection.[25] It can provide vital information on a systemic level and is sensitive to diseased or healthy states, genetic variances, environmental influences, lifestyle choices, dietary habits, and medications.[26] Hence, blood samples are frequently used in the diagnosis processes in medicine.[27] Plasma and serum samples obtained from blood are two of the utilization kinds of the blood. Plasma refers to the liquid base of the blood in which the solid portion comprises red blood cells, white blood cells, and platelets. Serum is obtained when clotting factors are removed from the plasma. Serum consists of proteins, enzymes, hormones, electrolytes, and vitamins.[28] Plasma/serum proteomics carries great potential for biomarker studies due to being easily accessible and involving most of the proteins. The protein concentration of the human plasma and serum is 60-80 mg/ml[29], and there are approximately 780[30,31] different proteins in them. However, twenty-two of these proteins constitute 99% of the total proteins in the plasma/serum.[32] These proteins are given in Figure 1.2.

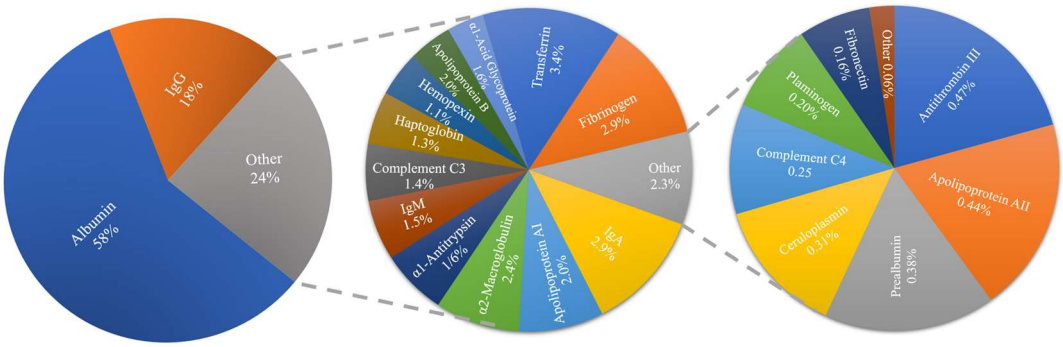


Figure 1.2. Common Proteins in Human Plasma and Their Percentage Values in Human Plasma[25]

Albumin is the protein that has highest abundance in the blood with 30-45 mg/ml concentration value.[33] It is one of the protein biomarkers as given in Table 1.2. A2MG and Fibrinogen given in Table 1.2 are also seen as abundant proteins in the blood as seen in Figure 1.2.

1.3 Clinical Proteomics-Based Protein Analysis

1.3.1 Protein Analysis Techniques in Clinical Proteomics

Proteomics is utilized in various fields, from nutritional studies[34,35] to space studies[36,37] to clinical studies.[38,39] There are several analysis methods that yield qualitative and quantitative results. Each analysis method has their advantageous and limitations.[40] Qualitative methods including UV-Vis[41] and Circular Dichroism[42] spectroscopy methods involve the interaction of light with residues and the secondary structure of the protein. Quantitative colorimetric protein quantification methods are quick, easy, and non-laborious methods.[43] And they are used for obtaining overall protein concentration in the samples. There are four commonly

used colorimetric assays in the literature: Biuret Assay, Lowry Assay, Bicinchoninic Acid (BCA) Assay, and Bradford Assay. Edman Degradation[44], Green Fluorescent Protein tagging[45], NMR spectroscopy[46], and X-ray crystallography[47] are other useful protein analysis tools. Enzyme-linked Immunosorbent Assay (ELISA) and Western Blotting are widely employed biochemistry methods in proteomics. The ELISA is a highly sensitive immunoassay that is extensively utilized for diagnostic purposes. However, it can be utilized for limited proteins[16,48], is open to cross-reactivity[49], and shows variations among batches.[50] Another analysis technique is Western Blotting and it can be used for identifying a single protein in a mixture of proteins.[51] Mass spectrometry-based proteomics is increasingly utilized in the identification of biomarkers for the early detection, prognosis, prediction, and monitoring of treatment response.[52,53] MS analysis can be used in identification as well as quantitation. It is a highly sensitive and accurate method that provides structural information. Additionally, there are methods with MS analysis that may target one specific protein in a mixture which is useful for samples like blood which have a complicated matrix and many proteins. Details are given below for selected protein analysis tools.

1.3.1.1 Absorption Spectroscopy Analysis in Proteomics

Absorption spectroscopy is a quick and easy qualitative protein analysis method since no additional reagents or incubation is involved in the sample preparation. However, it is open to interferences by buffers, detergents, and conditions of pH, polarity, and ionic strength of the environment.[54]

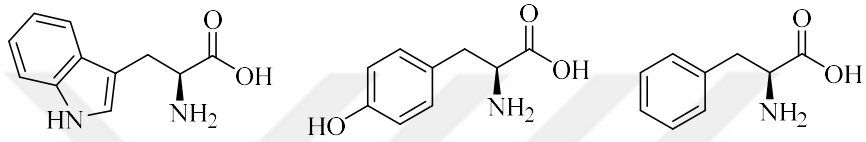


Figure 1.3. Amino Acids with Aromatic Side Chains

Several functional groups are contributing to the UV absorption of the proteins as given in Figure 1.3. Peptide bonds absorb between 180-230 nm due to $\pi \rightarrow \pi*$ transitions. Additionally, disulfide bonds of cystine residues make a weak absorption near 260 nm. Absorption between 230-300 nm is caused by residues containing aromatic side chains. Absorption is made at 280 nm by tryptophan with indole group, 275 nm by tyrosine with phenol group, and 257 nm by phenylalanine with phenyl group.[55–57] At the same concentration, tryptophan absorption at ~280 nm is greater in magnitude than tyrosine or phenylalanine absorption, as given in Figure 1.4. In addition, the weak absorption of the peptide bond and sulfide bond, makes them less effective for protein analysis.[55,58] Hence, absorbance signal at ~280 nm is widely used for protein analysis in the literature.[59] Additionally, it is possible to measure the concentration of proteins that lack aromatic residues with the absorbance signal at 205 nm[60] due to peptide bonds.[41]

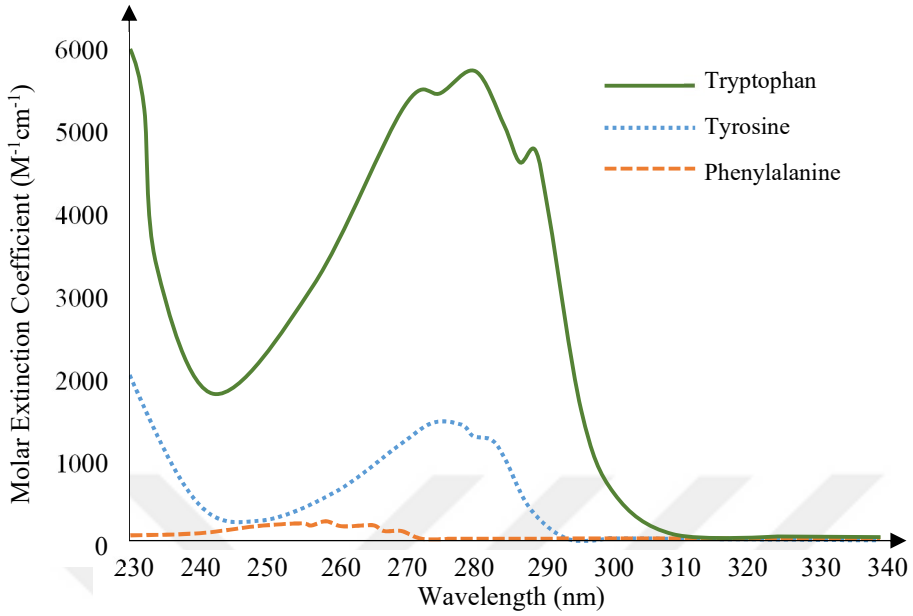


Figure 1.4. Absorption Spectra of Aromatic Amino Acids.
Absorption peak of proteins is predominantly effected by Trp residue since it gives the highest signal among other aromatic residue at same concentration.

1.3.1.2 Circular Dichroism Spectroscopy Analysis in Proteomics

Circular dichroism Spectroscopy (CD) is a highly effective method for quickly determining the secondary structure and folding characteristics of proteins. Protein CD is primarily employed to assess the folding status of an expressed and purified protein, as well as to evaluate the impact of mutations on its conformation and stability.[42] CD analyzes the uneven absorption of light that is polarized in a left-handed and right-handed circular manner.[61] A chiral molecule, such as a protein, cannot be superimposed on its mirror image and thus exhibits optical activity.[62] In other words, it gives a signal in the CD spectrum. Figure 1.5. shows the presence of a negative band around 220 nm in an α-helical protein is attributed to the highly hydrogen-bonded structure of this conformation. This transition is largely unaffected by the length of the helix. β-sheets in proteins exhibit a wide negative peak around

225 nm and a significant positive peak around 195 nm. In contrast, disordered extended chains display a weak, broad positive circular dichroism peak around 210 nm and a prominent negative peak near 200 nm.[63] Further, proteins with complete random structure or denatured proteins devoid of α-helixes and β-sheets show weaker peaks. These proteins lack chirality due to the free rotation of the residues. Small peaks occur due to backbone amide group transitions, which are n → π* peak occurring around 210-230 nm and π → π* peak around 190-200 nm.[61,64,65]

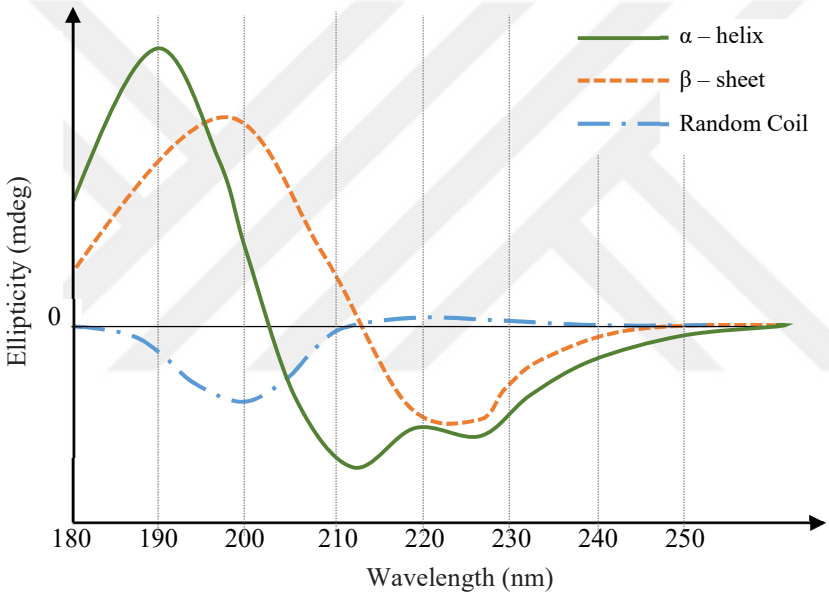


Figure 1.5. Circular Dichroism Spectrum of Secondary Structures of Proteins[61,63] Predominant signal of α–helix structure is observed around 220 nm and consists of two negative peaks. β–sheet structure gives a negative peak around 225 nm, however this signal cannot be observed if protein includes any α–helix structure since α–helix signal predominates the β – sheet signal. When protein only consists of random coils or it is fully denaturated, random coil structure is observed which is a weaker negative peak around neat 200 nm and a slight positive peak around 210 nm.

1.3.2 Colorimetric Assay Analyses in Proteomics

Colorimetric protein assays are used for quick protein quantification analyses. As mentioned above, four of the most used colorimetric assays in the protein analyses are Bradford Assay, Biuret Assay, Lowry Assay, and Bicinchoninic Acid (BCA) Assay. Bradford Assay's working principle includes the binding of Coomassie Brilliant Blue G-250 dye to proteins and the shift of absorption peak from 470 nm to 595 nm.[66] The next three assays depend on Cu^{2+} ion reduction by proteins to Cu^{+} ion in an alkaline medium given in Figure 1.6. This reaction is known as Biuret Reaction. During this reaction, the color of the solution changes from blue to purple since cupric ion solution are blue and cuprous ion-protein complex in alkaline medium are purple.[67,68] In literature, there are several studies investigating the mechanism of the Biuret Reaction. It is found that Cys, Trp, and Typ residues that have reducing groups effect the purple color formation in the Biuret Assay.[69] However, when peptides including three or more residues undergo Biuret Reaction, purple color formation is also observed.[70] This phenomenon occurs even when oligopeptides don't include Cys, Trp, and Typ residues.[70,71] Hence, in order to best represent the literature, protein itself given as oxidized counterpart while cupric ion is being reduced. The solution was then analyzed at 540 nm. As the protein concentration increases, purple formation increases in the solution. This assay is suitable for 5-160 mg/ml protein analysis.[67]

$$\text{Protein} + Cu^{2+} \xrightarrow{^{-}OH} \text{Protein}_{ox} + Cu^{+}$$

Figure 1.6. Reduction of cupric ion by protein to cuprous ion and oxidation of protein as a result of this reduction also known as biuret reaction.
Literature suggests that copper reduction is done by Cys, Trp, and Typ residues that have reducing side groups and peptide backbone of oligopeptides that has three or more residues/ proteins. Hence reaction is given as protein structure being oxidized rather than specific sites of the protein.

Lowry Assay has a similar working principle as Biuret Assay. Folin phenol reagent is added to the copper solution reduced by proteins. The blue to purple color development indicates the presence of proteins. Samples are then analyzed for their absorption of light at 750 nm.[72]. The working concentration range of this assay is 10 ug/ml to 1.0 mg/ml protein.

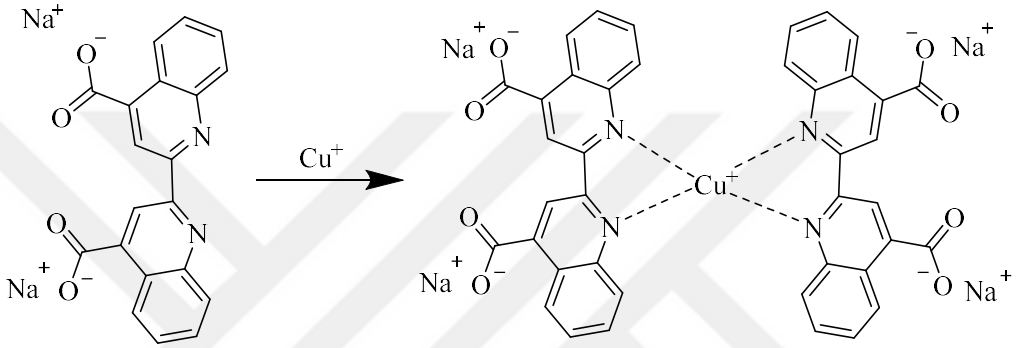


Figure 1.7. Chelation Reaction of Bicinchoninic Acid Disodium Salt and Cuprous Ion in BCA Assay[73]

BCA Assay is another analysis. Figure 1.7. shows color development reaction of BCA Assay and it happens as reduced copper ions chelate with BCA salt to develop a purple color that is analyzed at 562 nm. Protein concentrations are typically between 0.5 ug/ml and 1.5 mg/ml.[73] Due to its stability under alkali conditions, the BCA assay offers the advantage of being a one-step process, as opposed to the two-step process required by the Lowry assay. Additionally, this assay is more tolerant to detergents and denaturing agents than Lowry Assay.[74] Lastly, there is another salt that creates a colorful complex with cuprous ions named 2,9-dimethyl-4,7-diphenyl-1,10-phenanthroline or bathocuproine in short.[75] In literature, bathocuproine and its derivative bathocuproine disulfonate (BCS) are used in protein and peptide analysis.[76,77]The color development reaction of this assay is given in Figure 1.8. Firstly, a “reverse biuret method” is utilized where a protein solution is treated with

tartrate-complex cupric ion. Proteins displace cupric ions and form chelates. Excess cupric ions are reduced to cuprous ions by ascorbic acid, chelated by bathocuproine, and analyzed at 485 nm.

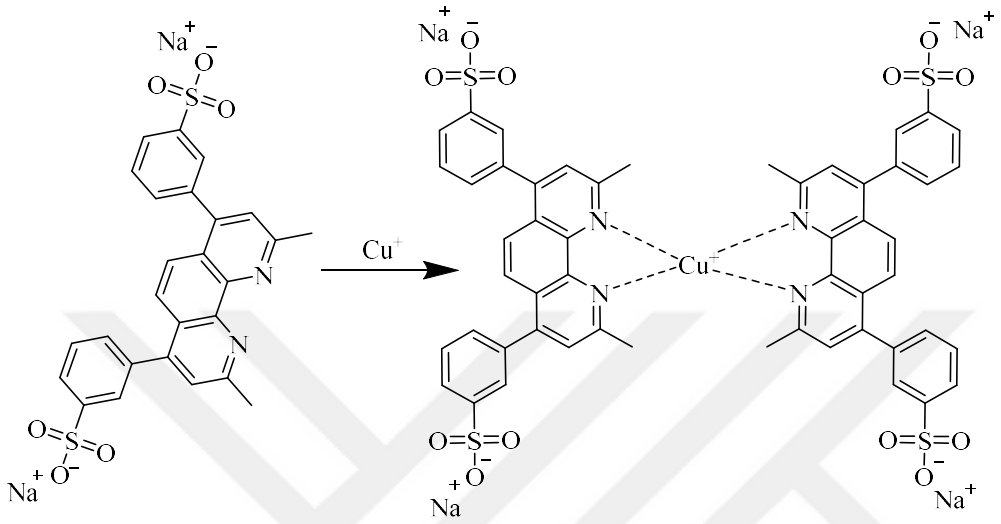


Figure 1.8. Chelation Reaction of Bathocuproine Disulfonate (BCS) Salt and Cuprous Ion in BCA Assay[76]

This method claims to be as sensitive as other methods, compatible with a wider range of proteins and improves on loss upon the adherence to cell walls.[77] It is observed that shorter peptide fragments showed better color formation in this method. Additionally, cystine, tyrosine and tryptophan residues improved to color development.[76]

1.3.3 Mass Spectrometry Analysis in Proteomics

Mass spectrometry is extensively utilized for the analysis of biological samples, therefore is an important instrument for proteomics research.[78] Protein analysis techniques given above yields information about total protein in a sample. However, they are not specific nor selective. MS is a more advanced analysis technique. It can

yield results including but not limited to identification, primary structure sequencing, interactions, post-translational and chemical (artifact) modifications, relative and absolute quantifications.[78–82]

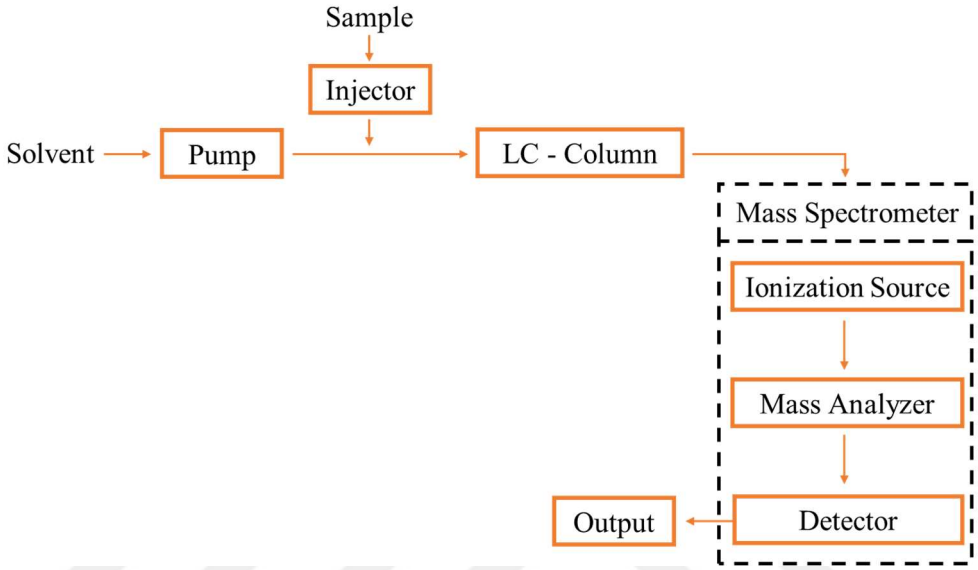


Figure 1.9. General Block Diagram of a LC/MS Instrument

In proteomics field, MS is generally coupled with a liquid chromatogram. The general scheme of the LC/MS is given in Figure 1.9. The working principle of the analysis is based on firstly the injection of a dissolved sample to the LC system. Here sample constituents, in proteomics case protein or peptides, are separated according to the interactions of them with mobile and solid phases. C8 or C18 reverse-phase columns are commonly used in proteomics studies.[83] C18 solid phase is utilized when a sample consists of unknown peptides or includes a wide range of peptides with varying polarities.[84] LC separated elutes reach the MS at different times. In MS, firstly solvent gets evaporated and sample ionizes at the ionization source. In proteomics, soft ionization sources are preferred for their low fragmentation property during ionization. LC/MS is commonly paired with Electron Spray Ionization (ESI). ESI sprays liquid samples consisting of proteins or peptides through a charged

capillary. Here, solvents evaporate, and analytes become charged positively or negatively depending on which one is desired.[85] Then, charged biomolecules are lead to the mass analyzer while neutral molecules are disposed. Mass analyzer differentiates the molecules according to their mass/charge (m/z) values. There can be several mass analyzers in a MS. These MSs are called Tandem Mass Spectrometry or MS/MS. In Figure 1.10, a representative scheme over Triple Quadruple MS (QQQ-MS) is given for MS and MS/MS analyses. A tandem mass spectrometer comprises of two mass analyzers combined by a collision/activation cell[86], which receives peptides ionized in the mass spectrometer as complete molecules and fragments them. Hence, the utilization of the MS/MS technique is imperative for the accurate analysis of peptide sequence through the identification of precursor ions and product ions.[87,88]

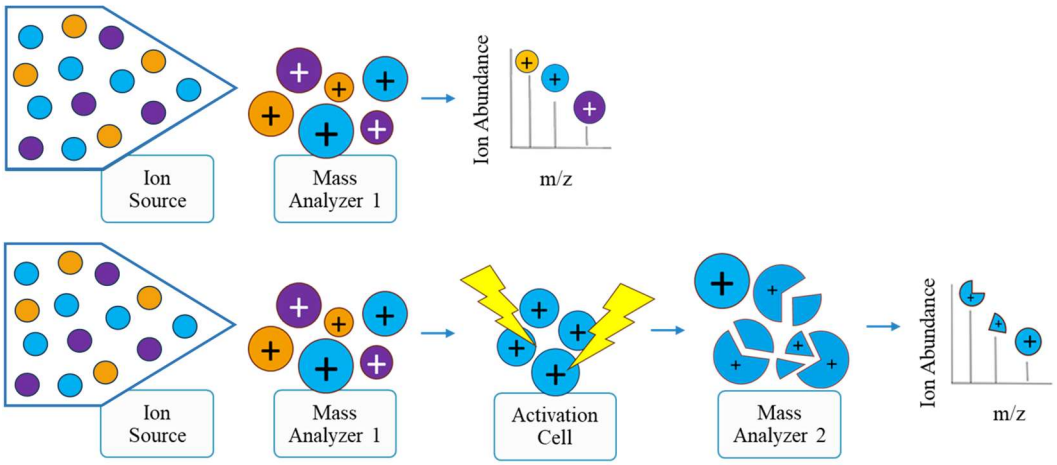


Figure 1.10. MS Analysis vs MS/MS Analysis.
MS analysis gives the m/z values of peptides in the injected sample. MS/MS analysis provides detailed information about the constituents of selected peptide or peptides in the sample by further fragmenting the bonds of these peptides via collision energy. Then gives the m/z values of formed fragments.

Mass analyzers can be divided into three categories as low, medium, and high resolution. High resolution mass spectrometer (HRMS) is a useful instrument to

identify the peptides in mixtures since it can separate the coeluted peptides with 0.01 to 0.001 Da accuracy.[89] HRMS also allows for detection and localization of modifications.[90] Examples to such mass analyzers can be given as Time-of-Flight (TOF), Fourier Transform Ion Cyclotron Resonance (FTICR) and Orbitrap. Continuing with the analysis, every molecule striking the detector face triggers a signal. Computational software utilizes this signal to create signal strength vs m/z value spectrum for every elution point of the sample. This results in a three-dimensional LC-MS Data.[91–93]

1.3.4 Database Search and Protein Identification

Database software is widely used in peptide centric proteomics studies. Software matches the theoretical mass spectrum libraries to experimental as shown in Figure 1.11. Then, every match gets a score. Peptide sequences above a predetermined threshold generally get considered as correct. Then peptide to protein relation is investigated according to preferred parameters and the type of the analysis utilized.[13,94]

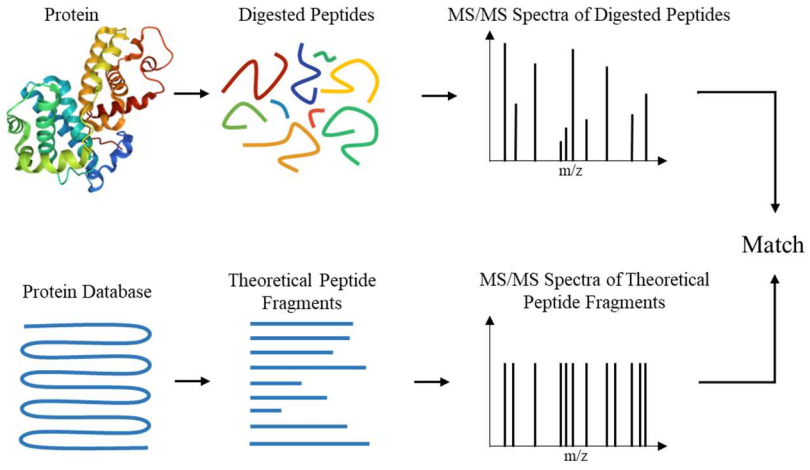


Figure 1.11. Software Match of *in silico* MS/MS Spectra with Experimental Results to Yield Qualitative and Quantitative Results

1.3.5 Peptide Centric Approach

There are two main approaches in MS analysis of proteins which are Top-Down Proteomics and Bottom-Up proteomics. General pathways of them are shown in Figure 1.12. The top-down approach is utilized for qualification analyses and native proteins are used as analytes. This approach yields detailed information about proteins and gives information about modifications in protein level.[95] Bottom-up approach, also called Shot-gun approach, is used for both identification and quantification proteomics.[96] The sensitivity, throughput, and proteome coverage of bottom-up proteomics surpasses top-down proteomics. [80,97] Sample preparation step is an important part of the bottom-up proteomics and effects the quality and the quantity of the peptides. The sequence of the obtained peptides changes according to the digestion type or utilized proteolytic enzyme. Some of these peptide sequences belong only to one protein which are called "unique peptides". Bottom-up approach utilizes one or more of this unique peptide level information to obtain the protein level information. This process is called "Peptide-Centric Approach".[98]

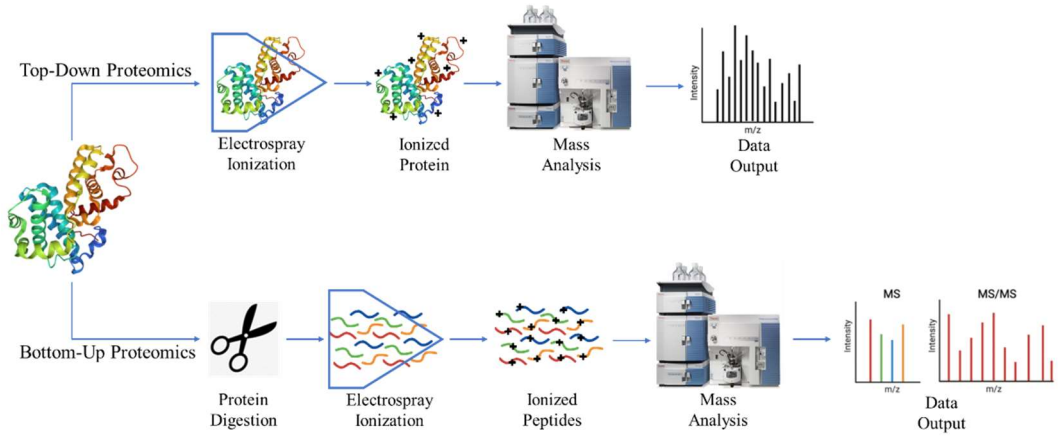


Figure 1.12. Schemes of Top-Down and Bottom-Up Approaches in Proteomics. Top-down proteomics is a qualitative analysis type. Here, proteins are ionized and analyzed as a whole macromolecule. Whereas, in bottom-up proteomics, proteins are digested into peptides via additional procedures. Then, these peptides are introduced to the instrument. The results of obtained peptides are considered to represent the mother protein. Bottom-up approach is a quantitative analysis type.

This approach is based on the assumption that information of one or a few unique peptides can represent the protein information.[80] However, information on the peptide level can be effected by several conditions according to digestion efficiency, modifications, peptide ionization during MS analysis and repression of lower signals.[13]

1.3.6 Sample Preparation for Bottom-Up Proteomics

Proteins are mostly in an unbound/free state in the serum samples and can be fragmented into peptides with proteolytic digestion. The most common proteolytic sample preparation types for bottom-up proteomics include in-solution digestion and in-gel digestion.[99] In-gel digestion is a technique that utilizes gel electrophoresis in order to separate proteins based on their molecular weights. Following the gel electrophoresis, desired protein band is cut out from the gel manually and digested. Gel separation may help to remove impurities. However, the process includes several steps that are error prone. Results may be non-reproducible depending on the sample content, gel separation and manual band cutting.[100] In-gel protocols include higher concentration of enzyme addition compared to in-solution digestion. This causes significant autolysis of enzyme hence increasing interferences in output spectra.[101] In-solution digestion is an alternative to gel digestion. In this case, the samples are not pre-separated. Instead, they undergo digestion process in buffer solutions. This process is more convenient, less susceptible to human errors, reduces sample loss,[100] and provides greater control over the process.[102] There are five steps in in-solution digestion which are extraction/solubilization, denaturation, reduction, alkylation and proteolytic digestion. Since proteins are mostly in free form in the serum samples, extraction step is not necessary and proteins in the serum sample can be directly solubilized in the buffers. In-solution digestion procedure steps are given in Figure 1.13.

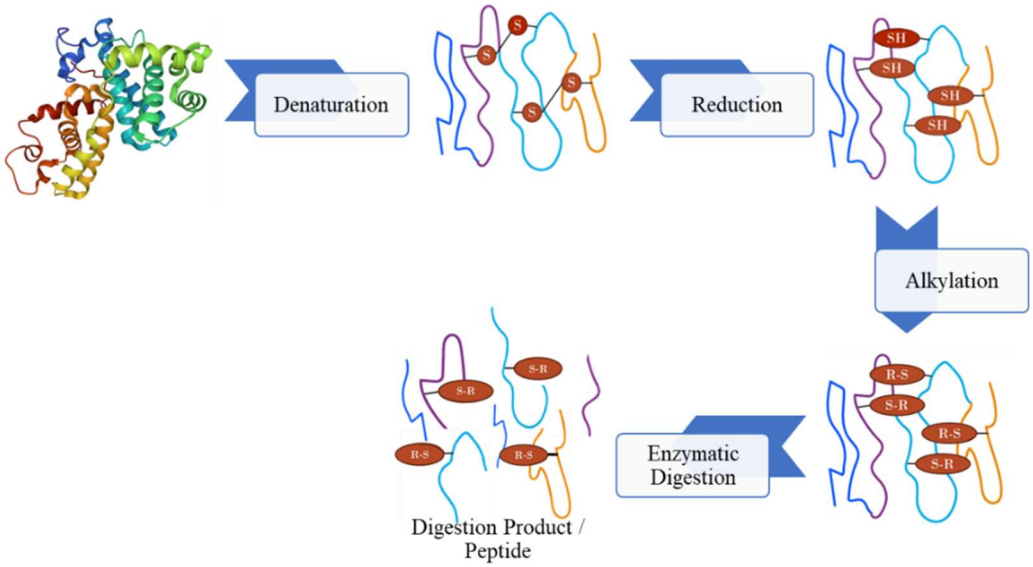


Figure 1.13. In-Solution Digestion Procedure Steps

1.3.6.1 Denaturation Process

The majority of proteins fold into distinct three-dimensional structures in their native states.[103] Tertiary structure of a protein is held by non-covalent interactions, ion pairs, disulfide bonds, and most importantly hydrophobic effect.[104] In order to ease the interaction of the enzyme with inside regions of the protein, it is denatured to lose its tertiary structure.[80] There are several denaturation reagents used in sample preparation such as urea or sodium deoxycholate (SDC).[105,106] Furthermore, some methods doesn't include reagents for this step and denatures the proteins thermally.[107] Heat denatures the proteins by effecting the non-covalent interactions. The reagents like urea and SDC denature the proteins by disturbing the hydrophobic core of the protein.[108] In literature, denaturation reagents are commonly used in heat incubation to improve denaturation.

1.3.6.2 Reduction Process

Denaturation step doesn't affect the covalent disulfide bonds in-between cystine residues. These bonds are broken commonly with Dithiothreitol (DTT) or Tris(2-carboxyethyl) phosphine hydrochloride (TCEP).[109] As seen in Figures 1.14 and 1.15, reduction reagents acts as nucleophiles and attacks the thiol group of cystine bridges. Both of the mechanism involves S_N2 nucleophilic substution.[110] Upon attack to the covalent bridge, thiol group collapses and one thiol group leaves freely while reduction reagent forms bond with other Cys. Then, DTT leaves by attacking itself from its other thiol group, creating a cyclic oxidized form of itself. Whereas phosphonium ion (TCEP) undergoes hydrolysis, freeing thiol of Cys and becoming a stable phosphine oxide.

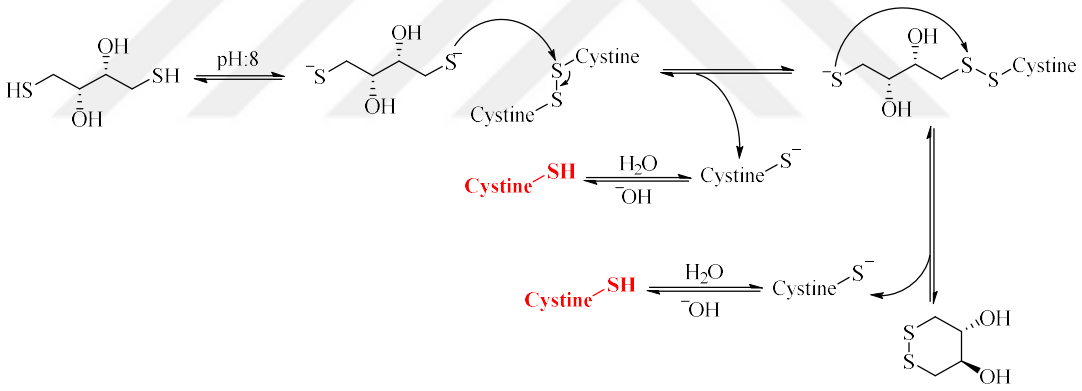


Figure 1.14. Reduction of Disulfide Bonds with DTT[111]

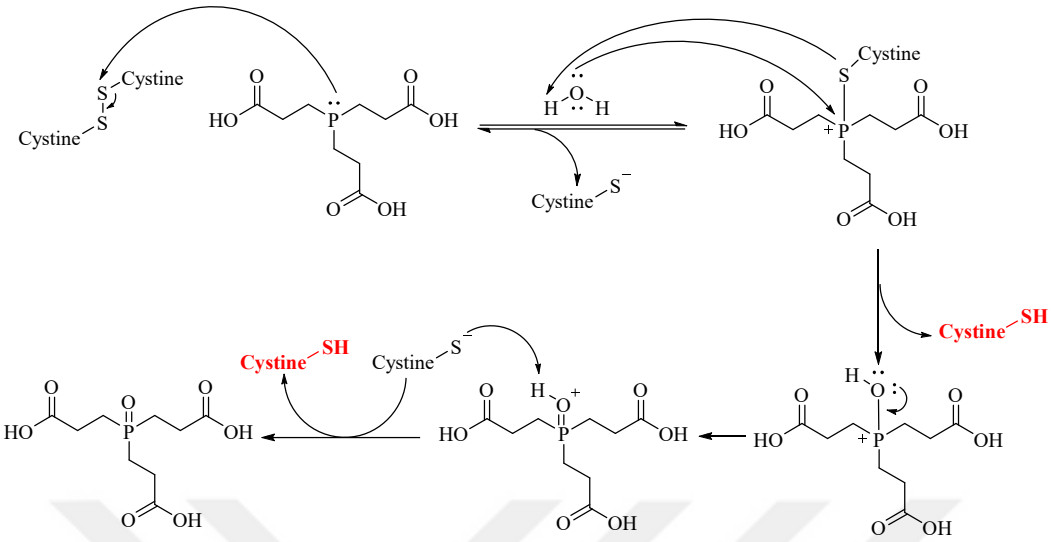


Figure 1.15. Reduction of Disulfide Bonds with TCEP[110]

1.3.6.3 Alkylation Process

Reduction cleaves disulfide bonds however leaves the sulfide sides active. So, in order to prevent the re-formation of disulfide bonds, thiol groups are quenched by alkylation as seen in Figure 1.16.[112] In literature, mostly iodoacetamide (IAA) is utilized. Nevertheless, there are some studies using chloroacetamide (CAA) instead of IAA in alkylation step.

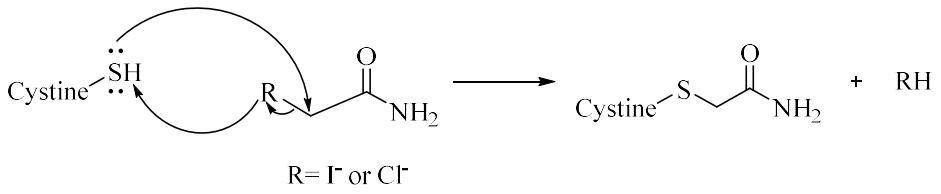


Figure 1.16. Alkylation of reduced Cys Residues with IAA or CAA[110]

1.3.6.4 Proteolytic Digestion Process

The next step in the processes after alkylation is proteolytic digestion. Several proteases can be used for this step, such as Chymotrypsin , LysC, GluC, AspN, and trypsin, to name a few. Nevertheless, trypsin is the most used enzyme for proteomics studies and its 3D structure is shown in Figure 1.17.[113] One of the most important reasons is the cleavage specificity of it. Trypsin cleaves the carboxy-terminal of Arginine and Lysine residues. This cleavage creates C-terminal basic peptides, which creates another advantage for mass spectrometry studies. Since the C-terminal basic peptides enhances the ionization of peptides in the positive mode.[114] However, there is a limitation for the enzyme. It cannot fragment the Arg and Lys if a Proline residue is following them in the sequence.[115] This is due to the same reason why trypsin is specific to Arg and Lys residues. The shape of the specificity pocket of the trypsin is long and narrow with negatively charged Asp residue present at the bottom. Hence positively charged and long residues are suitable to fit inside the specificity pocket like Arg and Lys. However, if these residues are followed by a Pro residue, the rigid shape constrains the overlapping of the Arg/Lys and the pocket, hence inhibiting the fragmentation of that site.[115]

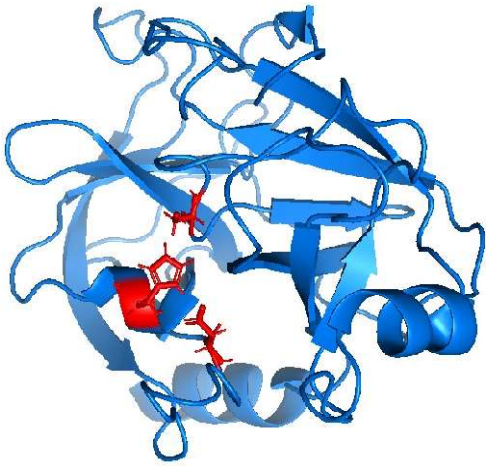


Figure 1.17. Trypsin Enzyme Structure with Catalytic Triad (Asp102+His57+Ser195) Highlighted[116]

Trypsin is a serine protease, and its active site consists of a catalytic triad of Asp102, His57, and Ser195.[117] This triad is located in the binding pocket and cleaves the C-Term of Arg and Lys residues, as mentioned before. The mechanism of the cleavage by trypsin is given in Figure 1.18.

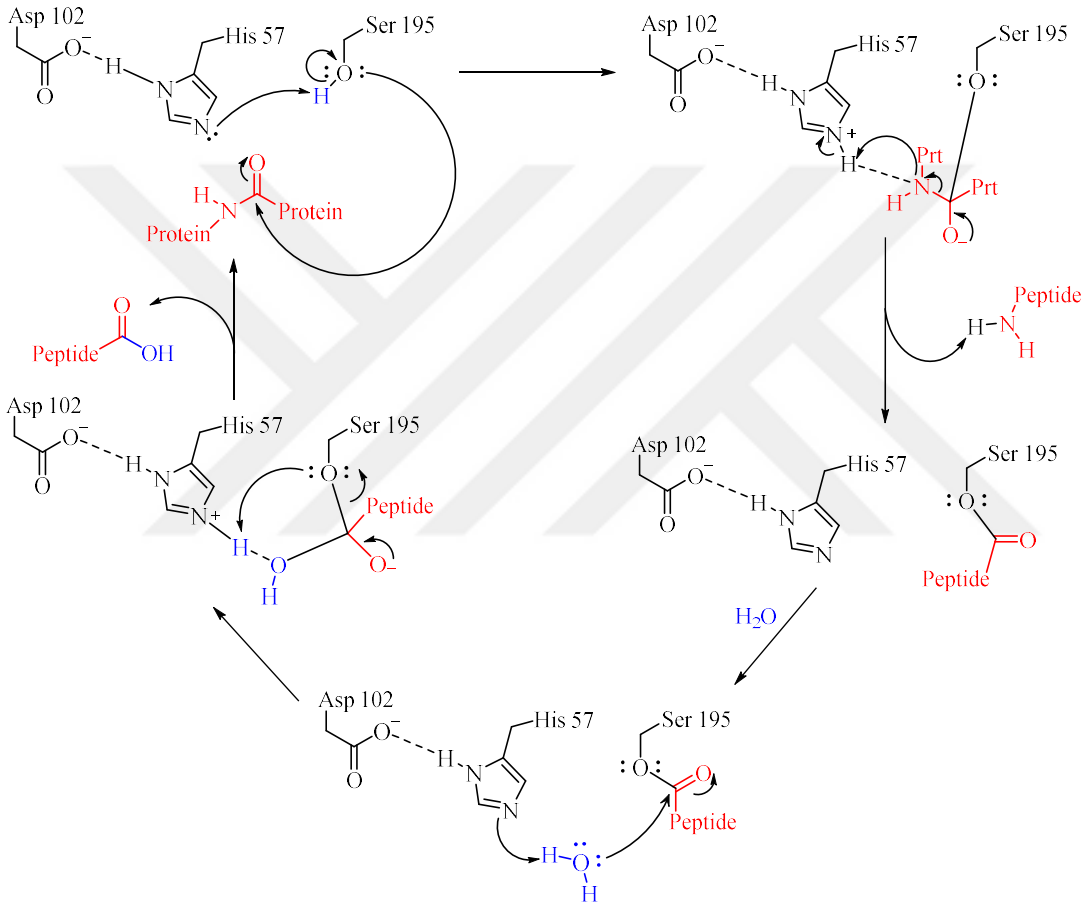


Figure 1.18. Reaction Mechanism of Cleavage Site of the Trypsin[114,118]

The protein chain enters the binding pocket. The carboxyl group of Asp102 makes a Hydrogen bond with the His57, orienting it properly while increasing its nucleophilicity. The amide group of His57 abstracts the proton of the hydroxyl group

of Ser195, increasing the nucleophilicity of the oxygen. Then, Ser195 attacks the carbonyl carbon of the polypeptide chain, creating a tetrahedral intermediate structure. Following, the carbonyl group reforms in the intermediate and collapses the structure, the amide of the polypeptide accepts a protein from His57, and cleavage occurs partially. Then, His57 increases the nucleophilicity of a water molecule, and oxygen of the water attacks the carbonyl carbon of the Ser195 attached carbonyl group of the chain. The formed intermediate collapses upon the reformation of the carbonyl group, and the protein cleavage ends with catalytic sites rejuvenated.[114,118]

1.4 Model Proteins

This study employs two proteins. For spectroscopic studies, BSA was utilized as model protein. It is the reference protein for BCA Assay[119] and Peptide Assay[120]. Additionally, there are numerous references for expected UV-Vis and CD spectra.[121–125] However, it is not suitable for MS analysis since it is not a human protein and this study based on human sera and plasma digestion procedures. Hence, for MS analysis part of the study, a human blood protein was utilized. This protein is selected as A2MG since it is has been used clinical MS studies in the literature.[126–129] Also PeptideAtlas database[130] includes a database for A2MG protein.

1.4.1 Bovine Serum Albumin (BSA)

Bovine Serum Albumin (BSA) is a widely used reference protein in proteomics studies. It is the most predominant protein in the cow blood.[131] BSA possesses 607 residues in its sequence and has the mass of 66.5 kDa.[132] Figure 1.19 shows that its secondary structure predominantly consists of α–helix. Its solubility is 40 mg/ml in deionized water.[133] Isoelectric point of BSA in water at 25°C is 4.7.[134]

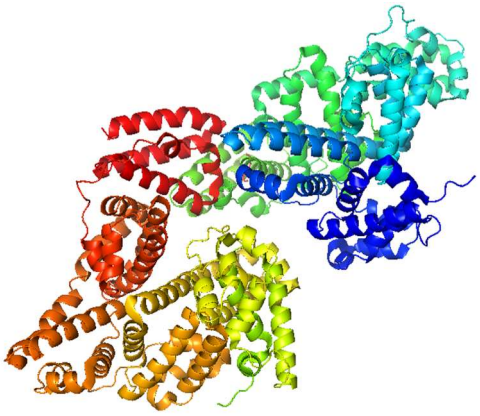


Figure 1.19. 3D Structure of BSA Protein

It provides several advantageous features, including sensitivity, and compatibility. Hence, it is used as primary standard for protein assays.[135] Additionally, many studies in the literature utilizes the analysis of secondary structure of BSA via CD.[121–125] Figure 1.20 shows the expected CD spectrum and absorption spectrum of BSA protein. Since it has α–helixes in its structure as seen in Figure 19, CD spectra consists of a negative double peak around 220 nm. Next, its UV-Vis spectrum is expected yield a peak near 277 nm.

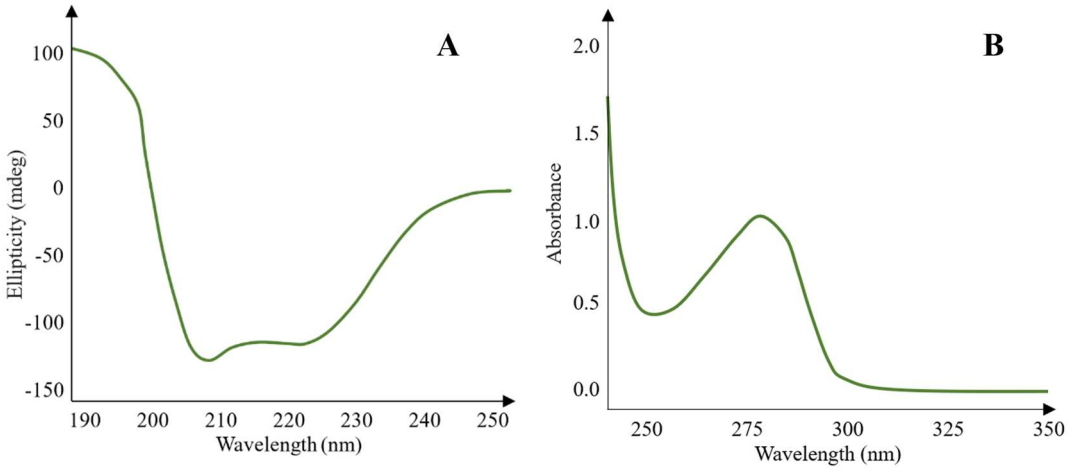


Figure 1.20. (a) CD Spectrum and (b) Absorption Spectrum of pure BSA.

The primary α-helix structure of BSA yields a CD spectra signal of a negative double peak at approximately 220 nm. [136] Its UV-Vis spectrum has a peak at 277 nm.[137]

1.4.2 Alpha-2-Macroglobulin (A2MG)

Alpha-2-macroglobulin (A2MG) is a highly large protein found in significant quantities in human serum. It functions as a broad-spectrum proteinase inhibitor of proteinase enzymes.[138,139] The A2MG is a versatile protein that has a vital function in maintaining the balance of blood components. In humans, the level of A2MG decreases as the individual gets older, going from around 1.5 to 4.0 mg/ml.[140,141] Therefore, it plays a role in homeostasis of blood and is highly significant in the development of age-related diseases.[142] A2MG is chosen as a reference protein due to its biomarker identity for a wide range of diseases such as Alzheimer's disease[143], diabetes[144], Parkinson's disease[139].

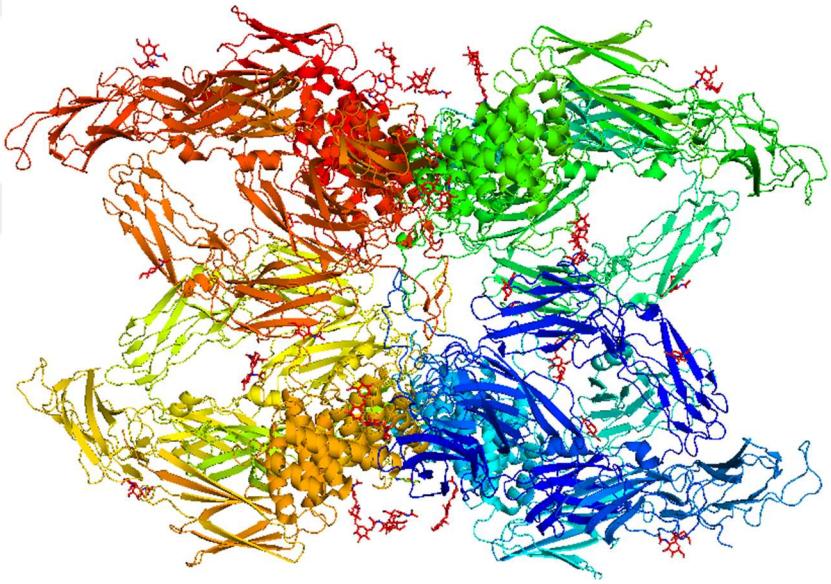


Figure 1.21. 3D Structure of A2MG Protein

The protein ID of the A2MG is P01023 and its sequence involves 1474 residues with a total mass of 720 kDa. Figure 1.21 shows that the secondary structure of the protein comprises α-helix, β-sheets, and random coils. PTMs of A2MG consist of disulfide bond, N-Glycoprotein, and cross-linkage. Additionally, it is translated from A2M gene and composed of four identical subunits.[132]

1.5 Motivation

In the bottom-up proteomics, protein analysis is done through its peptide fragments. Therefore, proteins are digested into peptides with a process called digestion. Digestion procedure can be described as a breaking down process of a protein to peptides which include several steps called denaturation, reduction, alkylation and proteolytic digestion.

Variations in the protein results for healthy vs. diseased states lead to biomarker detection. Hence, biomarker detection heavily depends on the reproducibility of the protein results. If the results do not align regarding the protein levels in a healthy individual, it is insufficient to make comparisons to identify a diseased state. Since peptides are obtained through the digestion process, it is an important part of the peptide-centric approach.

Literature search indicated that biomarker studies used various digestion methods without indications of the potential impact of sample preparation to protein product variation. This study examined whether different digestion processes produce protein products that were equivalent or each process had unique effects and outcomes. Additionally, equivalency of effects of different digestion procedure steps on proteins were investigated.

CHAPTER 2

METHODS AND MATERIALS

2.1 Materials

Bovine Serum Albumin (BSA), Alpha-2-Macroglobulin (A2MG), Ammonium Bicarbonate (ABC), Tris(hydroxymethyl)aminomethane (Tris), Sodium Deoxycholate (SDC), Dithiothreitol (DTT), Tris(2-carboxyethyl) Phosphine Hydrochloride (TCEP), Iodoacetamide (IAA) and Chloroacetamide (CAA) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Urea was obtained from Supelco. (Bellefonte, PA, USA). Formic Acid was purchased from Thermo Fischer (Waltham, MA, USA). Sequencing grade Trypsin and Trypsin/Lys-C mixture were obtained from Promega (Madison, WI, USA). Filter-Aided Sample Preparation kit was obtained as Abcam, FASP Protein Digestion Kit (ab270519). BCA Assay kit obtained as ThermoScientific Pierce® Microplate BCA Protein Assay Kit – Reducing Agent Compatible, 23252. Peptide Assay kit was obtained as ThermoScientific Pierce Quantitative Colorimetric Peptide Assay, 500 Assays, 23257. The enzyme activity test kit was obtained as the Trypsin Activity Colorimetric Assay Kit, MAK290. The UV-Vis Spectrophotometer utilized was the Thermo Scientific Multiskan Skyhigh Microplate Spectrophotometer. The CD Spectrophotometer utilized was the Jasco J-1500 Circular Dichroism Spectrometer.

2.2 Methods

2.2.1 Literature Search

Literature was searched for the in-solution digestion methods using human plasma and serum over databases of ACS Publications, Web of Science, Nature Portfolio, Google Scholar, ScienceDirect, and PubMed. Over a hundred methods were obtained from this search. Ten methods were chosen among them to represent the literature considering the frequency of the use of the method or the reagents and the parameters of the method. Then methods were experimentally compared with each other in terms of denaturation performances with UV-Vis and CD Spectroscopies, enzyme activities using Colorimetric Enzyme Activity kit, yielded peptide concentration using Peptide Assay kit and yielded peptide sequences, their modifications, and protein correlations using Mass Spectrometry analysis. BSA protein was utilized for spectroscopic and kit analyses. BSA is widely used as reference protein in the literature and it is compatible with BCA Assay and Peptide Assay. Also, it has spectra references for spectroscopic methods used in this study. However, BSA doesn't have references for MS studies. Also, since it is not a human sourced protein, it is not clinically relevant for biomarker studies. For this reason, A2MG was utilized as sample protein for the Mass Spectrometer analysis since it is a human biomarker found in blood.

2.2.2 Digestion Procedures

Experimental procedures of ten selected methods were scaled down to be compatible with 96-well plates. 20 mg/ml BSA solution[145] was prepared using deionized water, aliquoted and placed in a -20°C freezer. 1 mg/ml A2MG solution was prepared in 0.1% Formic Acid solution in deionized water, aliquoted and placed in a -20°C freezer. All aliquots were taken out before the experiment and thawed at +4°C in

order to ensure the equivalency of the protein states for all methods. BSA samples were prepared as six replicates and A2MG samples were prepared as three replicates. BSA samples were used in spectrometry and kit experiments and variations among sets are compared. In order to increase precision and reduce the effect of random error in this experiments, replicate number was chosen as six. When it comes to A2MG samples, they were pooled into one sample and then analyzed with MS. Due to pooling and also availability of this A2MG, replicate number was chosen as three. In all methods, 140.0 µg BSA and 11.2 µg A2MG proteins were digested per replicate. In order to prevent protein loss due to sticking to the well walls, Low-Binging 96-Deepwell plates were utilized during the procedures.

2.2.2.1 Method A

Method A was chosen as digestion process from Ozcan et al., 2017.[146–149] Proteins were added to wells, and volumes were adjusted to 190.0 µL with 50 mM ABC Buffer. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 10.0 µL of 650 mM DTT stock solution prepared in the ABC Buffer was added to reach 32.5 mM DTT concentration. The plate was again spun down for 10 seconds, sealed with a plate seal, and incubated at 60 °C for 30 minutes with 80 rpm mixing. Then, 100.0 µL 225 mM IAA stock solution was added to reach 75.0 mM IAA concentration and pipetted to mix. Plates were covered immediately with aluminum foil and incubated at room temperature for 30 minutes in a dark environment. Then 0.1 mg/ml trypsin solution prepared with trypsin reconstitution buffer was added to reach a 50:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 17 hours with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30°C for ~5 hours and reconstituted with 300.0 µL of 0.1% FA solution.

2.2.2.2 Method B

Method B was chosen as digestion process from Addona et al., 2009.[150] Proteins were added to wells, and volumes were adjusted to 20.0 µL with 100 mM Tris Buffer at pH 8.0. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 28.6 µL of 7.8 M Urea+ 26.0 mM DTT stock solutions prepared in the Tris Buffer was added to reach 6 M Urea and 20 mM DTT concentration. The plate was again spun down for 10 seconds, sealed with a plate seal, and incubated at 37 °C for 30 minutes with 80 rpm mixing. Then, 50.0 µL of 79.0 mM IAA stock solution was added to reach 40.0 mM IAA concentration and pipetted to mix. Plates were covered immediately with aluminum foil and incubated at room temperature for 30 minutes in a dark environment. Then samples were diluted with 299.4 µL Tris Buffer. Then 0.1 mg/ml trypsin solution prepared with trypsin reconstitution buffer was added to reach a 50:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 18 hours with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30 °C for ~5 hours and reconstituted with 300.0 µL of 0.1% FA solution.

2.2.2.3 Method C

Method C was chosen as digestion process from Geyer et al., 2016.[151–155] Proteins were added to wells, and volumes were adjusted to 10.0 µL with 100.0 mM Tris Buffer with pH 8.5. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 40.0 µL of 1.0% SDC+ 10.0 mM TCEP+ 40.0 mM CAA stock solution prepared in the Tris Buffer was added to samples. The plate was again spun down for 10 seconds, sealed with a plate seal, and boiled in water for 10 minutes. Plates were spun down. Then 0.01 mg/ml Trypsin/

Lys-C solution prepared with Trypsin/ Lys-C reconstitution buffer was added to reach a 100:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 1 hour with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30°C for ~5 hours and reconstituted with 300.0 µL of 0.1% FA solution.

2.2.2.4 Method D

Method D was chosen as digestion process from Kuzyk et al., 2009.[156] Proteins were added to wells, and volumes were adjusted to 20.0 µL with 25.0 mM ABC Buffer. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 20.0 µL of 12.5% SDC and 10.0 µL of 25 mM TCEP stock solution prepared in the ABC Buffer was added to reach 5% SDC+ 5.0 mM TCEP concentration. The plate was again spun down for 10 seconds, sealed with a plate seal, and incubated at 60 °C for 30 minutes with 80 rpm mixing. Then, 10.0 µL 60.0 mM IAA stock solution was added to reach 10.0 mM IAA concentration and pipetted to mix. Plates were covered immediately with aluminum foil and incubated at 37 °C for 30 minutes in a dark environment. After alkylation, samples were diluted with 140.0 µL ABC Buffer. Then 0.1 mg/ml trypsin solution prepared with trypsin reconstitution buffer was added to reach a 20:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 16 hours with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30°C for ~5 hours and reconstituted with 300.0 µL of 0.1% FA solution.

2.2.2.5 Method E

Method E was chosen as digestion process from Demichev et al ., 2021.[157] Proteins were added to wells, and volumes were adjusted to 11.2 µL with 100 mM ABC Buffer. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 14.0 µL of 12.5 M Urea+ 78.5 mM DTT stock solution prepared in the ABC Buffer was added to reach 7.3 M Urea and 45.8 mM DTT concentration. The plate was again spun down for 10 seconds, sealed with a plate seal, and incubated at 30 °C for 60 minutes with 80 rpm mixing. Then, 10.0 µL 26.2 mM IAA stock solution was added to reach 7.7 mM IAA concentration and pipetted to mix. Plates were covered immediately with aluminum foil and incubated at 23 °C for 30 minutes in a dark environment. Then samples were diluted with 135.0 µL ABC Buffer. Then 0.1 mg/ml trypsin solution prepared with trypsin reconstitution buffer was added to reach a 152:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 17 hours with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30°C for ~5 hours and reconstituted with 300.0 µL of 0.1% FA solution.

2.2.2.6 Method F

Method F was chosen as digestion process from Fortin et al., 2009.[158] Proteins were added to wells, and volumes were adjusted to 11.2 µL with 100 mM Tris Buffer with pH 8.0. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 50.0 µL of 7.2 M Urea+ 36.0 mM DTT stock solution prepared in the deionized water was added to reach 6.0 M Urea and 30.0 mM DTT concentration. The plate was again spun down for 10 seconds, sealed with a plate seal, and incubated at 40 °C for 30 minutes with 80 rpm mixing. Then, 10.0 µL 350.0 mM IAA stock solution prepared in deionized water was added to

reach 50.0 mM IAA concentration and pipetted to mix. Plates were covered immediately with aluminum foil and incubated at room temperature for 40 minutes in a dark environment. Then samples were diluted with 230.0 µL deionized water. Then 0.1 mg/ml trypsin solution prepared with deionized water was added to reach a 30:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 18 hours with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30°C for ~5 hours and reconstituted with 300.0 µL of 0.1% FA solution.

2.2.2.7 Method G

Method G was chosen as digestion process from Oberbach et al., 2014.[159] Proteins were added to wells, and volumes were adjusted to 50.0 µL with 25.0 mM ABC Buffer. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 50.0 µL of 10.0% SDC prepared in the ABC Buffer was added to reach 5% SDC and samples were incubated at room temperature for 5 minutes. Then 269.6 µL ABC Buffer was added to each well. Next, 41.1 µL of 50.0 mM TCEP and 45.6 µL 100 mM IAA stock solutions prepared in the ABC Buffer were added to reach 5.0 mM TCEP+ 10.0 mM IAA concentration. The plate was again spun down for 10 seconds, covered immediately with aluminum foil, and incubated at 37 °C for 30 minutes in a dark environment. Then 0.1 mg/ml trypsin solution prepared with trypsin reconstitution buffer was added to reach a 20:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 2 hours with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30 °C for ~5 hours and reconstituted with 300.0 µL 0.1% of FA solution.

2.2.2.8 Method H

Method H was chosen as digestion process from Hober et al., 2019.[160] Proteins were added to wells, and volumes were adjusted to 20.0 µL with 257.0 mM Tris Buffer with pH 8.0. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 25.0 µL of 11.1 M Urea+ 18.7 mM TCEP stock solution prepared in the Tris Buffer was added to reach 7.7 M Urea and 12.9 mM DTT concentration. The plate was again spun down for 10 seconds, sealed with a plate seal, and incubated at 37 °C for 30 minutes with 80 rpm mixing. Then, 20.0 µL of 97.2 mM IAA stock solution was added to reach 34.6 mM IAA concentration and pipetted to mix. Plates were covered immediately with aluminum foil and incubated at room temperature for 30 minutes in a dark environment. Then samples were diluted with 463.0 µL of 100.0 mM Tris Buffer. Then 0.1 mg/ml trypsin solution prepared with deionized water was added to reach a 15:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 17 hours with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30 °C for ~5 hours and reconstituted with 300.0 µL of 0.1% FA solution.

2.2.2.9 Method I

Method I was chosen as digestion process from Gaither et al., 2020.[161] Proteins were added to wells, and volumes were adjusted to 90.0 µL with 200.0 mM Tris Buffer with pH 8.0. The plate was spun down for 10 seconds at 4800 rpm to bring down the possible droplets stuck on the sides. Then 200.0 µL of 8.7 M Urea+ 18.9 mM DTT stock solution prepared in the deionized water was added to reach 6.0 M Urea and 13.0 mM DTT concentration. The plate was again spun down for 10 seconds, sealed with a plate seal, and incubated at 37 °C for 30 minutes with 80 rpm mixing. Then,

50 µL of 272.0 mM IAA stock solution prepared in deionized water was added to reach 40.0 mM IAA concentration and pipetted to mix. Plates were covered immediately with aluminum foil and incubated at room temperature for 30 minutes in a dark environment. Then 0.1 mg/ml trypsin solution prepared with trypsin reconstitution buffer was added to reach a 20:1 protein-to-enzyme ratio. The plate was sealed with a plate seal and incubated at 37 °C for 18 hours with 80 rpm mixing. After incubation, digestion was quenched with 10.0 µL formic acid solution with a final concentration of 0.1% FA. Then, the samples were fully dried using a condenser at 30 °C for ~5 hours and reconstituted with 300 µL of 0.1% FA solution.

2.2.2.10 Method J

FASP Kit was used for the alkylation and digestion of samples. As the kit suggests, proteins were denatured and reduced using a procedure of a clinical study[162] in literature utilizing the FASP Kit in their digestion process. For denaturation and reduction parts, proteins were added to Eppendorf tubes and 400.0 µL 8.0 M Urea+ 200.0 mM TCEP prepared in Tris Buffer provided by the FASP Kit. Then incubation was done at 60 °C for 30 minutes. The rest of the procedure was done according to the FASP Kit instructions. Samples were placed on a Spin Filter and centrifuge at 14,000 x g for 15 minutes. Then, 200.0 µL of Urea Sample Solution was placed onto the Spin Filter and centrifuge at 14,000 x g for 15 minutes. The flow-through was discarded. 10.0 µL Iodoacetamide Solution and 90.0 µL Urea Sample Solution were added to the Spin Filter and vortexed for 1 minute. Then samples were incubated without mixing for 20 minutes in the dark. After incubation, the Spin Filters were centrifuged at 14,000 x g for 10 minutes. Afterward, 100.0 µL of Urea Sample Solution was added to the Spin Filter and centrifuged at 14,000 x g for 15 minutes. This step was repeated twice, and the flow-through from the collection tube was discarded. Then 100.0 µL of 50.0 mM Ammonium Bicarbonate Solution provided was added onto the Spin Filter and centrifuged at 14,000 x g for 10 minutes. This

step was repeated twice. Next, a 75.0 µL digestion solution prepared with ABC buffer and enough trypsin solution to reach a 100:1 protein-to-enzyme ratio was added to the spin filter. Samples were vortexed for 1 minute, tubes were sealed with parafilm, and incubated for 16 hours at 37 °C. After digestion, the spin filter was transferred to a new collection tube. 40.0 µL of 50.0 mM Ammonium Bicarbonate Solution was added to the filters and centrifuged at 14,000 x g for 10 minutes. This step was repeated once more. Then, 50.0 µL of 0.5 M Sodium Chloride Solution provided was added to the filters and centrifuged at 14,000 x g for 10 minutes. Next, samples were acidified with 0.1% FA. Lastly, the samples were fully dried using a condenser and reconstituted with 300.0 µL of 0.1% FA solution.

2.2.3 Absorption Spectroscopy Analysis

Initially, buffers of the digestion methods were prepared and pH values for tris buffer were adjusted to 8.0-8.5 as given in the respective procedure. Then BSA protein was added to have 1.0 mg/ml concentration value. Six replicates for each method were placed onto a UV-VIS region suitable flat-bottomed 96-well plate. Then absorption values of the sample were scanned from 240 nm to 360 nm using the Thermo Scientific Multiskan Skyhigh Microplate Spectrophotometer UV-Vis Spectrophotometer.

2.2.4 Quality Control Samples for BCA Assay & Peptide Assay

In this study, several different quality control sample sets were utilized for different reasons. The first groups of QCs were to ensure the equivalency of the colorimetric assays done for different sets. These quality controls were coded as QC-B.P for BCA Assay plate sets and QC-P.P for Peptide Assay plate sets. These QCs were prepared in bulk, aliquoted equally, and immediately placed at -80°C. They were taken out

immediately before the start of the assay incubation time to ensure the elimination of any environmental effects on them.

The second group of quality controls were used to investigate the possible interference of buffers, reagents, and enzymes for both assay kits. So, deionized water, buffer, buffer + reagents, and buffer + reagents + enzyme solutions were prepared, as shown in Figure 2.1 for each method. Each QC was done in two replicates and coded as QC-BCA and QC-PA.

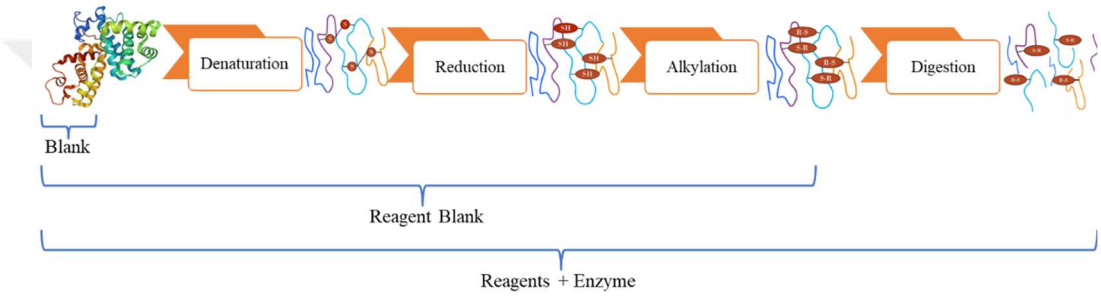


Figure 2.1. Digestion steps each QC represents.

The third QC group was used to observe the effect of denaturation on the BCA Assay. Initially, native proteins were analyzed using the BCA Assay Kit. Then, proteins were denatured using heat denaturation at 60 °C for 30 minutes to ensure the elimination of the possible interferences of the chemicals. Then, denatured proteins were analyzed using the BCA Assay Kit. Sets were prepared in deionized water and buffer of each method. This set was coded as QC-Δ.

The fourth QC group was used to investigate the specificity of the Peptide Assay Kit to peptides by analyzing the native proteins with it. Native proteins were prepared in deionized water and buffers of each method. This set was coded as QC-Reverse.

2.2.5 BCA Assay Procedure

Six replicates of 140.0 µg BSA were prepared for all ten methods chosen. To compare the peptide kit results in equivalent conditions, the concentration of proteins was determined for every replicate using the BCA Assay Kit, and values were used in normalization. Protein concentrations were measured in the respected buffers of the methods as native proteins. For every buffer of each method, calibration curves were drawn utilizing seven concentration points and three replicates for each point. Standard solutions were prepared from 2.0 mg/ml BSA stock solution provided by the BCA Assay Kit between 0.50 mg/ml and 2.0 mg/ml. To ensure the equivalency of every sample set, a quality control sample was analyzed with each set. Detailed information about QCs is given below. Blanks, protein samples, quality control samples, and standard samples were prepared according to the kit procedure. BCA Working Reagent was prepared in a 50:1 ratio of Reagent A to Reagent B provided by the kit. 10.0 µL of the sample was added to a 96-well microplate. Then 260.0 µL of BCA working reagent was added on top. Next, the plate was covered and incubated at 37 °C for 30 minutes. Upon incubation, samples were cooled to room temperature for five minutes. Then, the absorbance of samples was measured at 562 nm. Additionally, LOD and LOQ calculations were made using the measurements of ten blank samples. LOD was calculated by dividing the standard deviation of ten blank samples to slope of calibration curve of BCA assay and multiplying the results by 3.3. LOQ was calculated by multiplying the result of division with 10 instead of 3.3.

2.2.6 Circular Dichroism Spectra Procedure

After the BCA Assay, BSA samples were denaturated and reduced according to their procedure. At this stage, some of the denaturated BSA were taken and diluted to 0.05 mg/ml and samples were scanned from 190 nm to 400 mn using Jasco J-1500 Circular Dichroism Spectrometer.

2.2.7 Enzyme Activity Assay Procedure

Trypsin Activity Colorimetric Assay Kit's working principle is based on the formation of a p-NA molecule that is a specific product of the substrate provided by the kit and the trypsin enzyme. In order to replicate the environment, the enzyme shows activity in each method; reagent blanks were prepared with respect to the procedure of every method and used as buffers in the assay. Two replicates were prepared for every method. Additionally, two QC samples were prepared using the kit buffer provided by the kit. A calibration curve was drawn using the p-NA solution provided by the kit with two replicates for every point. Blanks, samples, QCs, and standard samples were prepared according to the kit's instructions.

For blanks, 50.0 µL respective buffer was placed in the wells. For the calibration curve, 2.0, 4.0, 6.0, 8.0, and 10.0 µL of p-NA stock solution were added to the 96-well plate wells then the volumes were adjusted to 50.0 µl. For samples and QCs, 0.01 mg/ml trypsin solution was prepared and added to the plate wells; then, volumes were adjusted to 50.0 µL with the respective reagent blank buffer. All samples were treated with 1.0 µL of 10.0 mM Chymotrypsin inhibitor provided by the kit and incubated at room temperature for 10 minutes. After incubation, reaction mixes were prepared using 2.0 µL and trypsin substrate and 48.0 µL respective buffer. The absorbance of samples was measured with the UV-Vis spectrophotometer at 405 nm at 25 °C for 100 minutes with 2-minute intervals.

2.2.8 Peptide Assay Procedure

Digested BSA samples, which were dried and reconstituted to 300.0 µL with 0.1% FA, were analyzed with a Peptide Assay Kit. The calibration curve was drawn using the 1.0 mg/ml peptide stock solution provided by the kit. Ten data points were used, with three replicates for each point. Additionally, a quality control sample was added to each plate set to ensure the equivalency of the sets. Detailed information about QCs is given below. Blanks, protein samples, quality control samples, and standard samples were prepared according to the kit procedure. Initially, a working reagent was prepared by mixing Reagents A, B, and C in 50:48:2 ratios. Then 20.0 µL of each sample was added to a 96-well microplate. A 180.0 µL of working reagent was added to the samples. Plates were covered immediately and incubated at 37 °C for 15 minutes. Then, the absorption values of the samples were measured at 480 nm. Additionally, LOD and LOQ calculations were made using the measurements of ten blank samples. LOD was calculated by dividing the standard deviation of ten blank samples to slope of calibration curve of peptide assay and multiplying the results by 3.3. LOQ was calculated by multiplying the result of division with 10 instead of 3.3.

2.2.9 Mass Spectrometry based Protein/Peptide Analysis

LC-MS/MS analysis were performed using Thermo Scientific™ UltiMate™ 3000 RSLC Ultra Nano ultra-performance liquid chromatography and Thermo Scientific Q Exactive™ HF mass spectrometer equipped with Thermo Scientific™ EasySpray™ ion source. Before starting the analysis, the instrument was calibrated. The samples, tryptic peptides, were separated by reverse-phase liquid chromatography according to their hydrophobicity before MS/MS analysis. A 5.0 µL of each sample was injected. Every sample was injected and analyzed twice. Peptides were retained in the trap column (C18 PepMap100, 5 µm, 100A, 300 µm i.d x 5mm) and then separated from the trap column by the acetonitrile gradient

created by the nano pump. These peptides were separated by C18 (Easy-Spray ES902 RSLC C18, 2 µm, 100A 75 µm x 25 cm) analytical column (column temperature 40 °C) with hydrophobic interactions by reverse phase chromatography for 50 min at a flow rate of 350 nL/min. 1000 ng of peptide was loaded onto a 250 mm Nano LC column. Mobile Phase A and Mobile Phase B were prepared as 0.1%FA, 98:2% H_2O: ACN, and 0.1%FA 98:2% ACN: H_2O, respectively. If the total gradient time is specified as x for all gradients, phase B increased from 3% to 8% in 0.09x time, to 24% in 0.76x time, and to 36% in 0.15x time. The flow rate was kept at 350 nL/min. Re-equilibration was performed by washing with 3% B solution for 10 min.

The Q Exactive HF-X instrument (Thermo Fisher Scientific, Bremen, Germany) was cleaned on the same day and calibrated with the Tune (Version 2.9) instrument control software. The spray voltage was set to 2 kV, funnel RF level to 50, and capillary temperature to 270 °C. The device was set to "Full MS/DD−MS/MS" configuration to perform DDA analysis. Full MS resolution was set to 60000 for m/z 200. Full MS AGC target is set to $3*10^6$ with IT 45 ms. The mass range is defined as 350−1400 Da. The MS/MS AGC target value was set to 10^5, the intensity threshold was set to $2*10^4$, and the isolation width was set to 1.3 m/z. Normalization collision energy was set to 28%. All data were obtained in positive ion mode.

2.2.10 Data Analysis

Protein search was done with MaxQuant 2.4.13 Software and Proteome Discoverer 3.1. The setting was made as given in Table 2.1 and 2.2. Other parameters were used in default settings.

Table 2.1. MaxQuant Parameters

Search Parameters	Setting
Variable Modification	Oxidation (M) Acetylation (Protein N-Term)
Fixed Modifications	Carbamidomethyl (C)
Enzyme	Trypsin
Max Missed Cleavages	2
Instrument	Orbitrap
Min Peptide Length	8
Max Peptide Length	25
Peptides for Quantification	Unique

Table 2.2. Proteome Discoverer Parameters

Search Parameters	Setting
Variable Modification	Oxidation (M,H,W,F) Carbamidomethyl (K, N-Term) Deamidation (N,Q) Dethiomethylation (M) Carbamylation (K, R, N-Term) Acetylation (N-term)
Fixed Modifications	Carbamidomethyl (C)
Enzyme	Trypsin
Max Missed Cleavages	2
Instrument	Orbitrap
Min Peptide Length	8
Max Peptide Length	25
Peptides for Quantification	Unique

CHAPTER 3

RESULTS AND DISCUSSION

This thesis investigates the impact of different in-solution sample preparation procedures on protein products. These procedures differ in the denaturation, reduction, alkylation, and digestion steps found in the literature. Protein products are analyzed using mass spectrometry and compared in cleavage differences caused by different enzyme activities. Additionally, artificial modifications formed due to varying reagents employed in the digestion procedures were investigated. Furthermore, effects of digestion steps on proteins were evaluated step by step using spectrophotometric methods.

3.1 Workflow

Figure 3.1. shows a schematic representation of the workflow. Firstly, literature is search for in-solution digestion methods that are applied to human serum/plasma in biomarker studies. Then, digestion procedures from the scanned articles were filtered according to their frequency of use, employed reagents and their concentrations, incubation temperatures and durations. Ten methods representing the literature search were chosen accordingly. Chosen procedures were applied to A2MG protein, a clinically relevant blood protein[139-143-144], and products of this digestion procedures were analyzed using LC/MS. Sample protein was chosen as a single blood protein instead of whole human serum proteome in order to observe the effects of digestion methods without suppression due to multiple proteins in the human serum. Data analysis included the comparison of MS data in terms of cleavage efficiency and artificial modifications formed on peptides due to procedures. Addition to the investigation of effects of different digestion procedures on the protein products,

state of the protein during the procedure were investigated step by step. Here, solubility effect, denaturation effect, enzymatic activity, and total peptide formation yields of digestion methods were investigated using spectroscopic methods. Model protein was chosen as BSA, a common model protein used in proteomics studies[163–165], for spectroscopic analysis.[121–125] Four QC Sets were employed for commercial kits (BCA Assay and Peptide Assay). QC sets were coded as QC-BCA, QC-PA, QC-Δ, and QC-Reverse. QC-BCA and QC-PA sets were designed to observe possible interferences of the buffers (ABC and Tris), reagents (DTT, TCEP, IAA, CAA), and the enzymes (Trypsin, Trypsin/Lys-C mix) on the signals of the assays. QC-Δ set was designed to observe the possible signal alteration of BCA assay due to protein denaturation compared to native protein state. In this set, BCA assay was applied to native BSA protein and heat-denatured BSA proteins and results were compared using t-test in 95% confidence level. QC-Reverse set was designed to test the specificity of the Peptide assay for peptides. In this set, Peptide assay was applied to samples of peptide-free protein mediums and responses were investigated.

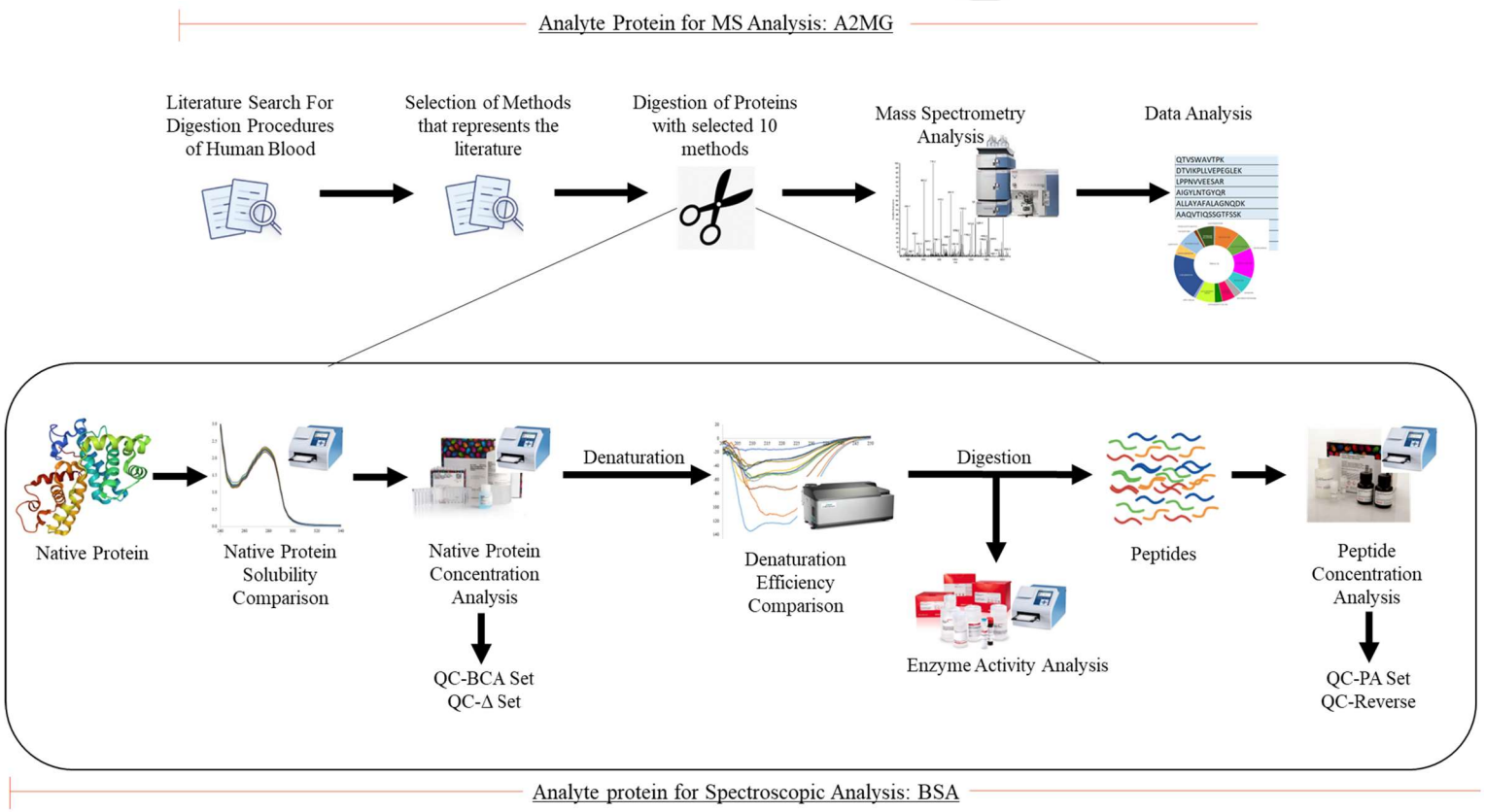


Figure 3.1. Workflow of this Study

3.2 Literature Search and Method Selection

Initially, literature was searched for in-solution digestion procedures in bottom-up proteomics based clinical studies. The search was conducted using the keywords "bottom-up proteomics, biomarker, in-solution digestion, clinical proteomics, validation, reproducibility, human serum, human plasma". To minimize the potential variations that might be caused by method adjustments due to different analytes, the search was restricted to methods that digest human plasma and serum. Over a hundred digestion methods were randomly chosen from the MS-based clinical studies. Table 3.1 shows the reagents and their concentration, incubation temperatures and durations of digestion processes found during the literature search. Each row of Table 3.1 represents a summarized digestion process and reference studies are given for each process. Procedures were filtered to determine their frequency of use. However, it was observed that only two of the methods were used multiple times the scanned articles. These methods were coded as Method A and Method C in Table 3.2. Thus, the frequency of use of reagents within the digestion method steps was considered. It is seen that Ammonium Bicarbonate (ABC) and Tris(hydroxymethyl)aminomethane (Tris) were the most used buffer reagents. At the denaturation step, two common reagents were observed which are urea and SDC alongside buffers. Urea concentrations were observed to vary within 6 M to 9 M whereas SDC constituted 1% to 10% of the denaturation solutions. Furthermore, some studies denatured proteins through heat without employing urea or detergents such as Method A. In the following step, which is the reduction step, used to break disulfide bonds, DTT and TCEP were commonly used in the literature. The concentration range of these reagents were observed to vary between 1 mM to 50 mM. In the alkylation step, IAA is observed to be used predominantly, with a concentration range of 7 mM to 75 mM. Additionally, CAA is also utilized in some methods such as Method C. Digestion with the trypsin enzyme was predominantly used in the searched studies, and trypsin/LysC enzyme mix was also observed.

Fourteen different enzyme-to-protein concentration ratios were seen, with the highest being 1:8 and the lowest being 1:152. Moreover, variations in incubation temperature and duration were abundant. Incubation temperatures included room temperature, 37 °C, 40 °C, 60 °C, 70 °C and 95 °C. Lastly, incubation durations for denaturation, reduction, and alkylation steps varied between 5 to 90 minutes. Meanwhile, the incubation for proteolytic digestion was done in a range of 1 hour to 18 hours in total. Additionally, it was observed that while reporting the concentration values of reagents, some studies reported concentrations of working solution while some reported the concentrations of stock solutions. Working solution concentrations were calculated if essential information was provided in the study. However, it's crucial to note that presenting the digestion procedure without clearly emphasizing the concentrations could lead to varied interpretations within the literature.

Table 3.1 Literature Search for In-Solution Digestion Products with Human Blood/ Plasma /Serum.
Literature was searched with keywords “clinical proteomics, biomarker, human blood, human plasma, human serum, bottom-up proteomics, and peptide centric approach” in order to find biomarker studies using human plasma and serum that applied peptide centric approach in their studies. Then digestion procedures used in these studies are collected and summarized in terms of reagent type and concentration, incubation temperature and duration. If a study were observed to not include the concentration of a regent, then only the name of the reagent was added to the table. Methods selected for the further steps of this study is written in bold.

DENATURATION				**REDUCTION**			**ALKYLATION**			**DIGESTION**			
Reagent	**Buffer**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Digestion (protein:enzyme)**	**Temp**	**Time**	**Ref**
9 M Urea	**300 mM Tris**	**-**	**-**	**12.9 mM TCEP**	**37 °C**	**30 min**	**34.6 mM IAA**	**-**	**30 min**	**Trypsin 15:1**	**37 °C**	**17 h**	160
1% SDC	**100 mM Tris**	**-**	**-**	**10 mM TCEP**	**-**	**-**	**40 mM CAA**	**95 °C**	**5 min**	**Lys-C/Trypsin 100:1**	**37 °C**	**1 h**	151–155
0.2% SDS	ABC	-	-	5 mM DTT	60 °C	20 min	20 mM IAA	RT	30 min	Trypsin	37 °C	3 h	166
8 M Urea	50 mM ABC	37 °C	60 min	10 mM DTT	37 °C	30 min	40 mM IAA	RT	90 min	Trypsin 50:1	37 °C	Overnight	167
10% SDS	100 mM bicine	-	-	9.4 mM DTT	70 °C	10 min	36.4 mM IAA	RT	45 min	Chymotrypsin 40:1	30 °C	50 min	168
0.1% RapiGest	100 mM bicine	-	-	9.4 mM DTT	70 °C	10 min	36.4 mM IAA	RT	45 min	Trypsin 40:1	37 °C	50 min	168
0.1% RapiGest	100 mM bicine	-	-	9.4 mM DTT	70 °C	10 min	36.4 mM IAA	RT	45 min	Lys-C 20:1	37 °C	50 min	168
0.1% RapiGest	100 mM bicine	-	-	9.4 mM DTT	70 °C	10 min	36.4 mM IAA	RT	45 min	Chymotrypsin 40:1	30 °C	50 min	168
Methanol	50 mM ABC	-	-	1.25 mM DTT	60 °C	30 min	2.65 mM IAA	RT	45 min	Trypsin	37 °C	overnight	169
8 M Urea	50 mM ABC	-	-	11.45 mM DTT	-	20 min	46.58 mM IAA	-	20 min	Trypsin 25:1	-	-	170
-	100 mM ABC	Boil	5 min	-	-	-	-	-	-	Trypsin 50:1	37 °C	15 h+5 h	25,171,172
1% SDC	0.1 M TEAB	-		10 mM TCEP	-	-	40 mM CAA	95 °C	10 min	Trypsin	37 °C	16 h	173

Table 3.1 Cont`d

DENATURATION				REDUCTION			ALKYLATION			DIGESTION			
Reagent	**Buffer**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Digestion**	**Temp**	**Time**	**Ref**
-	50 mM ABC	-	-	-	-	-	-	-	-	Trypsin	37 °C	18 h	[174]
-	100mM Tris			10 mM TCEP	-	-	40 mM CAA	95 °C	10 min	Trypsin/Lys-C 40:1	37 °C	2 h	[175]
1% HNO3	-	-	-	-	-	-	-	-	-	Pronase	37 °C	18 h	[176]
-	-	-	-	-	-	-	-	-	-	concentrated HNO3	250 °C	30 min	[177]
-	-	-	-	-	-	-	-	-	-	concentrated HNO3	250 °C	30 min	[174]
-	-	-	-	-	-	-	-	-	-	con HNO3	200 °C	7 min	[178]
-	-	-	-	-	-	-	-	-	-	con HNO3	200 °C	7 min	[178]
6 M guanidine	-	-	-	10 mM DTT	95 °C	30 min	20 mM IAA	-	40 min	Trypsin 25:1	37 °C	18 h	[179]
6 M guanidine	100 mM Tris	-	-	20 mM DTT	60 °C	30 min	28.6 mM IAA	37 °C	30 min	Trypsin 8:1	37 °C	15 h	[180]
6 M guanidine	100 mM Tris	-	-	20 mM DTT	60 °C	30 min	28.6 mM IAA	37 °C	30 min	Trypsin 16:1	37 °C	15 h	[180]
6 M guanidine	100 mM Tris	-	-	20 mM DTT	60 °C	30 min	28.6 mM IAA	37 °C	30 min	Trypsin 32:1	37 °C	15 h	[180]
6 M guanidine	100 mM Tris	-	-	20 mM DTT	60 °C	30 min	28.6 mM IAA	37 °C	30 min	Trypsin 64:1	37 °C	15 h	[180]
6 M guanidine	100 mM Tris	-	-	20 mM DTT	60 °C	30 min	28.6 mM IAA	37 °C	30 min	Trypsin 128:1	37 °C	15h	[180]
-	-	-	-	-	-	-	-	-	-	con HNO3	220 °C	15min	[181]
6 M urea	**100 mM Tris**	**-**	**-**	**20 mM DTT**	**37 °C**	**30 min**	**40 mM IAA**	**RT**	**30 min**	**Trypsin 50:1**	**37 °C**	**18 h**	[150]

Table 3.1 Cont`d

DENATURATION				REDUCTION			ALKYLATION			DIGESTION			
Reagent	**Buffer**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Digestion (protein:enzyme)**	**Temp**	**Time**	**Ref**
9 M urea	300 mM Tris	-	-	20 mM DTT	37 °C	30 min	40 mM IAA	RT	30 min	Trypsin 50:1	37 °C	18 h	[150]
6 M urea	100 mM Tris	-	-	20 mM DTT	37 °C	30 min	40 mM IAA	RT	30 min	Trypsin 50:1	37 °C	18 h	[182]
9 M urea	300 mM Tris	-	-	20 mM DTT	37 °C	30 min	40 mM IAA	RT	30 min	Trypsin 50:1	37 °C	18 h	[182]
10% SDC	25 mM ABC	-	-	5 mM TCEP	60 °C	30 min	10 mM IAA	37 °C	30 min	Trypsin 20:1	37 °C	16 h	[182]
5% SDC	**22.7 mM ABC**	**-**	**-**	**5 mM TCEP**	**60 °C**	**30 min**	**10 mM IAA**	**37 °C**	**30 min**	**Trypsin 20:1**	**37 °C**	**16 h**	[156]
9 M Urea	300 mM Tris	-	-	13.3 mM DTT	37 °C	30 min	40 mM IAA	RT	30 min	Trypsin 20:1	37 °C	18 h	[183]
0.1 M TFE	0.1 M ABC	RT	30 min	4 mM TCEP	RT	60 min	-	-	-	Trypsin	37 °C	16h	[184]
50% (v/v) TFE	-	37 °C	30 min	TCEP	-	-	IAA	-	-	Trypsin 25:1	-	-	[185]
-	50 mM ABC	-	-	10 mM DTT	60 °C	30 min	20 mM IAA	RT	30 min	Trypsin	37 °C	overnight	[186]
10 M urea	-	-	-	5 mM DTT	-	-	15 mM IAA	RT	30 min	Trypsin 50:1	37 °C	overnight	[187]
6 M urea	200 mM TrisHCl	37 °C	30 min	13 mM DTT	-	-	40 mM IAA	RT	30 min	Trypsin 20:1	37 °C	18 h	[161]
10% SDC	25 mM ABC	60 °C	30 min	5 mM TCEP	-	-	10 mM IAA	37 °C	30 min	Trypsin 20:1	37 °C	16 h	[188]
6 M urea	100 mM Ammonium Acetate	37 °C	30 min	18.2 mM DTT	-	-	56.8 mM IAA	RT	30 min	Trypsin 10:1	37 °C	22 h	[189]
-	50 mM ABC	-	-	5 mM DTT	60 °C	30 min	10 mM IAA	RT	30 min	Trypsin 50:1	37 °C	17 h	[190]
8 M Urea	50 mM ABC	-	-	10 mM DTT	60 °C	30 min	-	-	-	Trypsin 50:1	37 °C	5 h	[191]
8 M Urea	**100 mM ABC**	**-**	**-**	**45.8 mM DTT**	**60 °C**	**30 min**	**7.7 mM IAA**	**23°C**	**30 min**	**Trypsin 152:1**	**37 °C**	**17 h**	[157]

Table 3.1 Cont`d

DENATURATION				REDUCTION			ALKYLATION			DIGESTION			
Reagent	**Buffer**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Digestion (protein:enzyme)**	**Temp**	**Time**	**Ref**
10% SDC	25 mM ABC	-	-	5 mM TCEP	60 °C	30 min	10 mM IAA	37 °C	30 min	Trypsin 50:1	37 °C	16 h	[192]
10% SDC	25 mM ABC	-	-	5 mM TCEP	60 °C	30 min	10 mM IAA	37 °C	30 min	Trypsin 20:1	37 °C	16 h	[193]
2 M Urea	100 mM TEAB	-	-	9.5 mM TCEP	RT	60 min	9.6 mM IAA	RT	30 min	Trypsin	RT	Overnight	[194]
10% SDC	25 mM ABC	-	-	5 mM TCEP	60 °C	30 min	10 mM IAA	37 °C	30 min	Trypsin	37 °C	16 h	[195]
6 M Urea	5 mM ABC	-	-	15 mM DTT	60 °C	40 min	35 mM IAA	RT	40 min	Trypsin	37 °C	4 h	[196]
6 M Urea	5 mM ABC	-	-	15 mM DTT	60 °C	40 min	35 mM IAA	RT	40 min	Trypsin	37 °C	4 h	[196]
RapiGest	1 mM ABC			13.3 mM DTT	90 °C	15 min	25 mM IAA	RT	30 min	Trypsin	37 °C	90 min	[197]
-	40 mM ABC	-	-	-	-	-	-	-	-	Trypsin 20:1	37 °C	5h	[198]
-	40 mM ABC	-	-	-	-	-	-	-	-	Trypsin 100:1	37 °C	5h	[198]
-	40 mM ABC	-	-	-	-	-	-	-	-	Trypsin 1000:1	37 °C	5h	[198]
-	40 mM ABC	-	-	-	-	-	-	-	-	Trypsin 20:1	30 °C	5h	[198]
-	40 mM ABC	-	-	-	-	-	-	-	-	Trypsin 20:1	23 °C	5h	[198]
-	Sodiumcitrate	-	-	-	-	-	-	-	-	Trypsin 20:1	37 °C	5h	[198]
-	40 mM ABC	-	-	-	-	-	-	-	-	Trypsin 20:1	37 °C	5h	[198]
-	40 mM ABC	37 °C	1.5h	-	-	-	-	-	-	Trypsin/Lys-C 20:1	37 °C	3.5h	[198]

Table 3.1 Cont`d

DENATURATION				REDUCTION			ALKYLATION			DIGESTION			
Reagent	**Buffer**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Digestion (protein:enzyme)**	**Temp**	**Time**	**Ref**
-	40 mM ABC	-	-	2 mM TCEP	-	30 min	10 mM IAA	-	30 min	Trypsin 20:1	37 °C	5h	198
4 M urea	40 mM ABC	37 °C	3h	-	-	-	-	-	-	Trypsin/Lys-C 20:1	37 °C	21h	198
Rapigest 0.1%	40 mM ABC	37 °C	-	-	-	-	-	-	-	Trypsin 20:1	37 °C	5h	198
Rapigest 0.1%	40 mM ABC	100 °C	5h	-	-	-	-	-	-	Trypsin 20:1	37 °C	5h	198
10% SDC	25 mM ABC	RT	5 min	4.5 mM TCEP	-	-	10 mM IAA	37 °C	30 min	Trypsin	37 °C	1 h	159
5% SDC	**15 mM ABC**	**RT**	**5 min**	**4.5 mM TCEP**	**-**	**-**	**10 mM IAA**	**37 °C**	**30 min**	**Trypsin**	**37 °C**	**2 h**	159
10% SDC	25 mM ABC	RT	5 min	4.5 mM TCEP	-	-	10 mM IAA	37 °C	30 min	Trypsin	37 °C	4 h	159
10% SDC	25 mM ABC	RT	5 min	4.5 mM TCEP	-	-	10 mM IAA	37 °C	30 min	Trypsin	37 °C	8 h	159
10% SDC	25 mM ABC	RT	5 min	4.5 mM TCEP	-	-	10 mM IAA	37 °C	30 min	Trypsin	37 °C	12 h	159
10% SDC	25 mM ABC	RT	5 min	4.5 mM TCEP	-	-	10 mM IAA	37 °C	30 min	Trypsin	37 °C	24 h	159
8 M urea	0.1 mM ABC	-	-	1.67 mM TCEP	37 °C	60 min	15 mM CAA	37 °C	30 min	Glu-C 50:1	37 °C	Overnight	199
8 M urea	0.1 mM ABC	-	-	1.67 mM TCEP	37 °C	60 min	15 mM CAA	37 °C	30 min	Glu-C 100:1	37 °C	Overnight	199
8 M urea	0.1 mM ABC	-	-	1.67 mM TCEP	37 °C	60 min	15 mM CAA	37 °C	30 min	Lys-C 50:1	37 °C	Overnight	199
8 M urea	0.1 mM ABC	-	-	1.67 mM TCEP	37 °C	60 min	15 mM CAA	37 °C	30 min	Lys-C 100:1	37 °C	Overnight	199
-	Boiling Water	-	-	5 mM DTT	60 °C	60 min	15 mM IAA	RT	30 min	Trypsin 50:1	37 °C	Overnight	200
SDS 1%	50 mM TEAB	50 °C	20 min	18.2 mM TCEP	55°C	60 min	31.3 mM IAA	40 °C	60 min	Trypsin 50:1	37 °C	Overnight	201

Table 3.1 Cont`d

DENATURATION				REDUCTION			ALKYLATION			DIGESTION			
Reagent	**Buffer**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Reagent**	**Temp**	**Time**	**Digestion (protein:enzyme)**	**Temp**	**Time**	**Ref**
6 M urea	**10 mM Tris**	**-**	**-**	**30 mM DTT**	**40 °C**	**40 min**	**50 mM IAA**	**RT**	**40 min**	**Trypsin 30:1**	**37 °C**	**Overnight**	158
6 M urea	-	-	-	3.04 mM DTT	-	-	11.3 mM IAA	-	-	Trypsin 50:1	37 °C	Overnight	202
-	50 mM ABC	-	-	18.3 mM DTT	60 °C	30 min	46.6 mM IAA	RT	30 min	Trypsin 50:1	37 °C	Overnight	203
-	**50 mM ABC**	**-**	**-**	**32.5 mM DTT**	**60 °C**	**30 min**	**75 mM IAA**	**RT**	**30 min**	**Trypsin 50:1**	**37 °C**	**Overnight**	146–149
-	50 mM ABC	-	-	10 mM DTT	60 °C	30 min	10 mM IAA	RT	30 min	Trypsin 50:1	37 °C	Overnight	204
6 M urea	50 mM ABC	-	-	20 mM DTT	37 °C	30 min	50 mM IAA	RT	30 min	Lys-C 50:1+Trypsin 50:1	30°C + 37 °C	Overnight	205

Table 3.2. Selected Methods from Literature. Ten methods were coded Method A to J for ease of reference.

	DENATURATION				REDUCTION			ALKYLATION				DIGESTION			
	Reagent	**Buffer**	**Temp**	**Time (min)**	**Reagent**	**Temp**	**Time (min)**	**Reagent**	**Temp**	**Time (min)**	**Dilution**	**Reagent (prt:enzyme)**	**Temp**	**Time (h)**	**Ref**
Method A	-	50 mM ABC	-	-	32.5 mM DTT	60 °C	30	75 mM IAA	RT	30	-	Trypsin 50:1	37 °C	17	146–149
Method B	6 M Urea	100 mM Tris	-	-	20 mM DTT	37 °C	30	40 mM IAA	RT	30	100 mM Tris, pH 8.0	Trypsin 50:1	37 °C	18	150
Method C	1% SDC	100mM Tris	-	-	9.6 mM TCEP	-	-	38.4 mM CAA	95 °C	10	-	Lys-C/Trypsin	37 °C	1	151–155
Method D	5% SDC	22.7 mM ABC	-	-	5 mM TCEP	60 °C	30	10 mM IAA	37 °C	30	ABC Buffer	Trypsin 20:1	37 °C	16	156
Method E	7.3 M Urea	100 mM ABC	-	-	45.8 mM DTT	30 °C	60	7.7 mM IAA	23 °C	30	100 mM ABC Buffer	Trypsin 152:1	37 °C	17	157
Method F	6 M Urea	10 mM Tris	-	-	30 mM DTT	40 °C	40	50 mM IAA	RT	40	Distilled Water	Trypsin 30:1	37 °C	18	158
Method G	5% SDC	15 mM ABC	RT	5	5 mM TCEP	-	-	10 mM IAA	37 °C	30	-	Trypsin 20:1	37 °C	2	159
Method H	9 M Urea	300 mM Tris	-	-	12.9 mM TCEP	37 °C	30	34.6 mM IAA	-	30	100 mM Tris, pH 8.0	Trypsin 15:1	37 °C	17	160
Method I	6 M Urea	200 mM TrisHCl	-	-	13 mM DTT	37 °C	30	40 mM IAA	RT	30	-	Trypsin 20:1	37 °C	18	161
Method J	8 M Urea	100 mM Tris	-	-	200 mM TCEP	60 °C	30	FASP KIT				Trypsin 100:1	37 °C	16	162

Ten methods that represented the literature search were chosen as given in Table 3.2. Methods were coded alphabetically for ease of reference. Method A[146–149] was one of the selections with a high frequency of use as mentioned above. The reference studies to this method focused on the reproducibility of the results and developed Multiple-reaction Monitoring Mass Spectrometry techniques to be used in biomarker discovery and validation. Method A used ABC as the buffer and utilized no additional reagent in denaturation step. Reduction and alkylation reagents were the most common combination with DTT and IAA. Additionally, digestion enzyme, temperature, and duration represented the common parameters in the literature search.

Method B[150] was a validated method. In the reference study, Prostate-specific antigen (PSA) and C-reactive protein were chosen as target proteins. PSA was analyzed using two unique peptides, and CRP was analyzed with three unique peptides. Samples were prepared by the National Institute of Standards and Technology (NIST) and delivered to 8 different laboratory groups, where the samples were subjected to the Method B digestion procedure. The method involved Tris Buffer, the denaturation reagent that is Urea, DTT-IAA couple in reduction and alkylation steps, and trypsin as the enzyme.

Method C[151–155] was another highly referenced procedure in the literature and focused on the reproducibility of the results. Tris Buffer and SDC as the denaturant were utilized alongside a different reduction-alkylation couple, that is, TCEP and CAA. Additionally, the method involved a combined single incubation in boiling water for 10 minutes. Additionally, trypsin/Lys-C enzyme mixture was used rather than trypsin alone with the shortest proteolytic digestion time.

Reference study to Method D[156] developed MRM techniques for biomarker discovery. This study used TCEP as a reducing agent and IAA as an alkylation agent. Notable differences were that the method used one of the lowest reduction and alkylation reagent concentrations and one of the highest enzyme-to-protein ratios.

Method E[157] was a study that focused on the early diagnosis of COVID-19, and it

sampled 113 proteins in human plasma. It is the only method incubated at 30°C for reduction and 23°C for alkylation. This method uses the lowest enzyme to protein ratio among the literature search.

The reference study to Method F[158] is another developing MRM technique for biomarker discovery studies. The method was another urea-containing procedure with the DTT-IAA couple. However, the main difference of Method F was the utilization of distilled water after the initial denaturation step rather than the buffer. While other methods prepared reagents in and made dilutions with buffer solutions, Method F used distilled water, leading to a non-buffer environment when the trypsin enzyme was added.

Method G[159] was one of the validated methods. Four biomarker proteins were targeted in this study, with them being C3 (Complement 3), adiponectin, angiotensinogen, and RBP4 (Retinol Binding Protein 4). Each of the proteins was analyzed through one of their unique peptides. This method employed SDC as the denaturant in the ABC buffer. A different incubation pattern was observed with an incubation right after the denaturation step rather than common combined denaturation and reduction incubation. Next, incubation in the method was done after the simultaneous addition of reduction reagent, TCEP, and alkylation reagent, IAA. Additionally, it has one of the shorter proteolytic digestion durations among methods.

Method H[160] is another clinical study focusing on the high-sensitivity quantitative analysis of 13 high-sensitivity quantitative. This study further marks the first application of stable isotope-labeled recombinant protein fragments internal standards (SIS PrEST) on clinical samples. It employed the highest Tris and Urea concentrations and the highest enzyme-to-protein ratio among them.

Method I[161] was another MRM developing study that employed a higher enzyme-to-protein ratio while utilizing a DTT-IAA reduction-alkylation reagent couple.

The last method chosen, Method J, was a FASP kit method widely used in the literature.[206–210] It provided reagents for denaturation as urea and alkylation as IAA that were instructed to use subsequently. However, the kit suggested that cystine bridges of the protein samples to be reduced prior to loading the protein into the kit. This approach changed the arrangement of the digestion steps. The kit didn't provide or recommended a procedure for the reduction of cystine bonds, so a study from literature was referenced that uses same FASP Kit.[162] Combination of the reduction step from the reference and kit procedure formed Method J.

3.3 Investigation of Digestion Method

3.3.1 Native Protein Analysis/ Solubility Comparison

Solubility differences of buffers used in selected ten methods were compared in this section. Firstly, six experimental replicates of BSA protein dissolved in respective buffers of the chosen methods were qualitatively analyzed with UV-Vis spectroscopy. Then samples were qualitatively analyzed using BCA Protein Assay kit.

3.3.1.1 Solubility Comparison via Absorption Spectroscopy

Absorption spectra of proteins with a maxima point ~280 nm is a widely used qualitative analysis.[59] Proteins absorb the light ~280 nm due to residues with aromatic groups, mainly tryptophan.[55–57] Absorption spectra of 1 mg/ml native BSA protein in buffers of selected digestion procedures (For Method A, buffer was 50.0 mM ABC. For Methods B, C, and J buffer was 100.0 mM Tris. For Method D, buffer was 22.7 mM ABC. For Method E, buffer was 100.0 mM ABC. For Method F, buffer was 10 mM Tris, For Method G, buffer was 15.0 mM ABC. For Method H, buffer

was 300 mM Tris. For Method I, buffer was 200 mM TrisHCl.) were given in Figure 3.3.

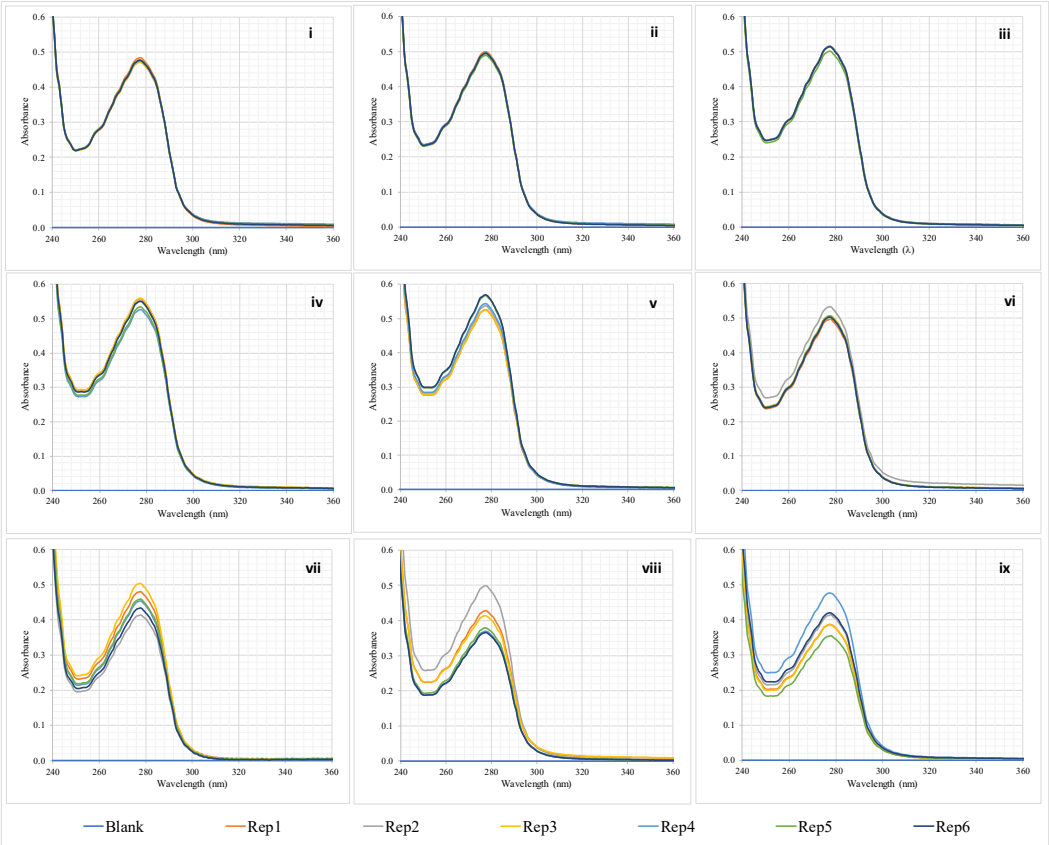


Figure 3.3. Absorption Spectrum of six experimental replicates of native BSA with peak maxima at 277 nm in (i)Deionized Water (QC) (ii) 15.0 mM ABC Buffer (Method G) (iii) 22.7 mM ABC Buffer (Method D) (iv) 50.0 mM ABC Buffer (Method A) (v)100.0 mM ABC Buffer (Method E) (vi) 10.0 mM Tris Buffer (Method F) (vii) 100.0 mM Tris Buffer (Methods B, C, J) (viii) 200.0 mM TrisHCl Buffer (Method I) (ix) 300 mM Tris Buffer (Method H)

In order to compare the relative solubilities of the BSA protein in the different buffers employed in selected digestion methods, average absorbance values of experimental replicates were calculated. Calculation was done by taking the average of six experimental replicates at each point of the spectrum and creating a representative

spectrum for each buffer. These representative spectrum for each buffer and DI water were given in Figure 3.4.

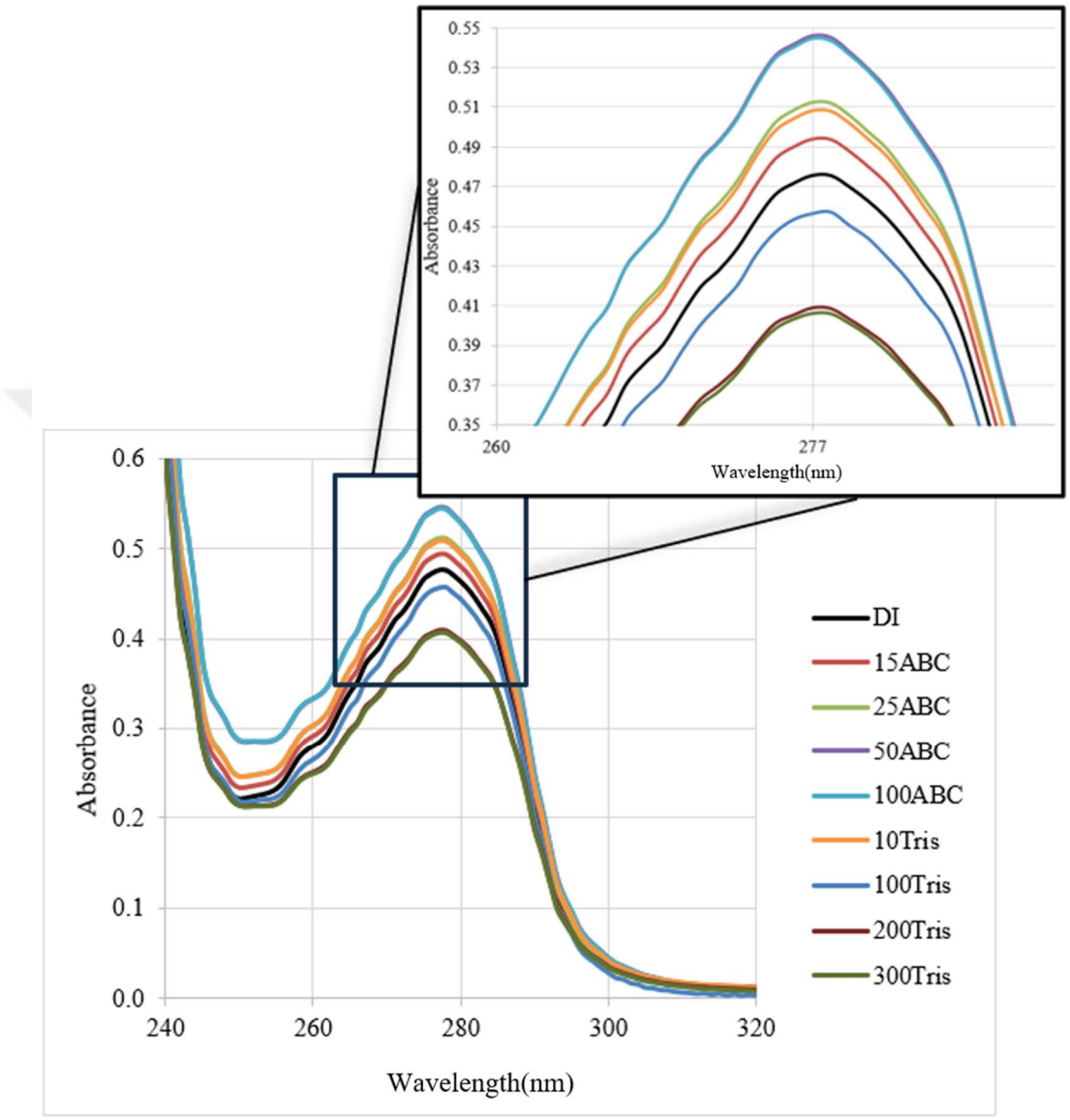


Figure 3.4. Average absorbance values of experimental replicates were calculated to compare BSA protein solubilities in selected digestion buffers.
The average of six experimental replicates at each spectrum point was used to calculate a representative spectrum for each buffer. Zoomed in version to the peaks were added on top.

Figures 3.3 and 3.4 showed peak maxima at 277 nm for BSA samples which is in agreement with the literature.[137] Figure 3.3. showed the absorption spectra of BSA samples dissolved in buffers of selected methods and in deionized water. Each spectra included six 1 mg/ml BSA replicates. Replicate precisions were observed to vary in different buffers.

$$A = -\log(T)$$

Figure 3.5. Equation to calculate transmittance percent (%T) from absorbance (A)

Percent transmittance values of peak maximas (λ=277 nm) of BSA replicates were calculated according to Figure 3.5 and $\%T_{max}$ and $\%T_{min}$ values in six replicates were given in Table 3.3.

Table 3.3. $\%T_{max}$ and $\%T_{min}$ values of six BSA replicates for each buffer and deionized water mediums at 277 nm. Difference between these values was given under $\%T_{max-min}$. $\%T_{average}$ values are calculated by averaging the percent transmittance values of each replicate in respective buffer medium.

Buffer	**Wavelength**	**$\%T_{max}$**	**$\%T_{min}$**	**$\Delta\%T_{max-min}$**	**$\%T_{average}$ (of six replicates)**
DI	λmax (277 nm)	33.9%	32.9%	1.0%	33.5%
15 mM ABC		32.4%	31.7%	0.8%	32.1%
23 mM ABC		31.5%	30.5%	1.0%	30.7%
50 mM ABC		29.8%	27.6%	2.2%	28.4%
100 mM ABC		29.8%	27.0%	2.9%	28.5%
10 mM Tris		31.9%	29.3%	2.6%	31.0%
100 mM Tris		38.7%	31.4%	7.2%	35.0%
200 mM Tris		43.1%	31.7%	11.4%	39.0%
300 mM Tris		44.3%	33.4%	10.9%	39.3%

Precision of replicates in lower concentrations of ABC buffer (15 mM, 23 mM) weren't observed to be negatively affected compared to precision of BSA samples in deionized water. Increase in the ABC concentration led to an increase of

$\%T_{max-min}$ from 1.0% to 2.9% as seen in Table 3.3. Whereas Tris-containing buffers exhibited a more pronounced impact on the precision of replicate measurements compared to buffers containing ABC. $\%T_{max-min}$ difference was observed to reach 11.4% for protein samples in Tris buffers.

Figure 3.4. showed the absorbance spectrum of average BSA sample spectra in buffers and deionized water. Table 3.3. presents the percent transmittance peak maximas in Figure 3.4. Increase in the absorption, hence, decrease in the transmittance were observed as the concentration of ABC buffer increased until 50 mM ABC buffer. Transmittance values of 50 mM and 100 mM ABC were observed to be similar with 28.4% and 28.5%, respectively. Here, we can deduce that ABC buffer increased the solubility of the BSA until 50 mM and increase from this concentration didn't affect the protein solubility. A dissimilar trend was observed for Tris buffers compared to ABC buffers. BSA samples in 10 mM Tris showed lower transmittance by 2.5%. However, as Tris concentration increased from 10 mM to 300 mM, transmittance increased to 39.3%, decreasing the absorbance values lower than DI samples. This trend can be explained by salting in and salting out effects of buffers. In literature, it is stated that reagents may create salting in and salting out effects on proteins.[211] ABC buffer were observed to create a salting in effect on proteins until 50 mM. It is also added that while lower concentrations of a reagent may create salting in affect, higher concentrations of same reagent may cause salting out.[211] This phenomenon explains the increase in solubility degree in 10 mM Tris buffer compared to DI by salting in effect and the decrease of it in rest of the higher concentrations by salting out effect. In conclusion, it was observed that choice of buffer effects the precision of replicates and degree of solubility of BSA protein.

3.3.1.2 Solubility Comparison via BCA Analysis

Absorption spectra of protein was used in qualitative analysis. Here, protein solubility was measured quantitatively.

3.3.1.2.1 QC Sets for BCA Assay

A commercial BCA Assay Kit was employed for the quantitative analysis. Before analyzing the samples, two quality control sets were design to observe the possible interference of buffer medium and the effect of state (native/ denatured) of the protein on the assay results.

3.3.1.2.1.1 QC-BCA Set

QC-BCA was designed to observe the possible interferences of buffers, reagents, and enzymes on the BCA assay results. All blank corrections were done according to the blank signal of deionized water to investigate the possible interferences of buffers. Table 3.4. shows the absorption values of QC samples at 562 nm. In "Buffer (Blank)" column, buffer reagents of selected methods were tested for interference. In "Reagent Blank" reduction and alkylation reagents were added to the buffers with respect to their digestion procedure given in Table 3.2. and incubation were done accordingly. In "Reagents & Enzyme", enzyme was added to the reagent blanks in the amount used in their respective procedures given in Table 3.2, then samples were incubated as given in their respective procedures. To obtain BCA signals, samples were mixed with BCA Kit reagents according to BCA Kit procedure and incubated at 37 °C for 15 minutes. Then absorbance of the samples was measured at $\lambda = 562$ nm as kit instructs.

Table 3.4. Quality Control Set for Investigation of BCA Assay (QC-BCA)
These control sets were done to observe if buffer reagents used during digestion procedures interfere with the BCA analysis. Absorption reading was done at 562 nm. (DI= Deionized Water)

	Absorbance Signals of QCs for BCA Kit (QC-BCA)	
	Buffer Samples (Corrected according to DI)	
Method	**Replicate 1**	**Replicate 2**
DI	0.00	0.00
A	0.01	0.01
B	0.03	0.03
C	0.03	0.03
D	0.00	0.00
E	0.00	0.00
F	0.01	0.01
G	0.00	0.00
H	0.05	0.05
I	0.06	0.06
J	0.01	0.01

Table 3.4. indicated that the buffers used in BCA Assay didn't show any interference at 562 nm.

3.3.1.2.1.2 QC-Δ Set

QC-Δ was designed to investigate the effect of denaturation on BCA Assay response. The cupric ion reduction in the Biuret Reaction was due to the protein backbone structure and the side groups of residues. The denaturation process increases the interaction of the inner part of the folded protein with the environment. Therefore, with QC-Δ set, possible increase of the BCA respond was investigated. QC- Δ set started with addition of native BSA protein to the respective buffer of each method to reach a final protein concentration of 1.5 mg/ml. BCA analysis was done for these native protein samples with two replicates. Protein samples were incubated for 30 minutes at 60 °C. After incubation, heat denatured samples were again analyzed with

BCA assay. Both absorbance measurements were done at λ = 562 nm. Then, protein concentration result was calculated using the respective calibration curve of buffers for BCA analysis. As an additional QC sample, the same application was performed in deionized water. Results were given in Table 3.5.

Table 3.5. Quality Control Set for Investigating BCA Assay Results Before and After Thermal Incubation (QC- Δ)
This set was done to observe the effect of denaturation on the results of BCA assay.

	Concentrations of Thermal Incubation QCs (QC-Δ)			
	BCA Assay			
	Native Protein Concentration (mg/ml)		**Protein Assay Results after Thermal Incubation (mg/ml)**	
Method	**Replicate 1**	**Replicate 2**	**Replicate 1**	**Replicate 2**
DI	1.2	1.3	1.4	1.4
A	1.2	1.4	1.4	1.5
B	1.2	1.3	1.3	1.4
C	1.3	1.2	1.2	1.3
D	1.3	1.4	1.3	1.4
E	1.3	1.2	1.2	1.5
F	1.4	1.5	1.4	1.5
G	0.8	0.9	0.8	0.9
H	0.8	0.9	0.8	0.8
I	1.2	1.1	0.8	0.9
J	1.3	1.3	1.4	1.4

The results were normalized according to the native protein results for every method and results were given in Table 3.6. In order to investigate the change in the signal due to denaturation was statistically significant change, Student's t-test was applied. Equation given in Figure 3.6. was used to calculate the experimental t_{calc} values of native protein concentration and concentration value after thermal incubation. t_{crit} value was taken as 4.30 in 95% confidence level at degrees of freedom of 2 since there are two groups with two replicates at each.[212]

$$t = \frac{\bar{x}_2 - \bar{x}_1}{s_{pooled} * \sqrt{\frac{N_2 + N_1}{N_2 * N_1}}}$$

Figure 3.6. Student's t-test for calculating differences in means of two independent groups[212]

$\bar{x}_1$ and $\bar{x}_2$ represents the means of two groups which are native protein concentration and protein assay results after thermal incubation. s_{pooled} refers to the pooled standard deviation values of these groups. N represents the number of replicates in groups.

Table 3.6. Normalized Values for QC- Δ.
Then CV values were calculated by diving the standard deviation to average value of their respective methods. For t value calculation, Student's t-test was used and resulted values are given in t_{calc} column. t_{crit} value was found as 4.30 at 95% confidence level with degrees of freedom of two. It was found that all t_{calc} values are smaller than t_{crit}.

	Normalized Concentrations of Thermal Incubation QCs (QC-Δ)								
	BCA Assay								
	Native Protein Concentration (mg/ml)			**Protein Concentration after Thermal Incubation (mg/ml)**			**t-test**		
Method	$\bar{x}_1$	s_1	CV_1	$\bar{x}_2$	s_2	CV_2	t_{calc}	$t_{crit, 0.05}$	Relation
DI	**1.0**	0.071	7.1%	**1.1**	0.045	4.0%	1.7	4.3	$t_{crit} > t_{calc}$
A	**1.0**	0.063	6.3%	**1.1**	0.040	3.6%	1.9		$t_{crit} > t_{calc}$
B	**1.0**	0.052	5.2%	**1.1**	0.046	4.3%	2.0		$t_{crit} > t_{calc}$
C	**1.0**	0.005	0.47%	**1.0**	0.048	4.6%	0.0		$t_{crit} > t_{calc}$
D	**1.0**	0.066	6.6%	**1.0**	0.027	2.8%	0.0		$t_{crit} > t_{calc}$
E	**1.0**	0.075	7.5%	**1.1**	0.14	13%	0.9		$t_{crit} > t_{calc}$
F	**1.0**	0.038	3.8%	**1.0**	0.075	7.3%	0.0		$t_{crit} > t_{calc}$
G	**1.0**	0.073	7.3%	**1.0**	0.086	8.5%	0.0		$t_{crit} > t_{calc}$
H	**1.0**	0.033	3.3%	**1.0**	0.059	6.1%	0.0		$t_{crit} > t_{calc}$
I	**1.0**	0.053	5.3%	**1.0**	0.061	7.8%	0.0		$t_{crit} > t_{calc}$
J	**1.0**	0.019	1.9%	**1.0**	0.024	2.2%	0.0		$t_{crit} > t_{calc}$

$\bar{x}$ = mean value of corresponding replicates given in Table 3.5

s = standard deviation of corresponding replicates in Table 3.5.

CV = coefficient of variation of corresponding replicates in Table 3.5 (CV = (s / $\bar{x}$) * 100%)

Normalization in Table 3.6 was done by firstly calculating the average value and standard deviation of native protein concentrations of two replicates for each selected

ten methods. Then, both average value and standard deviation were divided by the average value for each of the methods. The aim of QC-Δ set was to investigate the effect of possible denaturation in short time gap between preparation times of first to sixth samples and short duration differences that may occur while preparing different method sets. So, the condition was exaggerated by the thermal incubation. It was seen that there was no statistically significant difference between native and denaturated protein samples since all t_{calc} values were found to be smaller than t_{crit}.[212] This means that any possible denaturation of proteins during sample preparation were not statistically important and didn't affect the BCA assay results.

3.3.1.2.2 Quantitative Solubility Comparison

Effect of buffer choice on protein solubility was shown in Part 3.3.1.1. qualitatively. Here, concentrations of protein samples that were prepared in buffers of selected methods with equal amounts of proteins were quantitatively analyzed. This analysis was done with the BCA Assay Kit. Further, an equal amount of BSA was added to deionized water as control set. Separate calibration curves were drawn with three replicates of Standard BSA Solution (2.0 mg/ml) provided by the kit in seven concentration points. Another calibration curve was drawn for QC-DI-BCA with deionized water.

Also, LOD and LOQ calculations were done for each curve with their respective slope values and k constants taken as 3.3 for the LOD and 10 for the LOQ with the equations given Figure 3.7.[213]

$$\mathrm{LOD} = 3.3 * \frac{\text{standard devation of ten blank measurments}}{\text{slope of the calibration curve drawn with the same buffer as blank}}$$

$$\mathrm{LOL} = 10 * \frac{\text{standard devation of ten blank measurments}}{\text{slope of the calibration curve drawn with the same buffer as blank}}$$

Figure 3.7. Limit of detection (LOD) and Limit of linearity (LOL) equations

Calibration curves were given in Appendix starting with Figure 4.1. until Figure 4.8. LOD and LOQ values are given in Table 3.7.

Table 3.7. LOD and LOQ values of calibration curves in buffers of selected methods and DI

	Methods								
	DI	A	B, C, J	D	E	F	G	H	I
LOD (mg/ml)	0.024	0.042	0.062	0.024	0.022	0.024	0.024	0.11	0.030
LOL (mg/ml)	0.074	0.13	0.19	0.073	0.068	0.071	0.073	0.32	0.091

Initial native protein concentrations in the samples were then calculated using the respective calibration curves. Mean values, standard deviations, and CV values are also calculated as given in Table 3.8. Additionally, results are shown in graphs with their standard deviations in Figure 3.8.

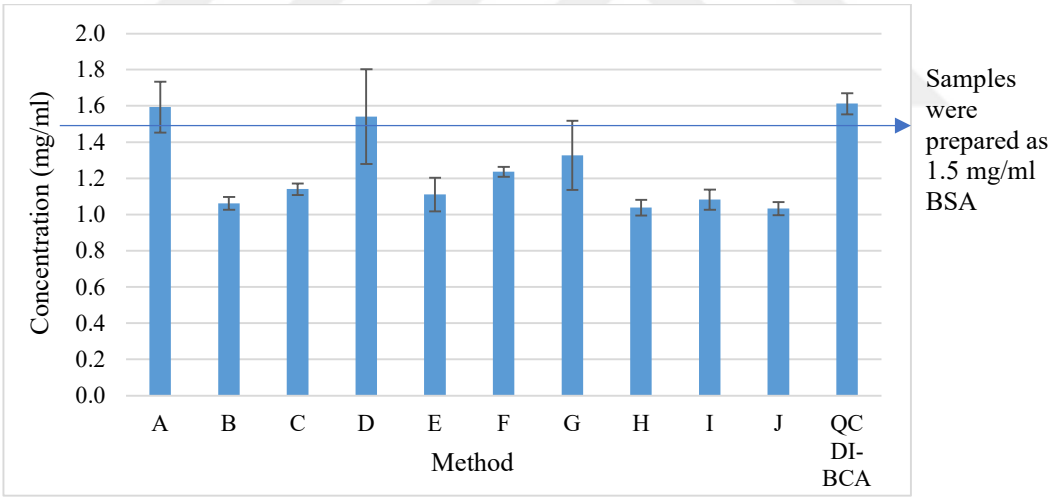


Figure 3.8. Concentrations of BSA Protein obtained from the Native Protein Analysis

Table 3.8. Concentrations of experimental replicates of BSA proteins

Method	Rep 1 (mg/ml)	Rep 2 (mg/ml)	Rep 3 (mg/ml)	Rep 4 (mg/ml)	Rep 5 (mg/ml)	Rep 6 (mg/ml)	Statistics (mg/ml)	
DI	1.58	1.62	1.65	1.58	1.55	1.71	Mean Std Dev CV	1.61 0.06 3.63%
A	1.50	1.50	1.45	1.52	1.81	1.74	Mean Std Dev CV	1.59 0.14 8.82%
B	1.08	1.06	1.02	1.05	1.05	1.12	Mean Std Dev CV	1.06 0.04 3.31%
C	1.16	1.13	1.19	1.14	1.10	1.13	Mean Std Dev CV	1.14 0.03 2.82%
D	1.36	1.32	1.48	2.04	1.53	1.52	Mean Std Dev CV	1.54 0.26 17.0%
E	1.11	1.13	1.03	1.00	1.27	1.13	Mean Std Dev CV	1.11 0.09 8.30%
F	1.25	1.25	1.26	1.18	1.23	1.24	Mean Std Dev CV	1.24 0.03 2.26%
G	1.20	1.28	1.13	1.22	1.55	1.58	Mean Std Dev CV	1.33 0.19 14.4%
H	1.07	1.07	1.03	0.97	1.08	1.02	Mean Std Dev CV	1.04 0.04 4.13%
I	1.04	1.06	1.13	1.09	1.01	1.16	Mean Std Dev CV	1.08 0.06 5.16%
J	1.01	1.03	0.98	1.03	1.06	1.09	Mean Std Dev CV	1.03 0.04 3.50%

It can be observed that similar concentrations were obtained in different methods using the same buffer condition due to solubility being equivalent with each other. Additionally, CV of these methods resulted to be similar. Methods B, C and J has100 mM Tris Buffer and all of them yielded ~1.0 mg/ml BSA concentration with ~3.0%

CV. Methods D and G has lower concentrations of ABC Buffer. These methods are the only two methods that showed high CV values with ~15-16%. Overall, it is also seen that ABC Buffer, Method A, D, E, G, yielded higher CV values than methods using Tris Buffer. In other words, Tris Buffer showed more producible results than ABC Buffer when it comes to protein solubility. However, when it comes to solubility efficiency, ABC Buffers showed greater performance, yielding ~1.5 timer higher native protein concentration. Also, Method A, which has 50 mM ABC Buffer yielded very similar results to the QC-DI-BSA. As a result of native protein concentration assay, it is seen that buffers used effect the solubility of the protein greatly. Spectra analysis results yielded higher solubility for Tris buffer whereas yielded lower solubility for BCA buffer. This contradiction may depend on the Tris acting as an interfering reagent for BCA Assay.[119] This would limit the color formation of the assay yielding lower results.

3.3.2 Denaturation Efficiency Comparison

Circular Dichroism Spectroscopy was utilized for analysis of tertiary structure of proteins, or the degradation of the tertiary structure thereof. Decrease in the peak ellipticity near 220 nm observed due to decrease of α-helix structure of the protein.[61,63] Protein denaturation causes α-helixes to unfold, hence decreasing the CD peak. Varying degrees of peak decrease indicates different degrees of denaturation. Six experimental BSA protein replicates were denaturated and reduced according to the method of choice. For comparison purposes, native BSA in respective buffers were also analyzed. CD spectra were given in Figure 3.9 and Figure 3.10.

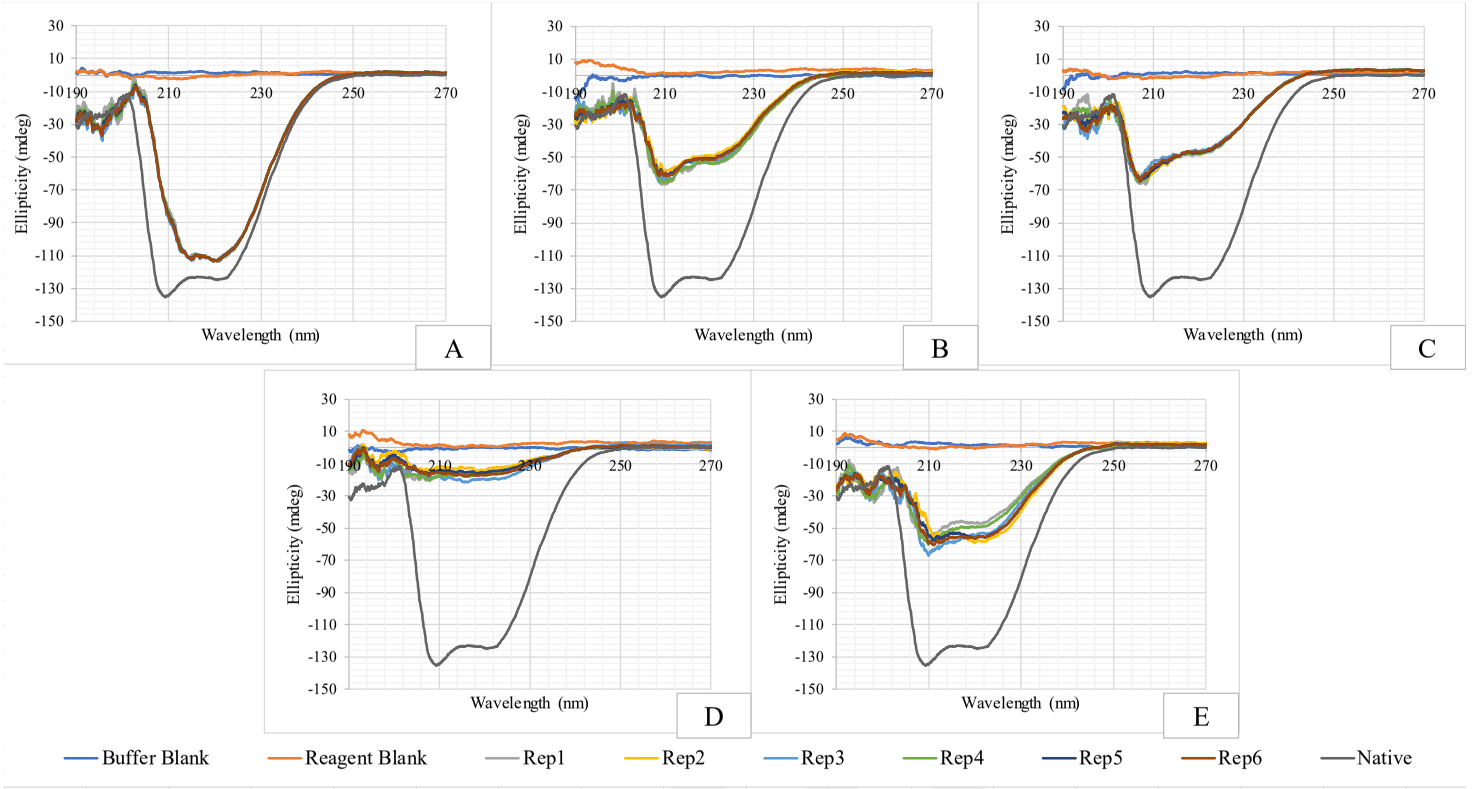


Figure 3.9. Native and denaturated & reduced protein sample peaks in CD spectra according to Methods A to E.
All sample sets include six experimental replicates denaturated according to selected method procedures and native protein dissolved in the respective buffer of the method. Samples were denatured using (A) Method A; 30 min incubation in 50 mM ABC Buffer with 32.5 mM DTT at 60 °C. (B) Method B; 30 min incubation in 100 mM Tris Buffer and 6 M Urea with 20 mM DTT at 37 °C. (C) Method C; 10 min incubation in 100 mM Tris Buffer and 1% SDC with 9.6 mM TCEP at 95°C. (D) Method D; 30 min incubation in 22.7 mM ABC Buffer and 5% SDC with 5 mM TCEP at 60 °C. (E)Method E; 60 min incubation in 100 mM ABC Buffer and 7.3 M Urea with 45.8 mM DTT at 30°C.

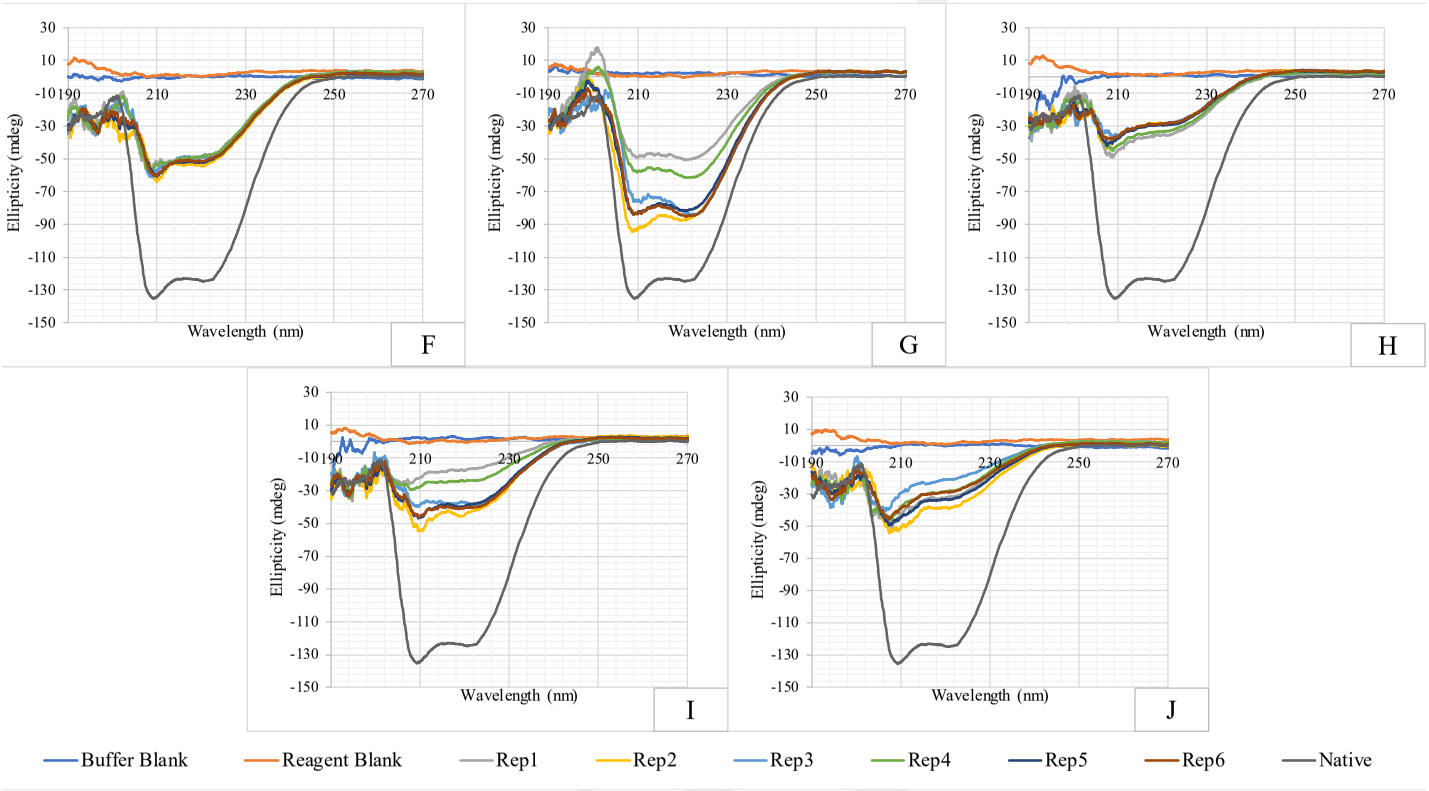


Figure 3.10. Native and denaturated & reduced protein sample peaks in CD spectra according to Methods F to J.
All sample sets include six experimental replicates denaturated according to selected method procedures and native protein dissolved in the respective buffer of the method. Samples were denatured using (F) Method F; 40 min incubation in 10 mM Tris Buffer and 6 M Urea with 30 mM DTT at 40°C. (G) Method G; 30 min incubation in 15 mM ABC Buffer and 5% SDC with 5 mM TCEP at 37 °C. (H) Method H; 30 min incubation in 300 mM Tris Buffer and 9 M Urea with 12.9 mM TCEP at 37 °C. (I) Method I; 30 min incubation in 200 mM TrisHCl Buffer and 6 M Urea with 13 mM DTT at 37 °C. (J)Method J; 30 min incubation in 100 mM Tris Buffer and 8 M Urea with 200 mM TCEP at 60 °C.

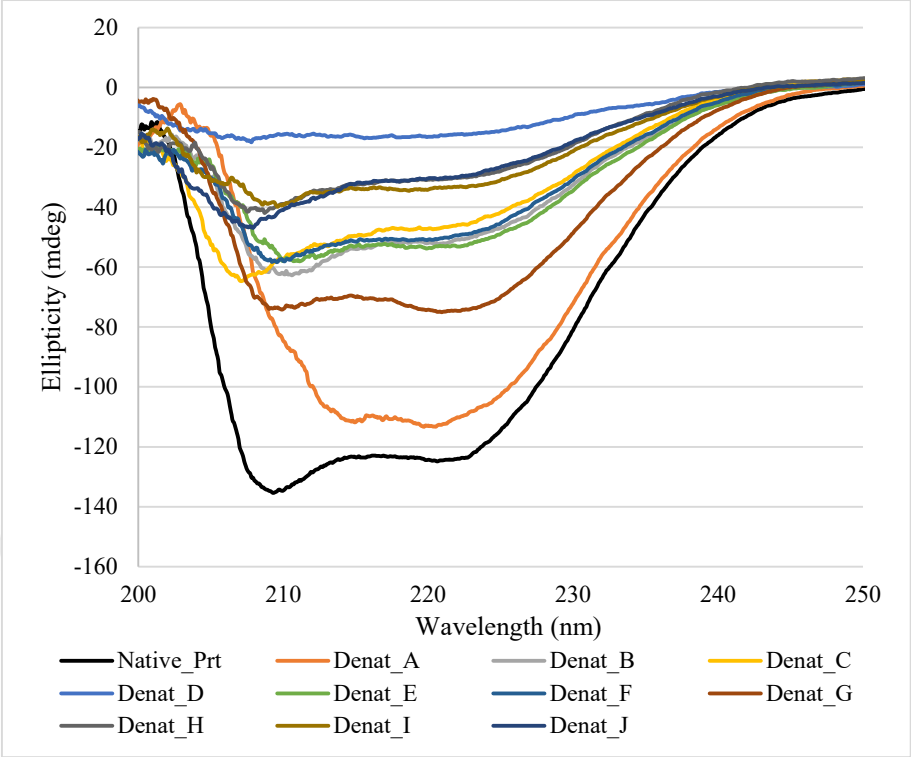


Figure 3.11. Comparison of CD Spectra of Denaturated and Reduced BSA Proteins according to methods and Native BCA Protein

Spectra for denatured BSA according to selected methods were given in one spectrum at Figure 3.11 by taking the average of six replicates within the same method to obtain a representative spectrum. Here, the greatest denaturation was observed for Method D which uses 5% SDC at 60 °C incubation. Followed by three methods, H, I, J.

Common features of these methods were utilization of urea and Tris buffer. H and I method procedures involve the samples incubation at 37 °C while incubation temperature in Method J was 60 °C. Consequently, Tris concentration of methods H and I were 2-3 fold of Method J. Here, equal denaturation performance of these methods may indicate the compensation of lack of higher temperature (60 °C) via 2-3 fold higher concentration of Tris buffer.

The temperature effect was also observed with the results of Method F and B. These methods also used urea at a similar temperature to Methods I and H. However, Tris concentration was lower than I and H. Since denaturation performance of F and B were lower than I and H, it can be concluded that denaturation effect of urea was increased via the presence of harsh buffer environment. In other words, 200 mM to 300 mM Tris concentration in 60 °C incubation temperature yielded better denaturation efficiencies.

Method G employed 5% SDC in the presence of low concentration of ABC, like Method D. However, there was a difference with the result of as can be seen in Figure 3.11. While best denaturation occurred during the Method D procedure, one of the worst denaturation efficiencies was observed for Method G. Main difference is the temperature for these methods, D procedure involves a 30 min incubation at 60 °C while G involves a 30 min incubation at 37 °C. This showed that denaturation does not depend on only harsh chemicals or temperature. Combining them yielded the best results.

3.3.3 Enzyme Activity Analysis

Most proteolytic digestion methods rely on the enzymes for the formation of peptides from proteins.[214] Hence, enzyme activity is a crucial criterion. Parameters like temperature and substrate concentrations affect the activity of the enzyme along with the buffers and pH of the environment.[215] The Trypsin Activity Colorimetric Assay Kit was utilized to compare the activities of the enzymes in samples with the aforementioned preparation methods. The working principle of the kit was based on the measurement of the p-NA molecule, which is a specific product of the hydrolysis of BAPNA with trypsin which is given as trypsin substrate in the kit.[216] So, the calibration curve of p-NA is drawn using the provided p-NA stock solution of the kit. The curve is given in Figure 3.12.

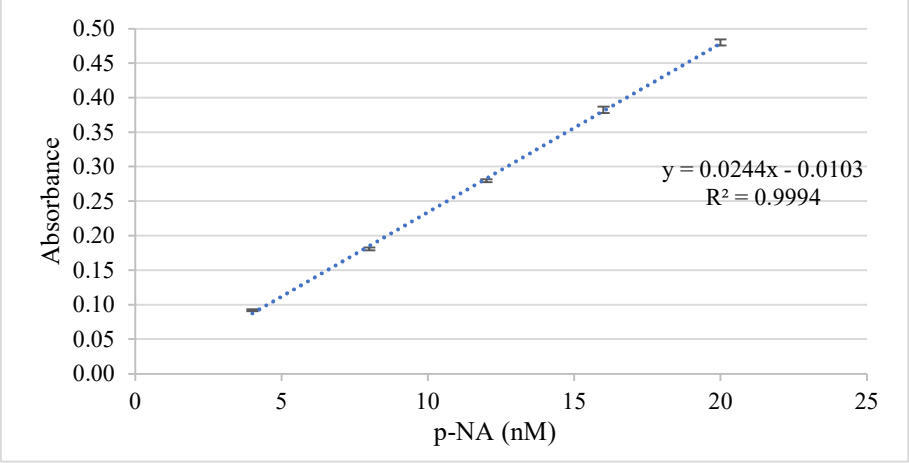


Figure 3.12. Calibration Curve of Enzyme Activity Kit measured at 405 nm

The linearity of the calibration curve was obtained as 0.9994, LOD concentration was calculated as 0.156 nM, and LOQ concentration as 0.474 nM.

During the sample preparation procedure, enzymes show activity in a medium where buffer, denaturation, reduction, and alkylation reagents were added prior, followed by the defined incubations according to the method. To obtain the activity of the

enzyme in their working medium, digestion sets were prepared in accordance with the methods and were used instead of the kit buffer. Additionally, a QC was designed utilizing the provided kit buffer. Figure 3.13 shows the p-NA formation during 0th and 100th minutes.

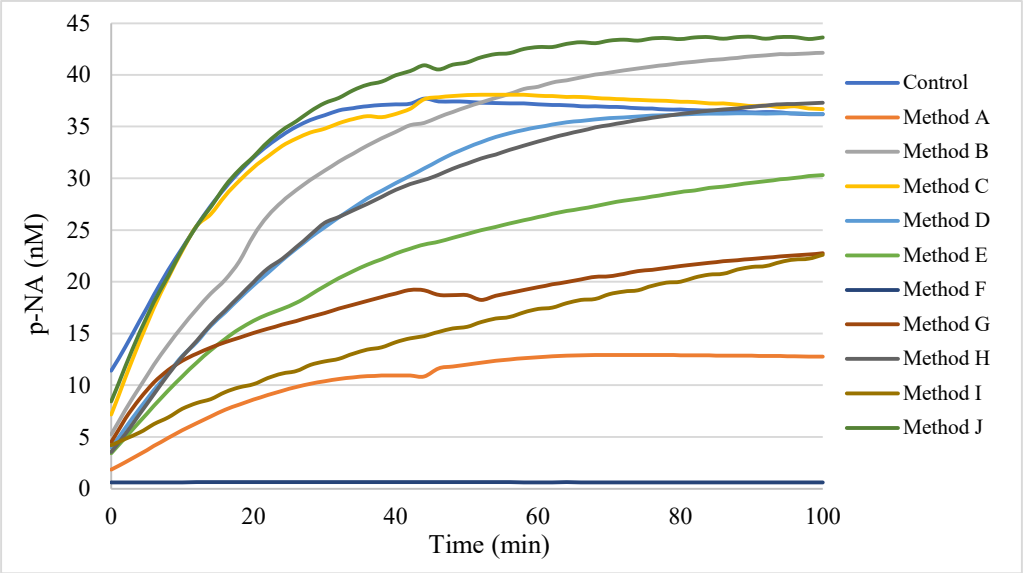


Figure 3.13. Trypsin Activity in samples prepared by Methods A-J measured with two-minute intervals at 405 nm and concentrations were calculated with calibration curve given of p-NA. pH values were set to 8.0 to 8.5 since procedures included this step except F. Method F procedure created a medium with pH 6.0.

The amount of product, p-NA, at 100th minute were given with standard deviation of two replicates at Figure 3.14. Reading was stopped at 100th minute because most of the samples stopped producing p-NA and was flat lines like Method A, C, D, F, J. Trypsin enzyme activity is negatively affected by the higher concentration of denaturation reagents.[217] Nevertheless, the concentration of these reagents were taken below the limitation amounts within each method.

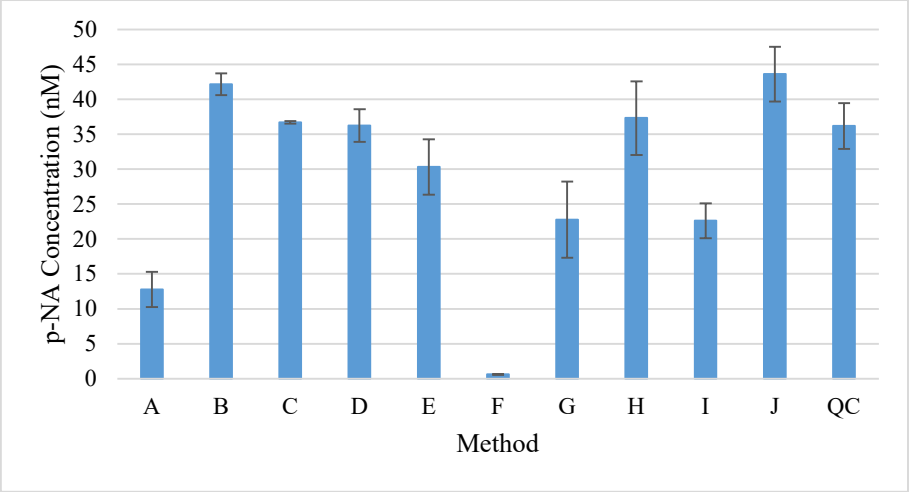


Figure 3.14. Final p-NA Concentration

Another limitation is the pH range for the trypsin enzyme, which is between pH 7 to 9. All methods utilize a buffer between pH 8.00 and 8.50 since procedures of digestion methods involved the pH arrangement to 8.00 to 8.50. However, dilution of Method F with distilled water without pH adjustment, as mentioned in methods and materials part, lowers the pH of the environment to 6.00. The inactivity of the trypsin enzyme in this method was observed in Figures 3.13 and 3.14. It should be noted that the procedure in the study of Method F calls for distilled water specifically without any disclosure of the freshness or the pH of the water. Moving further, it is seen that trypsin tends to show higher activities in Tris Buffers (Methods B, C, H, J) compared to ABC Buffer (Methods A, G). However, this is not the definite case due to the product ratio of Method D, which utilizes ABC Buffer, and the lower results of Method I. Lastly, the incubation time for digestion might be modified. As seen from Figure 3.13, the highest product formation was seen within the first 30 minutes. The product formation stops for Methods A, B, D, and J since increase in the p-NA production stops and samples start to give flat lines in Figure 3.13. For the rest of the methods, the product formation ratio was significantly lower compared to the initial 30 minutes. This may lead to the questioning of the common proteolytic digestion

incubation duration in the literature which is 16-18 hours as seen in Table 3.1. Similar conclusion was made by the study group of Method C.[152] This method gave the lowest CV value with 0.55% compared to average CV value of 11% and highest one being 23%. Overall, it was observed that enzyme activity under different conditions of the varying methods was not affected by one parameter, rather it was a combination of reagents, pH, and digestion duration. Digestion temperature was fixed to 37 °C in all the methods investigated, so there were no varying effects due to temperature.

3.3.4 Total Peptide Analysis

A commercial Peptide Assay Kit was employed for the quantitative total peptide analysis. Before analyzing the samples, two quality control sets were designed to observe the possible interference of reagents used in selected methods and the specificity of the assay towards peptides were analyzed.

3.3.4.1 QC Sets for Peptide Assay

3.3.4.1.1 QC-PA Set

QC samples were prepared for Peptide Assay (QC-PA) similarly to QC-BCA samples. Absorbance signals of buffers, reagent blanks, and reagent & enzyme solutions of ten methods analyzed with peptide assay at λ = 480 nm were given in Table 3.9. To obtain Peptide assay signals, samples were mixed with Peptide assay kit reagents according to Peptide assay kit procedure and incubated at 37 °C for 30 minutes. Then absorbance of the samples was measured at λ = 480 nm as kit instructs. All of these buffer samples were prepared with Deionized Water (DI). In order to control the interferences of chemicals, each set was blank corrected according to DI.

Table 3.9. Quality Control Set for Investigation of Peptide Assay (QC-PA)
Buffer, Reagent Blank and Reagent & Enzyme samples were prepared according to the selected ten methods. Absorption reading was done at 480 nm. (DI= Deionized Water)

	Absorbance Signals of QCs for Peptide Assay Kit (QC-PA)					
	Buffer (Blank)		**Reagents Blank**		**Reagents & Enzyme**	
Method	**Replicate 1**	**Replicate 2**	**Replicate 1**	**Replicate 2**	**Replicate 1**	**Replicate 2**
DI	0.00	0.00	-	-	-	-
A	0.13	0.12	**> Max**	**> Max**	**> Max**	**> Max**
B	-0.03	-0.04	0.01	0.03	0.00	0.02
C	-0.02	-0.03	-0.03	-0.03	-0.04	-0.04
D	-0.04	-0.05	-0.03	-0.03	-0.04	-0.03
E	-0.06	-0.07	**> Max**	**> Max**	**> Max**	**> Max**
F	-0.04	-0.04	**> Max**	**> Max**	**> Max**	**> Max**
G	-0.06	-0.08	-0.04	-0.03	-0.04	-0.04
H	-0.06	-0.06	-0.01	-0.01	-0.02	-0.01
I	-0.02	-0.03	0.13	0.03	0.01	0.01
J	0.01	0.00	0.02	0.03	0.01	0.01

Since first steps of both BCA assay and Peptide assay are based on the Biuret reaction, it was expected to observe interferences at the same methods. However, this carries importance since Peptide Assay was performed on the final step of digestion. The digestion methods were unable to be altered to remove interferences because the main objective of this study is to compare commonly used digestion methods in proteomics research. Hence, altering the procedure would change the efficiency of the method, resulting in the deviation from the literature. The Peptide assay step served as a control to monitor the impact of selected methods on protein digestion. The QC-PA set shows that peptide assay results from Methods A, E, and F should not be discussed as they do not accurately reflect the peptide concentration in the sample due to reagent interferences.

3.3.4.1.2 QC-Reverse Set

QC-Reverse set was designed to investigate the specificity of the Peptide Assay towards peptides rather than proteins. Resulting medium after digestion contained peptides. Results of this QC set determined the possibility of measuring undigested proteins after the digestion. So, Peptide Assay was applied to protein samples that didn't include peptides. Concentrations of native proteins analyzed with BCA assay are shown in Table 3.10. Along with responses of Peptide assay for samples which didn't contain any peptides. Samples only contained native BSA proteins.

Table 3.10. Quality Control Set for Investigating Peptide Assay Results in a Complete Protein Environment (QC-Reverse).
Native BSA protein samples were prepared with 1.5 mg/ml concentration. Samples were analyzed with both BCA assay (λ = 562 nm) and Peptide assay (λ = 480 nm).

	Concentrations of Protein QCs for Peptide Kit (QC-Reverse)			
	BCA Assay		**Peptide Assay**	
	Native Protein Concentration (mg/ml)		**Peptide Assay Results for Protein Samples (mg/ml)**	
Method	**Replicate 1**	**Replicate 2**	**Replicate 1**	**Replicate 2**
DI	1.2	1.3	0.83	0.81
A	1.2	1.4	1.0	1.1
B	1.2	1.3	0.92	0.94
C	1.2	1.2	0.91	0.93
D	1.3	1.4	1.0	1.1
E	1.3	1.2	0.89	1.0
F	1.3	1.5	1.1	1.1
G	0.80	0.87	0.63	0.73
H	0.81	0.85	0.70	0.71
I	1.2	1.1	0.75	0.72
J	1.3	1.3	1.0	0.96

To observe results more clearly, normalization was done on the results same as the third QC set. Results were given in Table 3.11.

Table 3.11. Normalization values of QC-Reverse Samples.
Normalization was done by firstly calculating the average value and standard deviation of native protein concentrations of two replicates for each method. Then, both average value and standard deviation were divided by the average value for each of the methods. Then, CV values were calculated by diving the standard deviation to average value of their respective methods.

	Normalized Concentrations of Protein QCs for Peptide Kit (QC-Reverse)					
	BCA Assay			**Peptide Assay**		
	Native Protein Concentration (mg/ml)			**Peptide Assay Results after Thermal Incubation (mg/ml)**		
Method	**$\bar{x}$**	**s**	CV	**$\bar{x}$**	**s**	CV
DI	**1.0**	0.071	7.1%	**0.66**	0.0086	1.3%
A	**1.0**	0.063	6.3%	**0.81**	0.029	3.6%
B	**1.0**	0.052	5.2%	**0.75**	0.013	1.7%
C	**1.0**	0.005	0.47%	**0.74**	0.0096	1.3%
D	**1.0**	0.066	6.6%	**0.76**	0.012	1.5%
E	**1.0**	0.075	7.5%	**0.76**	0.080	10%
F	**1.0**	0.038	3.8%	**0.78**	0.0016	0.21%
G	**1.0**	0.073	7.3%	**0.83**	0.064	7.7%
H	**1.0**	0.033	3.3%	**0.85**	0.012	1.4%
I	**1.0**	0.053	5.3%	**0.64**	0.015	2.4%
J	**1.0**	0.019	1.9%	**0.78**	0.037	4.8%

Results showed that Peptide assay yielded signals for samples that contained no peptides. These samples contained only proteins. Hence, this QC set indicated that possible undigested proteins in a medium contains peptides cannot be measured accurately using Peptide assay.

3.3.4.2 Quantitative Total Peptide Concentration Comparison

After the digestion of BSA samples, total peptide concentration analysis was performed for every six replicates of each method. To have results in equivalent conditions, all solvent in the medium was evaporated and samples were reconstituted in 300 µL of 0.1% Formic Acid solution. Formic acid is widely used as the solvent since it is able to solubilize proteins and peptides by agitating intramolecular hydrogen bonds, promoting repulsions through protonation, and dissolving polar residues. The acid is also used as a component of mobile phase during liquid chromatography in LC-MS analysis. Calibration curve for the Peptide Assay was drawn using 0.1% FA as shown in Figure 3.15.

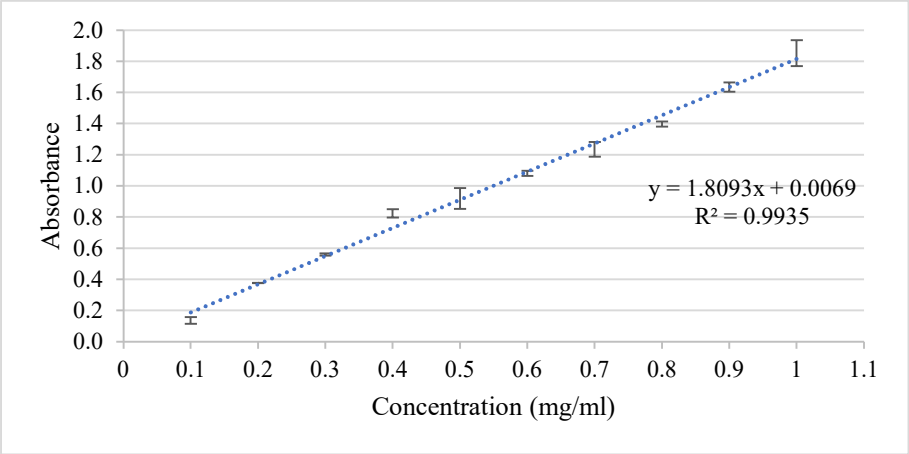


Figure 3.15. Peptide Assay Calibration Curve measured at 480 nm

The linearity of the calibration curve was obtained as 0.9935, LOD concentration was calculated as 0.0123 mg/ml, and LOQ concentration as 0.0374 mg/ml in 0.1%Formic Acid Solution. The concentration of obtained peptides after digestion process is given at Table 3.12 and Figure 3.16.

Table 3.12. Yielded Peptide Concentrations of Different Digestion Methods Obtained from Equivalent Proteins

Method	Rep 1 (mg/ml)	Rep 2 (mg/ml)	Rep 3 (mg/ml)	Rep 4 (mg/ml)	Rep 5 (mg/ml)	Rep 6 (mg/ml)	Statistics (mg/ml)	
A	0.68	0.90	0.87	0.87	0.94	0.96	Mean BSA Std Dev CV	0.87 0.10 11%
B	0.69	0.69	0.58	0.80	0.81	0.62	Mean BSA Std Dev CV	0.70 0.09 13%
C	0.42	0.41	0.34	0.29	0.47	0.34	Mean BSA Std Dev CV	0.38 0.07 17%
D	0.46	0.44	0.27	0.34	0.41	0.44	Mean BSA Std Dev CV	0.40 0.07 18%
E	0.73	0.86	1.01	1.01	0.92	1.06	Mean BSA Std Dev CV	0.93 0.12 13%
F	1.43	1.61	1.58	1.58	1.40	1.34	Mean BSA Std Dev CV	1.49 0.11 8%
G	0.56	0.51	0.72	0.70	0.86	0.69	Mean BSA Std Dev CV	0.67 0.12 19%
H	0.26	0.25	0.24	0.22	0.26	0.22	Mean BSA Std Dev CV	0.24 0.02 8%
I	0.45	0.48	0.71	0.47	0.54	0.41	Mean BSA Std Dev CV	0.51 0.11 21%
J	0.40	0.40	0.39	0.36	0.37	0.41	Mean BSA Std Dev CV	0.39 0.02 5%

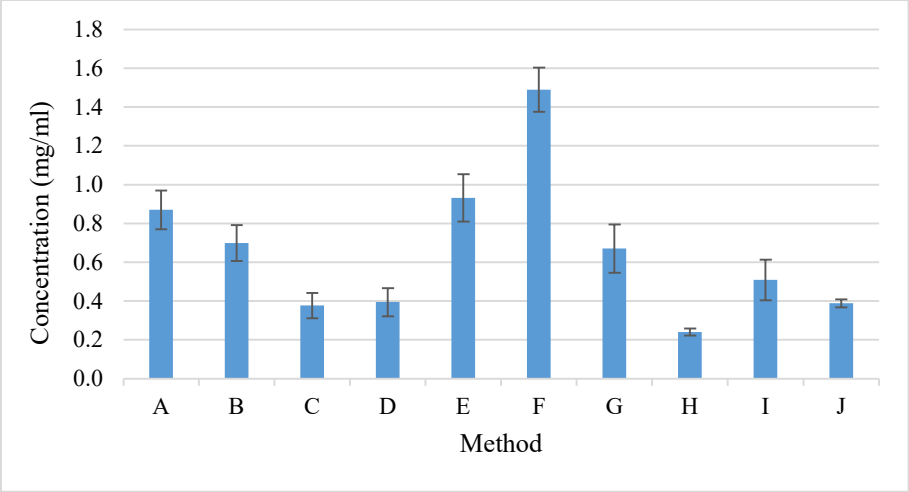


Figure 3.16. Yielded Peptide Concentrations as the result of Different Digestion Methods Obtained from Equivalent Proteins

Figure 3.17 shows obtained peptide product amount from the different digestion methods. It is clearly seen that even though starting samples were equivalent, peptide formation is greatly affected by the digestion method. This is another indication that digestion methods cannot be utilized interchangeably.

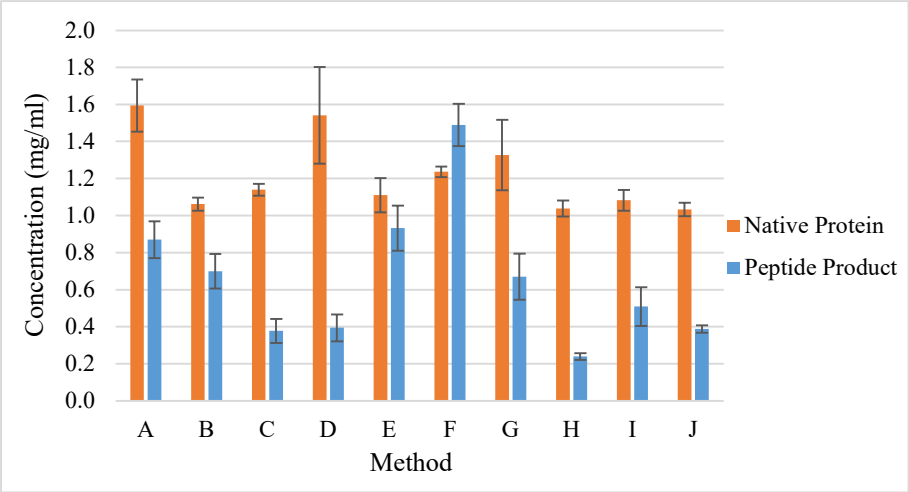


Figure 3.17. Varying Native Protein Concentrations due to Solubility Difference in different buffers associated with different Methods and Yielded Peptide Concentrations

Nevertheless, it was observed during the native protein analysis that the solubilities of proteins were also showing variance with respect toto the type of the buffer. So, different amounts of proteins were available to digest for different methods as can be seen in Figure 3.18.

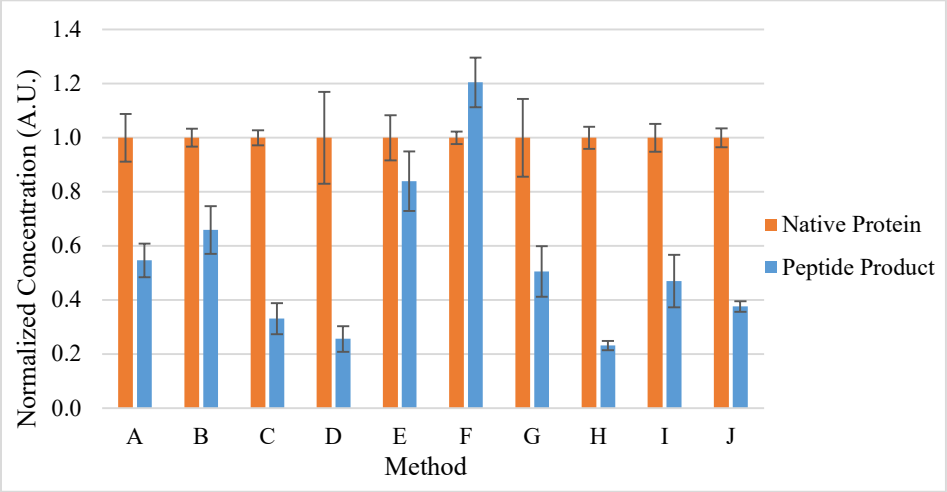


Figure 3.18. Normalized Native Protein and Yielded Peptide Concentrations

These differences are caused by the variances between the methods used. However, in order to further investigate the digestion efficiencies based on the method selected, each peptide amount was normalized according to the initial native proteins. Values are given in Figure 3.18.

Method F yielded higher amounts of the peptide than the added protein. The reason is the reagent interferences as given in the QC-PA. It was also given that reagent interferences were observed in Method A and E, so they were considered inconclusive and left out of the discussion.

The differences in the results were not dependent on one variable. Rather, cumulation of variances affected the peptide formation. However, overall, it was observed that methods utilizing DTT yielded higher amounts of peptide than methods with TCEP. Another important result was that the amount of enzyme used didn't show direct

correlation with the peptide formation. Method H, which utilized the highest amount of enzyme per protein yielded the lowest amount of peptide. Meantime, Method B, which used ~3 times less enzyme resulted in a ~3 fold increase in peptide amount. In enzyme activity analysis, Method B showed slightly higher activity than Method H. While this is one of the factors on better performance of Method B, it is not enough to explain ~3-fold difference. This shows the importance of the entirety of the digestion process. Surprisingly, Method B and H are two of the most similar processes, from a reagent standpoint. They only differ in terms of the reducing reagent which DTT and TCEP, respectively. However, concentrations of the reagents were varying. This shows that not only the type of reagents, but their concentrations were also greatly affecting the efficiency of the digestion process.

3.4 Investigation of Digestion Products

In this part, A2MG protein was digested with selected ten digestion methods. Three replicates were digested for each method and replicates were pooled when digestion process was completed. Pooled A2MG products were analyzed with untargeted tandem MS. LC/MS chromatograms of digested A2MG samples were given in Appendix B. Digestion results were investigated for protein group results and peptide group results.

3.4.1 Protein Groups

Protein group results were obtained via data analysis with default parameters in MaxQuant software. Table 3.13 showed that analyzed protein were identified as A2MG for all samples. However, other three parameters given in Table 3.13 showed variations for equivalent protein samples that were digested with different digestion methods. Number of peptides showed the amount of different peptide sequences and derivatives (caused by artificial modifications) software detected. Here, default parameters of MaxQuant software analyses peptide sequences with 8 to 25 residues. Number of theoretical tryptic peptide sequences with 8-25 residues was sixty-five. Therefore, higher number of peptides than sixty-five indicates the presence of miscleaved and artificially modified peptide sequence for samples digested with selected methods except Method F. This method was the only method yielded lower number of peptides than theoretical peptide count for A2MG. Lack of tryptic peptides may be due to inactive trypsin enzyme in the medium. Consequently, enzyme activity test in part 3.3.3. showed that trypsin enzyme in medium of Method F shows limited activity.

Table 3.13. Protein groups results of A2MG samples digested with Methods A to J. "Peptides" column refers to number of identified peptide sequences that belong to A2MG protein within each sample. "Sequence coverage %" column refers to completion percent of A2MG protein amino acid sequence with the peptides found in the samples. MS/MS count refers to number of tandem spectra MS instrument was able to analyze.

Method	Protein names	# of Peptides	Sequence coverage [%]	MS/MS count
A	Alpha-2-Macroglobulin	102	70.9	1278
B		100	70.6	946
C		126	74.2	1443
D		92	68.7	2386
E		134	74.2	1942
F		23	28	172
G		94	56	428
H		110	72.8	3056
I		113	69.3	1960
J		90	65.7	1843

Next, "Sequence coverage %" column yields the completion of residue sequence of A2MG protein obtained with peptides obtained. Here, number of peptides didn't directly correlate to sequence coverage. An example can be given as follows; two consequent peptide sequences may be AAAAAK and GGGGGR. These correctly cleaved peptides were counted as two by the software. However, if a miscleavage occurred, AAAAAKGGGGGR peptide would also be present in the sample and be counted as third peptide. Yet, the sequence coverage wouldn't have changed by the presence of this third peptide. In Table 3.13, results of Method G indicated that number of peptide found was 94, which was higher than three methods. However, it has the lowest sequence coverage except for the results of F. It's possible that number of peptides were due to derivatives of a limited number of peptides while some tryptic peptides weren't formed here. This may be explained with denaturation efficiency test done in part 3.3.2 and enzyme activity test done in part 3.3.3. Method G was one of the lower performing methods in denaturation efficiency and. Hence,

during digestion, trypsin might not be able digest peptide sequences that were not denatured yielding a lower sequence coverage.

Last column of Table 3.13. shows the MS/MS count obtained with each method. MS/MS spectra were used in matching the fragmentations of peptides and determining the peptide sequence its originated from.[219] It was used to identify and locate the modifications within a peptide. Hence, higher MS/MS counts yielded detailed information about peptides obtained with digestion methods. More detailed investigation of MS/MS count was done in further part of 3.4.2.2.

3.4.2 Peptide Groups

Peptide group results were analyzed under three categories which were unique peptide distribution, cleavage efficiencies and artificial modifications. Here, MaxQuant software was not able to successfully complete the data analysis when more than two artificial modification were scanned. Instead, Proteome Discoverer (PD) software was employed for data analysis with five modifications which was the maximum number of modification software allowed to scan in one analysis. Two missed cleavage sites were allowed in the data analysis which is the default setting in PD as well as MaxQuant. Unique peptides of A2MG were obtained from UniProt BLAST database and given in Table 3.14.[132] Data analysis results were filtered for unique peptides since bottom-up proteomics employs unique peptides to identify and quantify proteins.[220]

Table 3.14. Unique Peptide List of A2MG Protein obtained from UniProt BLAST database.

Protein Name	UniProt ID	Peptide Sequence
A2MG	P01023	AAQVTIQSSGTFSSK
		DMYSFLEDMGLK
		DTVIKPLLVEPEGLEK
		FEVQVTVPK
		HNVYINGITYTPVSSTNEK
		IAQWQSFQLEGGLK
		LHTEAQIQEEGTVVELTGR
		LLIYAVLPTGDVIGDSAK
		LVHVEEPHTETVR
		QFSFPLSSEPFQGSYK
		QGIPFFGQVR
		TEVSSNHVLIYLDK
		VGFYESDVMGR
		VSNQTLSLFFTVLQDVPVR
		VSVQLEASPAFLAVPVEK

It was found that fifteen peptide sequences given in Table 3.14 were the unique peptide sequences of the A2MG protein.

3.4.2.1 Unique Peptide Distribution and Cleavage Efficiencies

Trypsin enzyme should fragment the protein at carboxy-terminal of Arginine and Lysine residues.[115] However, trypsin enzyme may not cleave every K or R site. Peptide fragments that contain K or R residues besides the C-terminal are called “miscleaved peptides". Here, correct cleavage performance of the enzyme within the selected methods was investigated. Figure 3.19 shows the unique peptide ratios and contributions of miscleaved peptides obtained with selected digested methods.

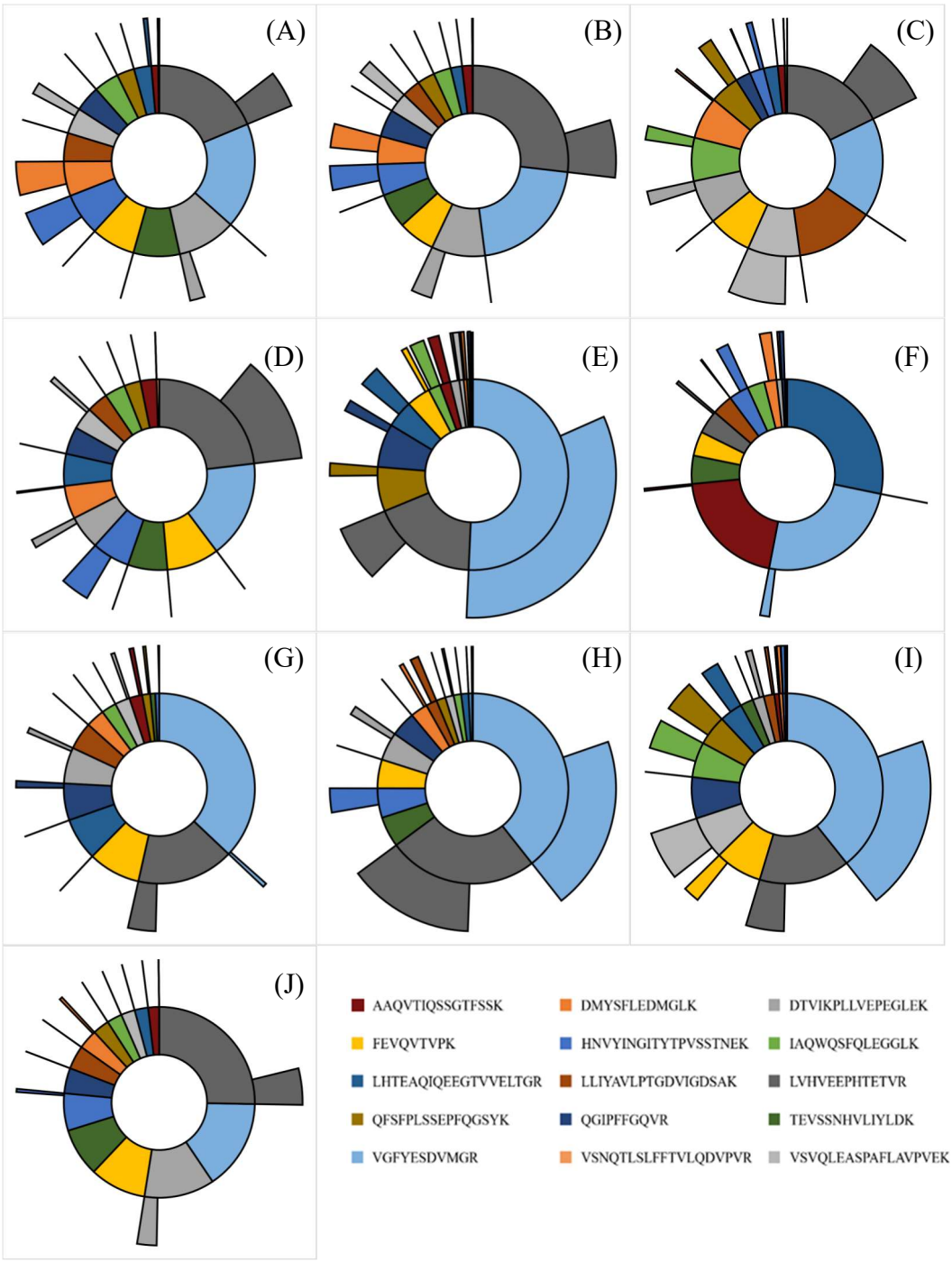


Figure 3.19. Unique Peptide Ratios Every Method Yield and Their Miscleaved Forms (Given in The Outer Ring of The Diagram)

Figure 3.19. showed that Method F didn't yield two of the unique peptides which were QGIPFFGQVR and VSNQTLSLFFTVLQDVPVR. Due to low peptide yield, sequence coverage % and MS/MS count given in Table 3.13, Method F wasn't involved in further discussions. Rest of the methods yielded all fifteen unique peptides; however, distribution of these peptides was not equivalent. To further evaluate the distribution, number of studies on human plasma and serum employed unique peptides of A2MG were searched in PeptideAtlas database.[221] Wordcloud given in Figure 3.20. represents frequency of use of unique peptides.



Figure 3.20. Frequency of use of unique peptides in studies on human plasma/ serum Text size represents the frequency of use in the literature.

Five frequently used unique peptides were chosen as VGFYESDVMGR, DMYSFLEDMGLK, LLIYAVLPTGDVIGDSAK, LVHVEEPHTETVR, and QFSFPLSSEPFQGSYK.

VGFYESDVMGR was observed to contribute to 15% to 50% to total of unique peptides, lowest yield in Method J and highest yield in Method E. Contribution of miscleaved peptides to this sequence ranged from 0% to 64%. 0% miscleavege was observed in five methods (A,B,C,D,J) while three methods (E,H,I) yielded more than 50% miscleavege for VGFYESDVMGR sequence. Number of miscleavege scanned

in data analysis software were an optional entry and it was possible to enter zero miscleavege or any higher number. Here, an example can be given for two biomarker studies employing Method E in equivalent conditions but number of miscleavage entry (one with zero miscleavege allowance and other with two). The difference for equivalent samples would vary ~3-fold for these studies.

DMYSFLEDMGLK peptide sequence was observed to contribute 0.5% to 5.9%. Here, the difference between contribution of VGFYESDVMGR, one of the highest, and DMYSFLEDMGLK, the lowest ratio, was observed. These two peptides were two of the most used sequences chosen to represent A2MG protein in literature. Yet, the difference in percentage yield for DMYSFLEDMGLK varied ~24-fold when different digestion procedures were used while VGFYESDVMGR varied 3-fold as given above. Additionally, contribution of miscleaved peptides for DMYSFLEDMGLK ranged from 3.0% to 65%.

Lowest LLIYAVLPTGDVIGDSAK peptide sequence contribution (0.2%) was observed within Method E, unlike VGFYESDVMGR which showed its highest with this method. Whereas highest contribution (13%) of LLIYAVLPTGDVIGDSAK was seen in Method C. This created a 60-fold difference within contribution ratios. Additionally, lowest miscleaved ratio of 0.3% was also seen with this method.

LVHVEEPHTETVR was the second peptide besides VGFYESDVMGR that was observed in all methods (except Method F) in Figure 3.19. Its contribution varied from 15% to 27%. Its miscleaved ratio varied from 16% to 56%. This unique peptide was observed to be most consistent sequence to contribute to total unique peptides with less ~2-fold change within methods. However, due to miscleaved peptides ratios, this peptide may show variances in studies with different miscleavage allowances.

QFSFPLSSEPFQGSYK was one of the most used unique peptide sequences as given in Figure 3.20. However, it is one of the lowest contributing sequences with 1.4% in

Method G to 7.5% in Method E. In method where it showed the least contribution, it showed the highest miscleavege ratio with 20%. Lowest miscleavage of QFSFPLSSEPFQGSYK was in Methods D and J with 0.2% which created a 100-fold variation range in terms of miscleaved peptide contribution.

It was observed that most frequently used unique peptides representing A2MG weren't in agreement in terms of contribution ratio to total amount of unique peptides. Two of these peptides, VGFYESDVMGR and LVHVEEPHTETVR, showed higher ratio between 15% to 50%. While other three unique peptides showed contribution as small as 0.2%. Variation range of unique peptides reached 100-fold difference between lowest and highest contribution ratios. Miscleaved peptide contribution also showed variation for unique peptides obtained with different digestion methods. Unique peptides such as VGFYESDVMGR showed little to no miscleavege in seven of the digestion methods while it showed up to 64% in remaining seven methods. Another unique peptide, LVHVEEPHTETVR, yielded higher contributing miscleaved peptides forms in all of the methods, making it a variation source upon usage of different allowed miscleavege number during data analysis.

Method D showed best cleavage efficiency for QFSFPLSSEPFQGSYK while it showed the least for LVHVEEPHTETVR. Method E yielded lowest contribution ratios for DMYSFLEDMGLK and LLIYAVLPTGDVIGDSAK while it showed highest contribution to VGFYESDVMGR and QFSFPLSSEPFQGSYK. However, highest miscleaved peptide form of VGFYESDVMGR was also seen with Method E. Method J showed to best cleavage efficiency of four peptides except DMYSFLEDMGLK. Overall, it was observed that methods utilizing lower amounts of enzyme may yield better cleavage performance such as Methods C and J (protein to enzyme ratio of 100:1) compared to Method H (protein to enzyme ratio of 15:1). In conclusion, it was observed that a specific digestion method didn't show greater performance among selected ten methods. Methods may yield better peptide ratios

and cleavage performance for a unique peptide sequence while it yields less than ideal results for another unique peptide sequence.

3.4.2.2 Artificial Modifications

Artificial modifications are uncontrolled modifications that may occur on residues during digestion process due to reagents combined with incubation temperature and durations.[7] Table 1.1 listed a portion of these modifications and residues they are seen on. However, default setting of data analysis software, MaxQuant and PD, includes only oxidation on Met, acetylation on N-term, and carbamidomethylation on Cys (also known as alkylation after the reduction of cysteine residues).

Software identifies the peptides by matching MS and MS/MS spectra of raw data obtained after LC/MS analysis to a library of spectra. More specifically, software matches the total mass of precursor ion to theoretical mass within the mass ppm range that was entered to software. To further verify the results, software also matches the fragmented precursor peptide sequences to theoretical masses of these fragments. Most common fragmentation occur on peptide bonds between carbonyl carbon and nitrogen of amino group on the backbone of the peptide. Fragment that includes C-terminal are referred to as “y fragments” whereas fragments that include N-terminal are referred to as “b fragments” in Roepstorff–Fohlman scheme.[222] Number of residues each fragment include is written as a subscript on the right side of the letter.

Artificial modifications cause mass shifts in mass of precursor and the mass of fragments. In a case of not entering the artificial modifications to software, it cannot match spectra of modified peptides to sequence. This may affect the data output qualitatively and quantitatively. Effects of this phenomenon may be observed in Table 3.13. Results given in this table were obtained via data analysis with default settings. As a results, it was observed that while MS/MS count of Method A was

higher than Method B by 332 spectra, only 2 additional peptides were identified. Another example can be given for Methods H and I. Here, Method H yielded 3056 MS/MS spectra and 110 peptides were able to be identified. Whereas Method I yielded 1960 MS/MS spectra, and more peptides (113) were able to be identified. These additional MS/MS spectra may represent artificially modified peptide sequences that were not able to be identified due to not entering the corresponding artificial modification to the software.

In this part of the study, five artificial modifications were investigated which were carbamidometylation of Lys and N-term; dethiometylation of Met; oxidation of Met, His, Phe, and Trp; carbamylation of Lys, Arg, and N-term; deamidation of Asn and Gln. To exemplify the mass shift with these artificial modifications, chromatogram, MS spectrum and MS/MS spectrum of unmodified and modified versions of DMYSFLEDMGLK unique peptide sequence are given Figure 3.21 to Figure 3.30.

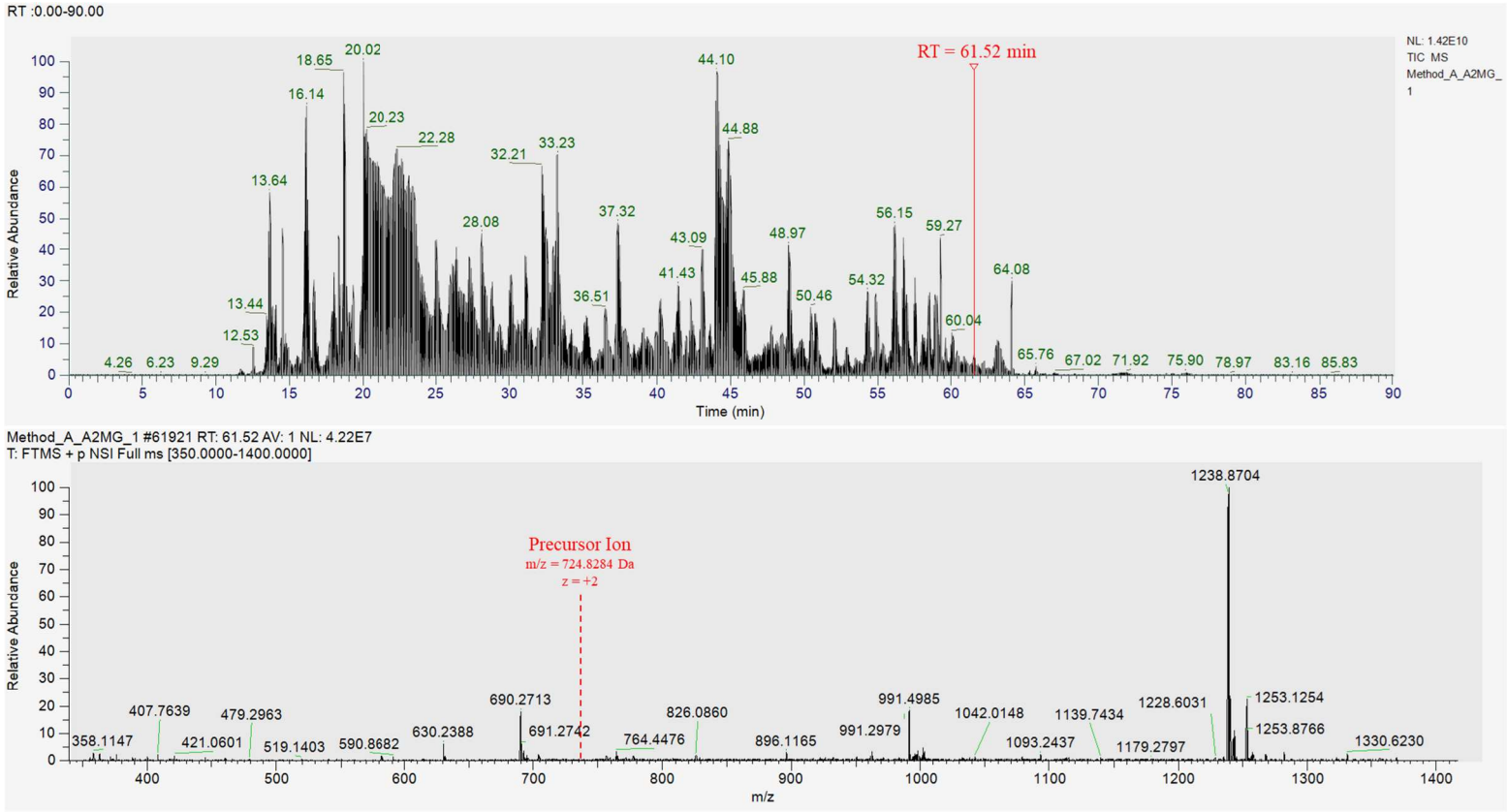


Figure 3.21. Chromatogram and mass spectrum for unmodified DMYSFLEDMGLK peptide

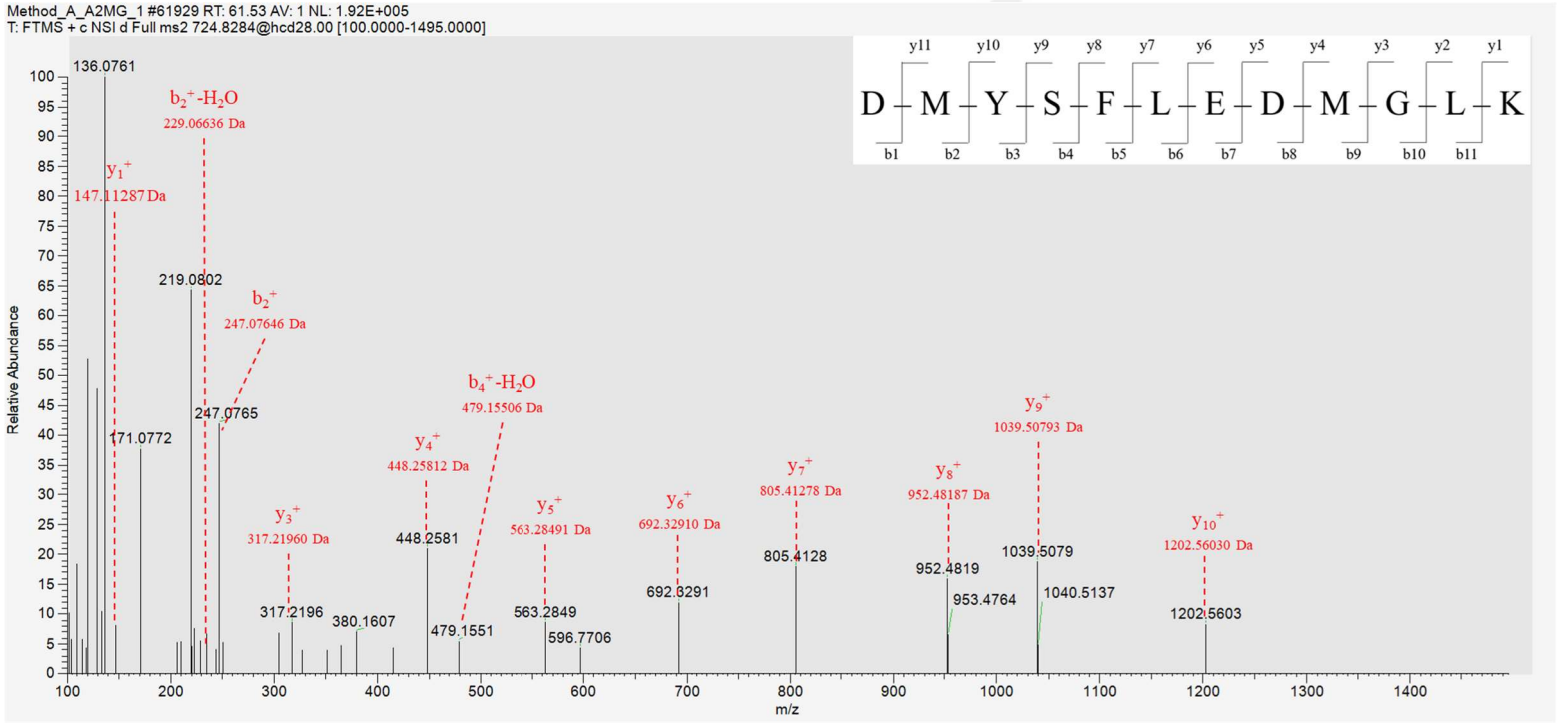


Figure 3.22. MS/MS spectrum for unmodified DMYSFLEDMGLK peptide

Table 3.15. Monoisotopic mass values in Da for unmodified DMYSFLEDMGLK peptide sequence

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	116.03422	58.52075	D			12
2	247.0747	124.04099	M	1333.61667	667.31197	11
3	410.13803	205.57265	Y	1202.57619	601.79173	10
4	497.17006	249.08867	S	1039.51286	520.26007	9
5	644.23847	322.62288	F	952.48083	476.74405	8
6	757.32254	379.16491	L	805.41242	403.20985	7
7	886.36513	443.6862	E	692.32835	346.66781	6
8	1001.39207	501.19968	D	563.28576	282.14652	5
9	1132.43256	566.71992	M	448.25882	224.63305	4
10	1189.45402	595.23065	G	317.21833	159.1128	3
11	1302.53809	651.77268	L	260.19687	130.60207	2
12			K	147.1128	74.06004	1

*Mass values given in red represents the matched values with MS/MS spectra given in Figure 3.22. within 10 ppm range

Table 3.15 showed the monoisotopic mass values of b and y fragments of DMYSFLEDMGLK unique peptide sequence. Red values represented the matched masses in respective unmodified DMYSFLEDMGLK sequence in Figure 3.22. Two different charges were observed for fragments which were +1 and +2. Peak masses were calculated by adding the mass of charge times monoisotopic mass of hydrogen (1.00783 Da) and dividing this sum to charge. Hence, b^{+2} and y^{+2} fragments yield half of the mass of singly charged fragment plus half of the monoisotopic mass of hydrogen. Additionally, Figure 3.21 showed that retention time of unmodified sequence was 61.52 min over the elution time of 90 mins. m/z value of precursor ion was 724.8284 Da with charge (z) of +2.

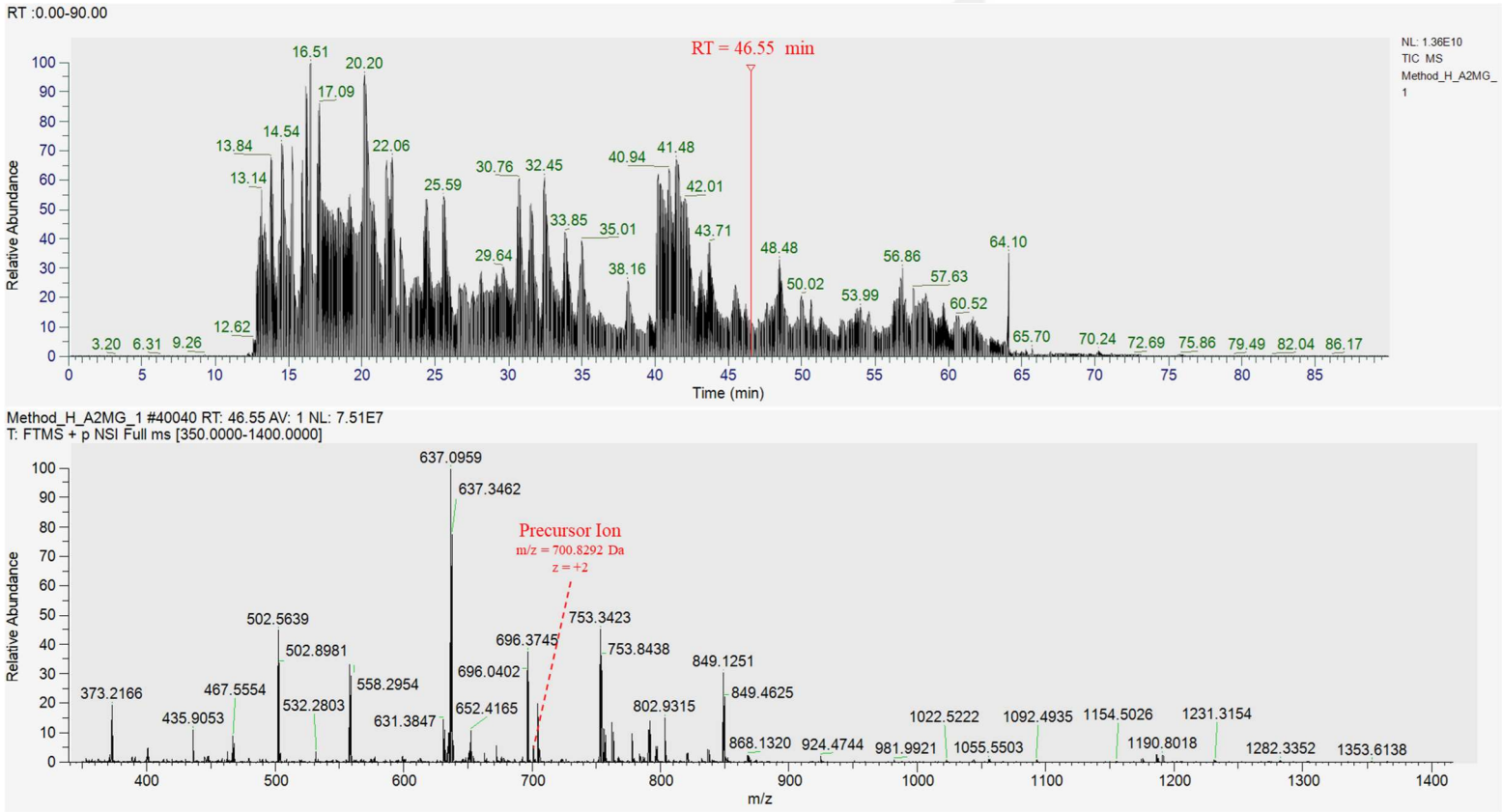


Figure 3.23. Chromatogram and mass spectrum of DMYSFLEDMGLK peptide which was dethiometylated on M2

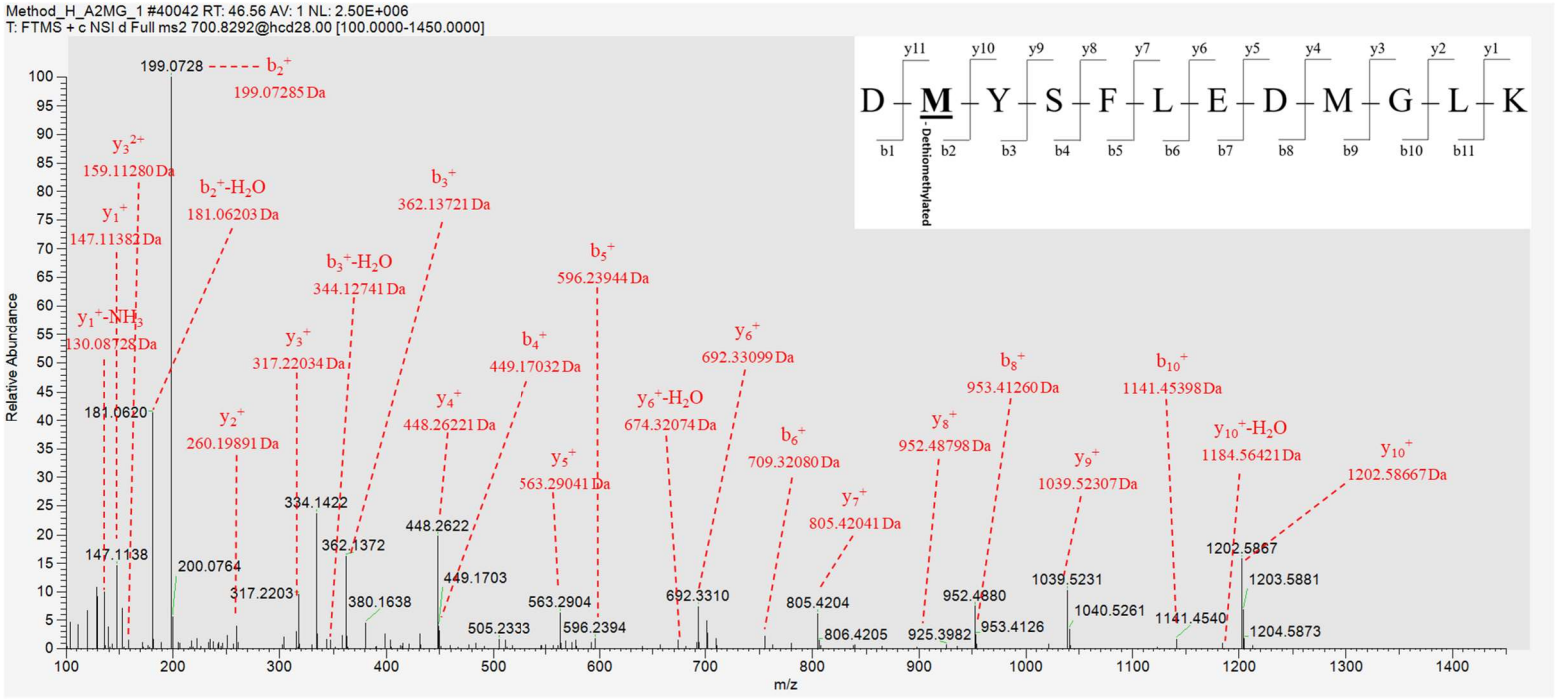


Figure 3.24. MS/MS spectrum of DMYSFLEDMGLK peptide which was dethiometylated on M2

Table 3.16. Monoisotopic mass values in Da for DMYSFLEDMGLK peptide which was dethiometylated on M2

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	116.03422	58.52075	D			12
2	199.07133	100.0393	M-Dethiomethyl	1285.6133	643.31029	11
3	362.13466	181.57097	Y	1202.5762	601.79173	10
4	449.16669	225.08698	S	1039.5129	520.26007	9
5	596.2351	298.62119	F	952.48083	476.74405	8
6	709.31917	355.16322	L	805.41242	403.20985	7
7	838.36176	419.68452	E	692.32835	346.66781	6
8	953.3887	477.19799	D	563.28576	282.14652	5
9	1084.4292	542.71823	M	448.25882	224.63305	4
10	1141.4507	571.22896	G	317.21833	159.1128	3
11	1254.5347	627.771	L	260.19687	130.60207	2
12			K	147.1128	74.06004	1

Mass shift of -48.00337 Da due to dethiomethylation

*Mass values given in red represents the matched values with MS/MS spectra given in Figure 3.24 within 10 ppm range

Table 3.16 showed the monoisotopic mass values of b and y fragments of DMYSFLEDMGLK unique peptide sequence that carries an artificial modification, dethiomethylation, on M2 residue. Red values represented the matched masses in respective dethiometylated DMYSFLEDMGLK sequence in Figure 3.24. Here, the mass shift of -48.00337 Da can be observed effecting b_2 to b_{11} and y_{11} fragments compared to unmodified peptide fragments. Figure 3.23 shows that retention time of precursor ion has changed, and it eluted earlier, indicating that polarity of the peptide changed. Since LC-column used in analysis was a reversed phase C18 column, it can be concluded that net dipole moment (polarity) of the peptide increased, causing it to elute faster from the non-polar stationary phase. Also, the mass of precursor (with +2 charge) decreased by 23.9992 Da which is in agreement with the dethiomethylation modification.

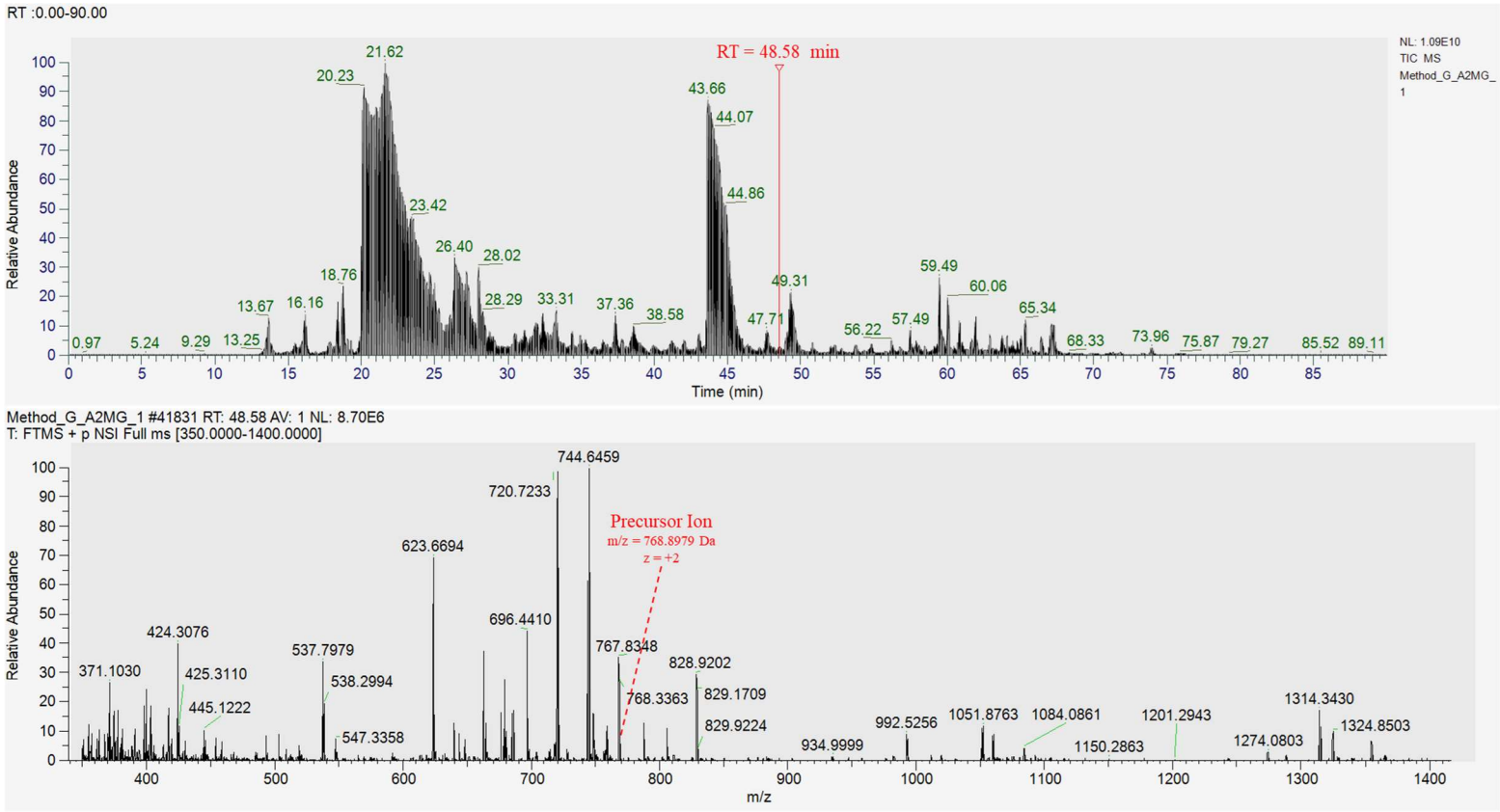


Figure 3.25. Chromatogram and mass spectrum of DMYSFLEDMGLK peptide which was carbamylated on N-terminal and K12

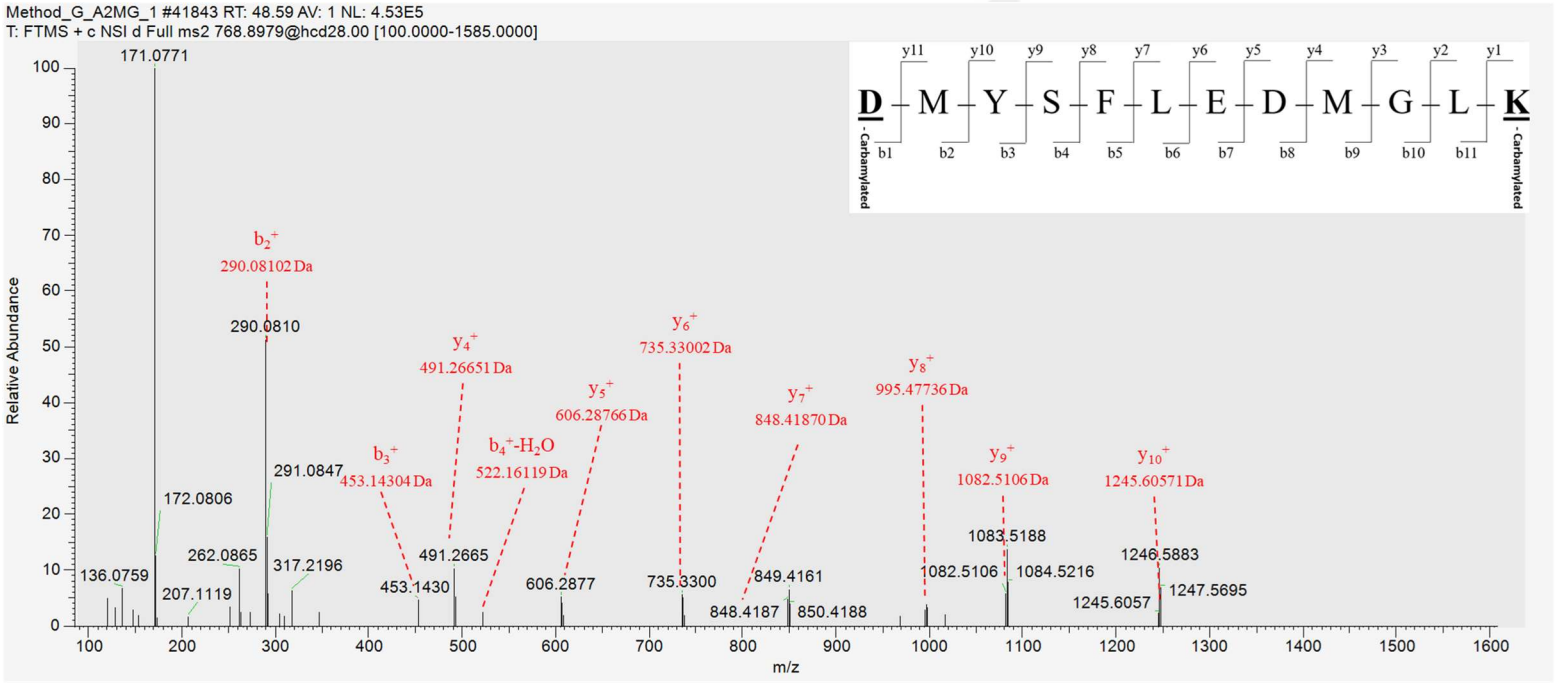


Figure 3.26. MS/MS spectrum of DMYSFLEDMGLK peptide which was carbamylated on N-terminal and K12

Table 3.17. Monoisotopic mass values in Da for DMYSFLEDMGLK peptide which was carbamylated on N-terminal and K12

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	159.04003	80.02366	D-Carbamyl			12
2	290.08052	145.5439	M	1376.6225	688.81488	11
3	453.14385	227.07556	Y	1245.582	623.29464	10
4	540.17587	270.59158	S	1082.5187	541.76297	9
5	687.24429	344.12578	F	995.48664	498.24696	8
6	800.32835	400.66781	L	848.41823	424.71275	7
7	929.37095	465.18911	E	735.33417	368.17072	6
8	1044.3979	522.70258	D	606.29157	303.64943	5
9	1175.4384	588.22282	M	491.26463	246.13595	4
10	1232.4598	616.73356	G	360.22415	180.61571	3
11	1345.5439	673.27559	L	303.20268	152.10498	2
12			K-Carbamyl	190.11862	95.56295	1

Mass shift of 43.00581 Da due to carbamylation

*Mass values given in red represents the matched values with MS/MS spectra given in Figure 3.26. within 10 ppm range

Table 3.17 showed the monoisotopic mass values of b and y fragments of DMYSFLEDMGLK unique peptide sequence was carbamylated on N-term and K12 residue. Red values represented the matched masses in respective modified DMYSFLEDMGLK sequence in Figure 3.26. Here, the mass shift of 43.00581 Da can be observed affecting all fragments compared to unmodified peptide fragments. Since modifications occurred on either end of the peptide, each fragment only shifted by one carbamylation mass. However, Figure 3.25 shows that precursor ion showed a shift of 44.06950 Da which corresponds to half of the mass of two carbamylation with +2 charge within 10 ppm range. In other words, mass shift due to double carbamylation was observed for precursor ion. Additionally, retention time of precursor ion has decreased, indicating an increase in the net dipole moment.

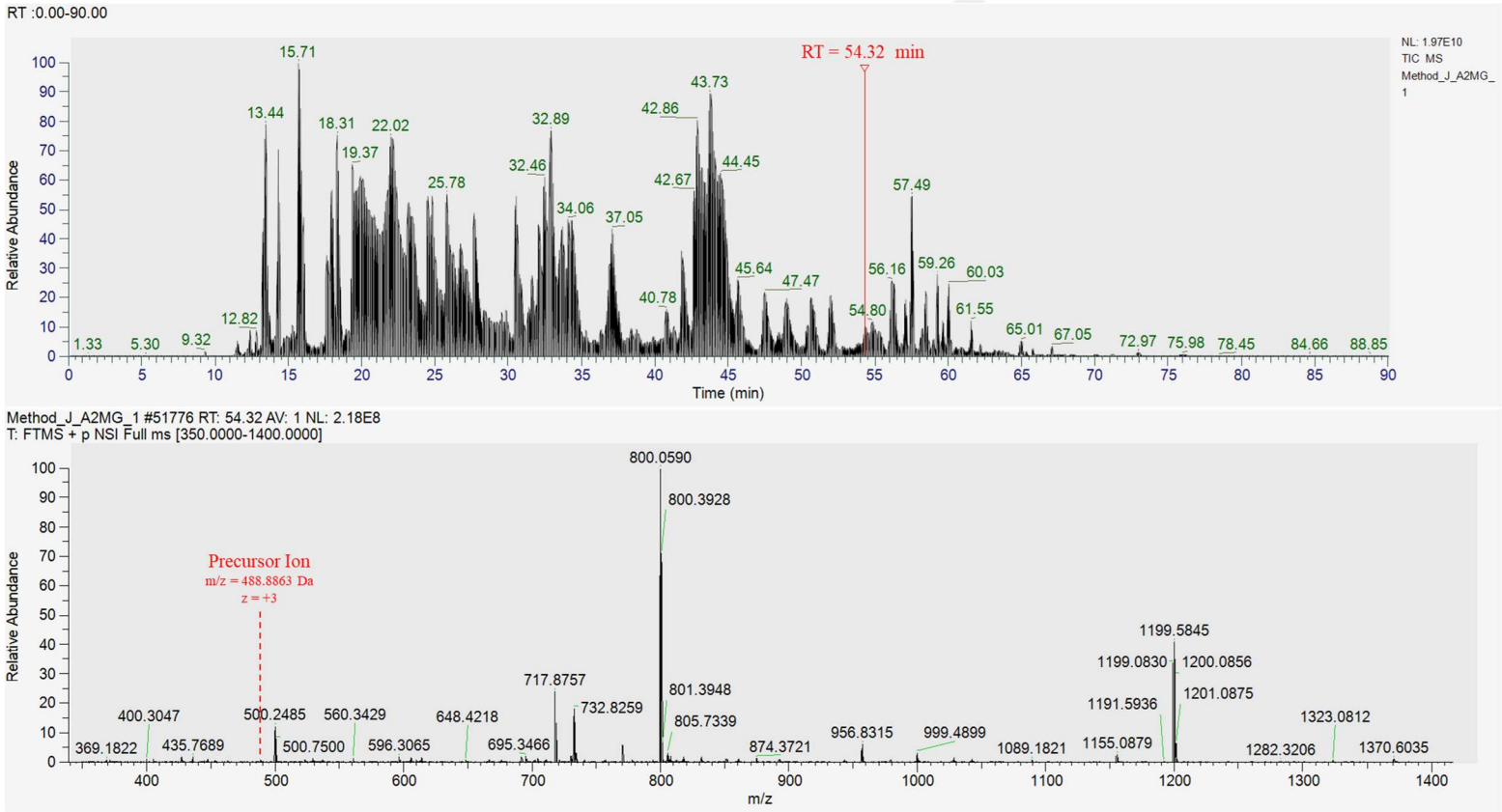


Figure 3.27. Chromatogram and mass spectrum of DMYSFLEDMGLK peptide which was oxidized on M9

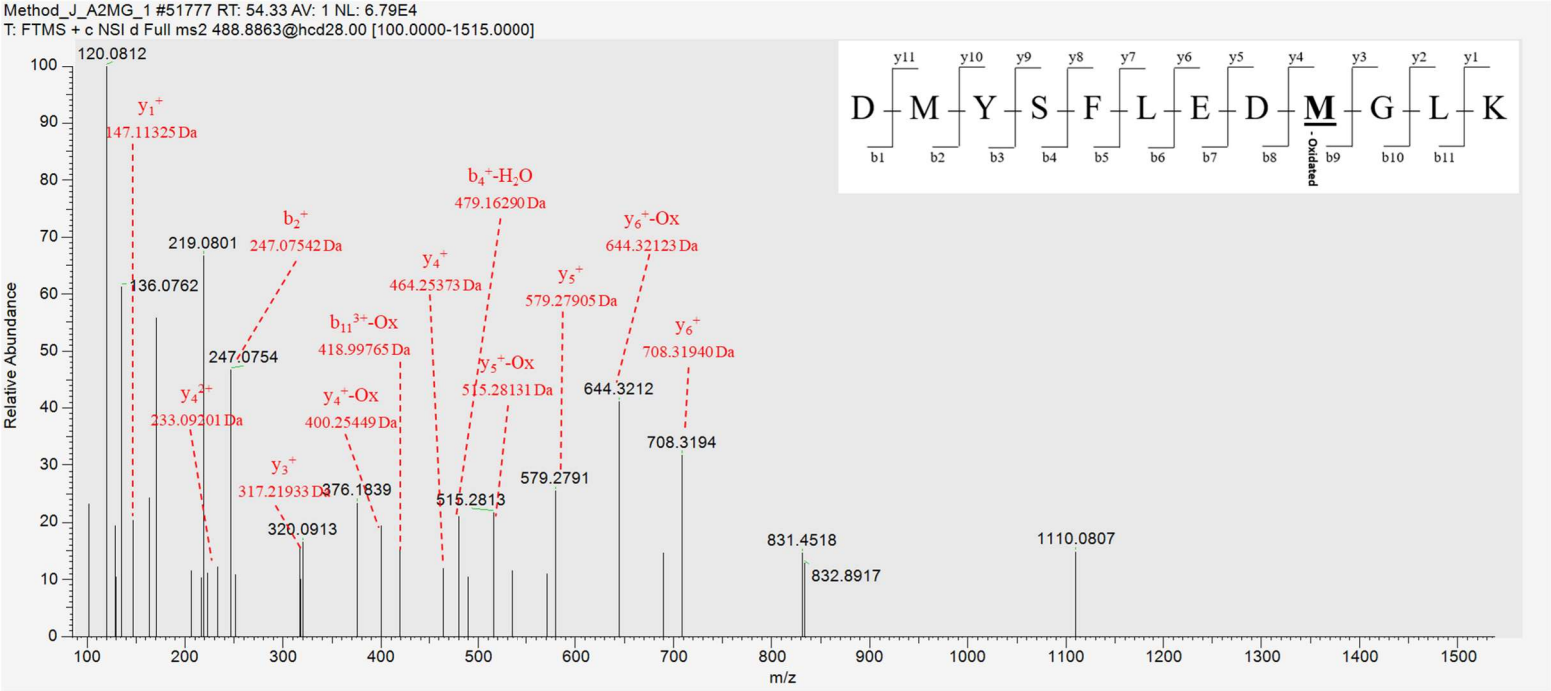


Figure 3.28. MS/MS spectrum of DMYSFLEDMGLK peptide which was oxidized on M9

Table 3.18. Monoisotopic mass values in Da for DMYSFLEDMGLK peptide which was oxidized on M9

#1	b^{+}	b^{2+}	b^{3+}	Seq.	y^{+}	y^{2+}	y^{3+}	#2
1	116.03422	58.52075	39.34959	D				12
2	247.0747	124.04099	83.02975	M	1349.61159	675.30943	450.54205	11
3	410.13803	205.57265	137.3842	Y	1218.5711	609.78919	406.86189	10
4	497.17006	249.08867	166.39487	S	1055.50777	528.25753	352.50744	9
5	644.23847	322.62288	215.41768	F	968.47575	484.74151	323.49677	8
6	757.32254	379.16491	253.11236	L	821.40733	411.2073	274.47396	7
7	886.36513	443.6862	296.12656	E	708.32327	354.66527	236.77927	6
8	1001.39207	501.19968	334.46888	D	579.28067	290.14398	193.76508	5
9	1148.42747	574.71738	383.48068	M-Oxidation	464.25373	232.6305	155.42276	4
10	1205.44894	603.22811	402.48783	G	317.21833	159.1128	106.41096	3
11	1318.533	659.77014	440.18252	L	260.19687	130.60207	87.40381	2
12				K	147.1128	74.06004	49.70912	1

Mass shift of 15.99491 Da compared to oxidation

*Mass values given in red represents the matched values with MS/MS spectra given in Figure 3.28. within 10 ppm range

Table 3.18 showed the monoisotopic mass values of b and y fragments of DMYSFLEDMGLK unique peptide sequence was oxidized on M9 residue. Red values represented the matched masses in respective modified DMYSFLEDMGLK sequence in Figure 3.26. Here, the mass shift of 15.99491 Da can be observed effecting b_9 to b_{11} and y_4 to y_{11} fragments compared to unmodified peptide fragments. Here, a +3 charge was observed for precursor ion and peptide fragments. Figure 3.27 showed that precursor ion with 488.88630 Da mass. Here, peptide mass can be calculated as 1463.63541 Da. Whereas, mass of unmodified precursor with 724.82840 m/z (z=+2) given in Figure 3.21 can be calculated as 1447.64114 Da. It can be observed that mass shift among these precursors correspond to 15.99427 Da which was in agreement with the oxidation modification. Additionally, oxidation modification was observed to hasten the elution of the peptide, increasing its polarity like prior modifications. However, with retention time shift of -7.20 min compared to prior shifts due to dethiomethylation, -14.97 min, and double carbamylation, -12.94 min, showed that net dipole moment of oxidized DMYSFLEDMGLK sequence on M9 was lower than sequences with given modifications.

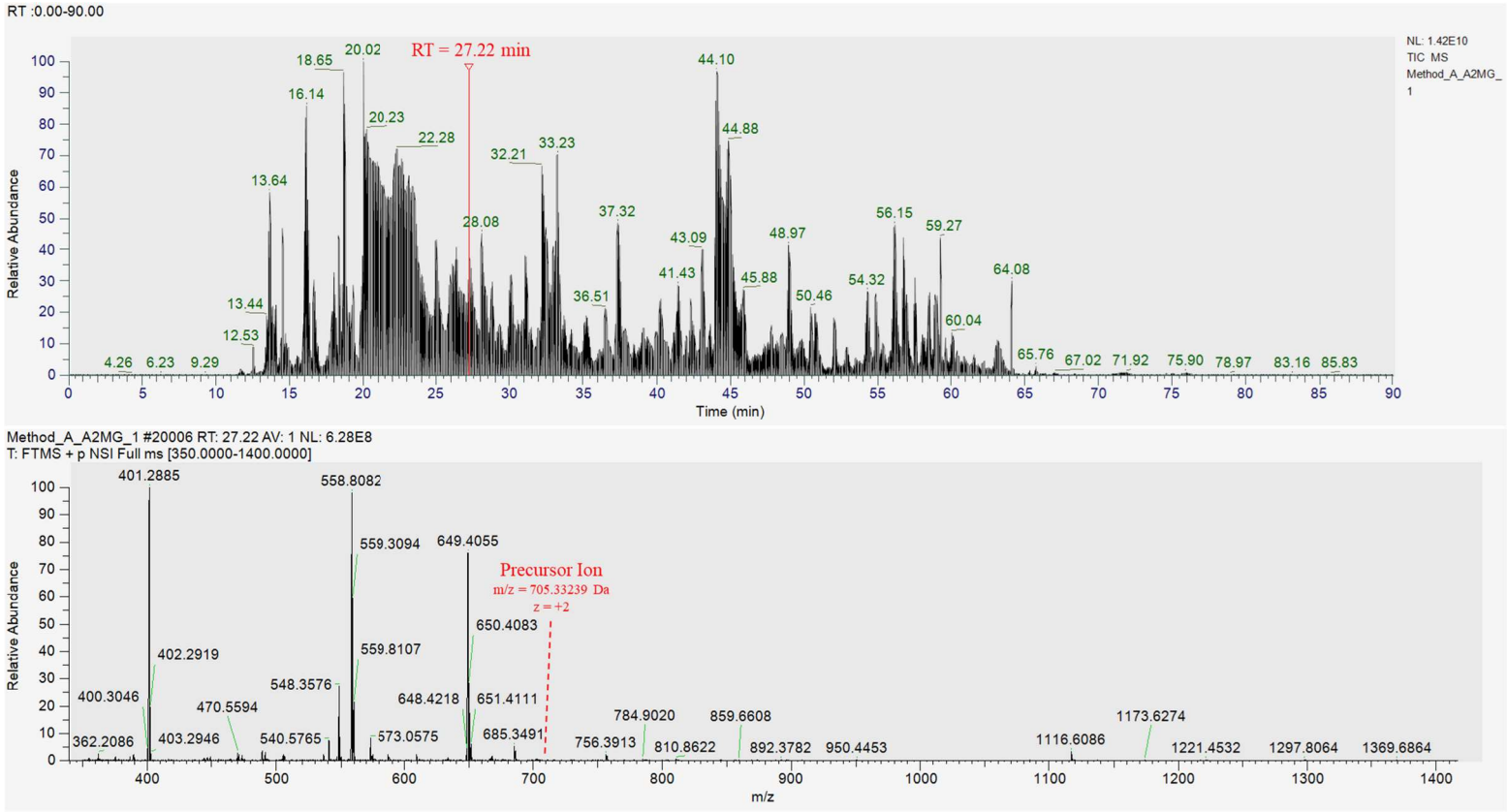


Figure 3.29. Chromatogram and mass spectrum of DMYSFLEDMGLK peptide which was carbamidomethylated on N-terminal, dethiometylated on M2 and M9

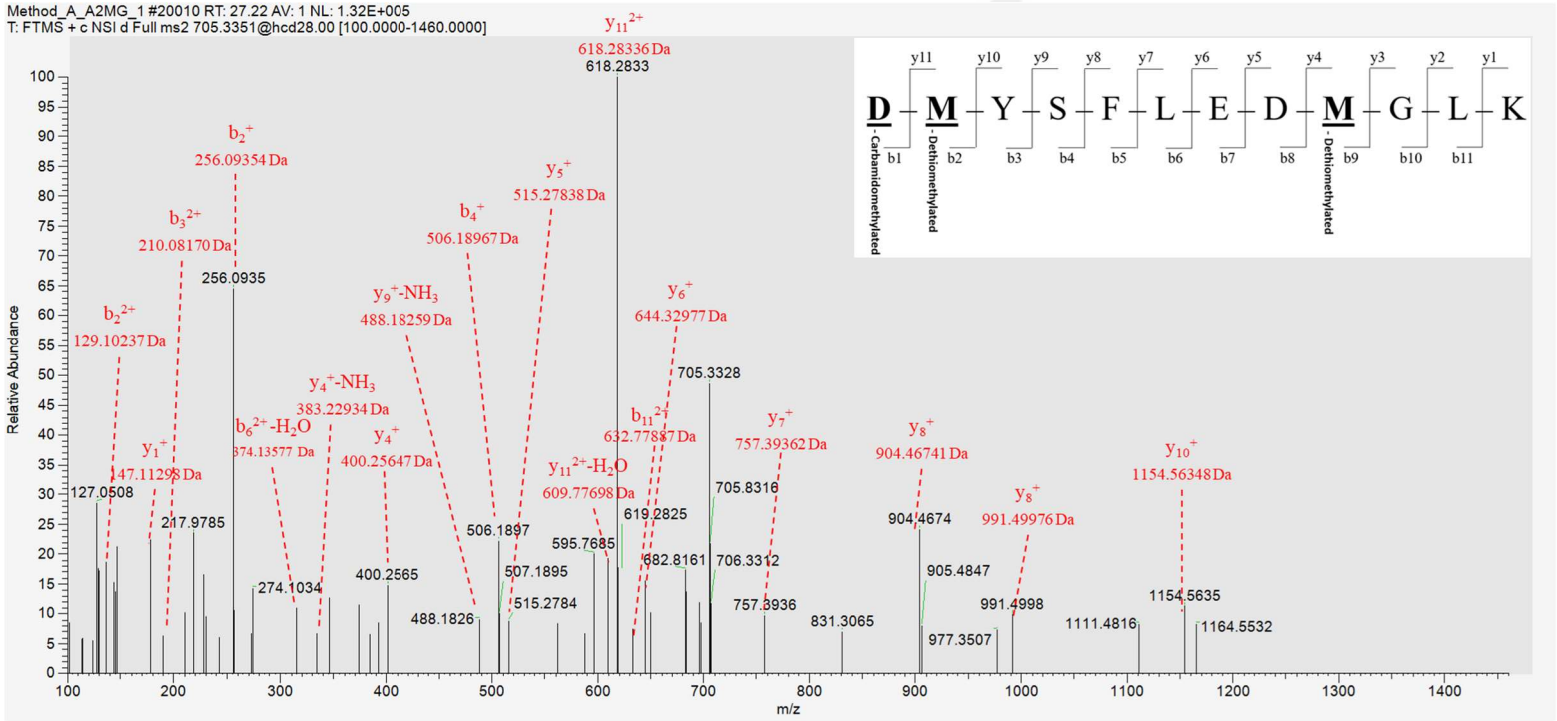
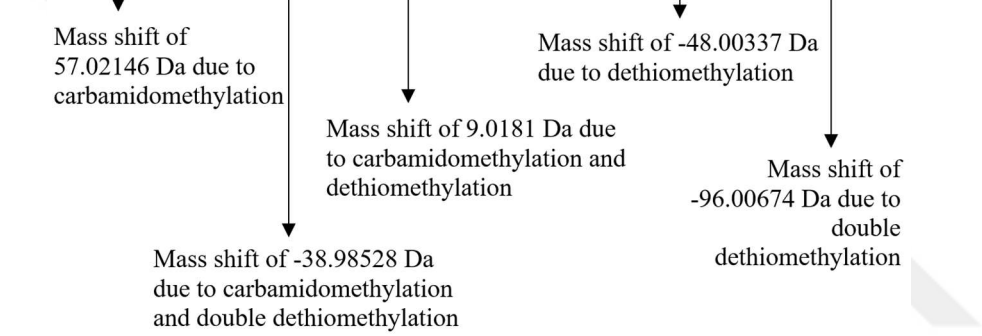


Figure 3.30. MS/MS spectrum of DMYSFLEDMGLK peptide which was carbamidomethylated on N-terminal, dethiometylated on M2 and M9

Table 3.19. Monoisotopic mass values in Da for DMYSFLEDMGLK peptide which was carbamidomethylated on N-terminal, dethiometylated on M2 and M9

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	173.05568	87.03148	D-Carbamidomethyl			12
2	256.0928	128.55004	M-Dethiomethyl	1237.60993	619.3086	11
3	419.15613	210.0817	Y	1154.57282	577.79005	10
4	506.18815	253.59772	S	991.50949	496.25838	9
5	653.25657	327.13192	F	904.47746	452.74237	8
6	766.34063	383.67395	L	757.40905	379.20816	7
7	895.38322	448.19525	E	644.32498	322.66613	6
8	1010.41017	505.70872	D	515.28239	258.14483	5
9	1093.44728	547.22728	M-Dethiomethyl	400.25545	200.63136	4
10	1150.46874	575.73801	G	317.21833	159.1128	3
11	1263.55281	632.28004	L	260.19687	130.60207	2
12			K	147.1128	74.06004	1



*Mass values given in red represents the matched values with MS/MS spectra given in Figure 3.30. within 10 ppm range

Table 3.19 showed the monoisotopic mass values of b and y fragments of DMYSFLEDMGLK unique peptide sequence was carbamidomethylated on N-term and dethiomethylated on M2 and M9 residues. Red values represented the matched masses in respective modified DMYSFLEDMGLK sequence in Figure 3.26. Here, due to three artificial modifications on three different positions, mass shift values varied for fragments. In detail, b_1 fragment shifted 57.02146 Da due to carbamidomethylation on the N-term. b_2 to b_8 fragments shifted 9.01810 Da due to effects of carbamidomethylation and dethiomethylation on M2. b_9 to b_{11} fragments shifted -38.98528 Da due to carbamidomethylation and double dethiomethylation on M2 and M9. For y fragments, y_1 to y_3 fragments did not shift since there were no

modifications in these sequences. y_4 to y_{10} fragments shifted -48.00337 Da due to dethiomethylation of M9 residue. Lastly, y_{11} fragment shifted -96.00674 Da due to double dethiometylation in its sequence. Figure 3.29 showed the precursor ion mass shifted in agreement with the mass shift due to carbamidomethylation and double dethiomethylation. Shift was observed to be 19.49550 Da with the charge of +2. Additionally, retention time of precursor ion observed to be 27.22 min compared to unmodified sequence retention time of 61.52 min. This peptide eluted quickest among the modified DMYSFLEDMGLK peptides given above. This may indicate that as the number of artificial modifications increases, net dipole moment of the peptide increases. Overall, it was observed that artificial modifications effect retention times and mass shifts of precursor ions as well as y and b fragments.

Examples to chromatogram, MS and MS/MS matched were given for carbamidomethylation, dethiomethylation, oxidation, and carbamylation on DMYSFLEDMGLK unique peptide as well as its unmodified form. Deamidation modification could not be shown on DMYSFLEDMGLK peptide because it was found that deamidation occurred on miscleaved sequences (on N residue) including DMYSFLEDMGLK sequence.

Further, PD software identified and matched raw data obtained after the LC/MS analysis similar to examples given in Figures 3.21 to 3.30. Matching was done for all of the protein products obtained with selected ten digestion procedures. Ratio of amounts of artificially modified unique peptides to unmodified unique peptides were given in Figure 3.31. Similar to prior section, most used five unique peptides of A2MG that were given in Figure 3.20 were selected for representation of results. (VGFYESDVMGR, DMYSFLEDMGLK, LLIYAVLPTGDVIGDSAK, LVHVEEPHTETVR, and QFSFPLSSEPFQGSYK)



Figure 3.31. Modification Types and Ratios Form on VGFYESDVMGR, DMYSFLEDMGLK, LLIYAVLPTGDVIGDSAK, LVHVEEPHTETVR, and QFSFPLSSEPFQGSYK

Figure 3.31 showed that while some unique peptide sequences such as VGFYESDVMGR, DMYSFLEDMGLK that include residues with reactive side chains were more prone to artificial modification, other peptides may not show wide range of artificial modifications. However, it was observed that carbamidomethylation was present in peptides that did not carry any other artificial modification. This modification was caused by IAA or CAA reagents used during alkylation.[112] Indeed, carbamidomethylation on Cys is one of the main steps of digestion procedures which is referred to as "alkylation step." Due to this being an

intentional modification, carbamidomethylation of Cys is not an artificial modification.[132] Additionally, this modification is one of the three default modification in MaxQuant and PD software. However, effects of IAA and CAA on other residues are not included in the default settings. In this study, carbamidomethylation was scanned on K residue and N-term of peptides as well as on C residue. Results given for carbamidomethylation in Figure 3.31 only included modifications on these sites (not C residue). It was observed that this modification was most commonly seen on the peptides formed via Method A and I. These methods were the methods that employed highest concentration of IAA with 75 mM and 40 mM, respectively. Method C also used ~40 mM alkylation reagent, however, Figure 3.31 showed lower ratio of this modification for Method C compared to Method I. This may be due to reagent this method employed which was CAA. In literature, it is given that CAA causes less artificial modification than IAA.[112] Further, least carbamidomethylation was seen with Method E. This method employed the lowest concentration of IAA. Additionally, it was the only method that emlyoed lower concentration of alkylation reagent than reducing reagent. This was an important point since IAA reacts with reduction reagent and loses its reactivity, hence loses its ability to artificially modify residues.[110]

Moreover, method J procedure was observed to yield the least amount of modifications compared to other procedures. This can be clearly seen on modification ratios of QFSFPLSSEPFQGSYK peptide. This method is the only one within the selected methods that uses Filter Aided Sample Preparation (FASP). FASP procedure is based on adding of reagents, incubating protein with them, and then filtering away the excess portions. Hence, FASP procedures provides a medium without reagents during the overnight digestion incubation. This might be the reason why peptides formed with Method J showed least amount of modifications.

CHAPTER 4

CONCLUSION

This study investigated the importance of the digestion procedure in the peptide-centric biomarker studies. Firstly, literature search was performed for digestion procedures used in the clinical biomarker studies. This first step aimed to select the most used procedures. However, it was observed that studies applied different digestion procedures. Moreover, although common reagents were observed within scanned studies, concentrations were observed to show variances. Hence, common reagents used in the literature were filtered and ten methods included these common reagents were selected.

In the second step, the effects of selected different reagents and their varying concentrations on the BSA protein were compared using spectrophotometric analysis methods. Comparison started with solubility and quantitative and qualitative analysis methods. The qualitative analysis included UV-Vis spectroscopy. Each digestion method selected from the literature included different buffers that vary in type (ABC and Tris) and concentration (10 mM to 300 mM). So, six replicates of 1.0 mg/ml BSA samples were prepared in each buffer. It was found that ABC solutions created a salting-in effect for the protein, increasing its solubility. This was also valid for 10 mM Tris Buffer. However, higher concentrations of Tris buffer (100 mM, 200 mM, and 300 mM) salted out the protein, lowering its solubility. Quantitative protein analysis utilized BCA Assay Kit that forms a chelate with the protein, creating a purple color that absorbs light at 562 nm. Here, results also showed variances among buffers; however, salting in effects were not observed as clearly as UV-Vis spectroscopy results. Second, the denaturation of proteins with each procedure were compared using CD Spectroscopy. Here, the decrease in the ellipticity of the native

BSA peak at ~220 nm was observed. Peak intensity decrease indicated the unfolding, in other words, denaturation of the BSA. Here, it was observed that neither heat nor harsh reagents are solely enough for the complete unfolding of the protein. Rather, it was dependent on both. Best denaturation was observed at 5% detergent solutions at 60°C incubation temperature Thirdly, the digestion efficiency of the trypsin enzyme in converting proteins into peptides was examined. Surprisingly, it was found that Method F didn't provide the appropriate pH conditions for trypsin to be active. Additionally, methods showed varying efficiencies yielding different amounts of peptides in different method conditions. FASP method yielded the best trypsin activity compared to others. FASP method filters out all of the reagents and resolves the protein samples in 50 mM ABC buffer. Hence, creating a reagent-free medium. 50 mM ABC buffer was also used in Method A. However, this medium contains all of the reagents used during prior steps of this method. This indicated the reduced enzyme activity due to reagents. Hence, prior to enzyme addition, reagents may be removed from digestion medium for best enzyme activity. Fourthly, for comparing digestion procedure efficiencies, total peptide amount formed after digestion was quantitatively measured using Peptide Assay Kit. Results of Methods A, E, and F were inconclusive due to reagent interferences with the kit as seen in the QC-PA quality control set. Here, it was observed that the Peptide kit was susceptible to interferences. Results obtained for samples with interfering reagents yielded misinformation about the peptide concentration. Hence, the compatibility of the mediums containing peptides should be checked before analysis. Nevertheless, except for interfering methods, it was found that methods yielded different concentrations of peptide products, varying ~3 folds among the lowest and the highest values.

In the last part of this study, ten digestion methods were applied to A2MG protein. Protein and peptide groups obtained from methods A to J were compared based on unique peptides and their derivatives including miscleaved peptides and artificial

modifications. It was found that unique peptides obtained from different digestion procedures weren't linearly related to each other. Furthermore, Method F didn't yield two of the unique peptide sequences. Additionally, it was seen that miscleavege sites varied in methods. For instance, VGFYESDVMGR peptide was found to be used the most frequently in literature. VGFYESDVMGR was observed to contribute 15% to 50% to the total of unique peptides, lowest yield in Method J and the highest yield in Method E. 0% miscleavege was observed in five methods (A, B, C, D, J) while three methods (E, H, I) yielded more than 50% miscleavege for VGFYESDVMGR sequence. Second, the artificial modification of the unique peptides was investigated over the five most frequently used sequences, which were VGFYESDVMGR, DMYSFLEDMGLK, LLIYAVLPTGDVIGDSAK, LVHVEEPHTETVR, and QFSFPLSSEPFQGSYK. Artificial modifications are the uncontrolled side reactions that occur during the digestion procedure. Since they cause mass shifts, unidentified artificial modification damages the integrity of the quantitative results used in peptide-centric biomarker studies. It was observed that some unique peptides were more prone to these modifications. Unique peptides like DMYSFLEDMGLK and VGFYESDVMGR, which include residues with active side groups like Met and Lys, were more susceptible to modifications. Additionally, it was observed that the type and amount of reagents used in digestion processes affect the formation of artificial modifications on residues. Alkylation reagents, IAA and CAA, were observed to frequently cause carbamidomethylation on the N-term of the peptides. This artificial modification was seen in unique peptide sequences which showed no other artificial modifications such as LLIYAVLPTGDVIGDSAK, LVHVEEPHTETVR, and QFSFPLSSEPFQGSYK. It was observed that the default setting in data analysis software, such as MaxQuant and PD, didn't include carbamidomethylation on N-term and K. This may cause misidentification or exclusion of unique peptides with artificial modifications.

This study showed that different reagents and their concentrations used in the literature showed varying effects on proteins and their products. More specifically, protein dissolved and denatured differently, and trypsin enzyme showed varying activities in different mediums of methods as seen with spectrophotometric analyses. Furthermore, LC/MS analyses showed that different digestion procedures yield inequivalent protein products for equivalent samples. Unique peptides showed varying miscleavege sites. Additionally, different types of artificial modifications were seen on varying residue sites. These miscleavege and artificial modifications were observed to effect retention time and mass of precursor as well as masses of the respective fragments. Hence, miscleavege and artificial modifications were found to be included in the software data analysis for a complete identification and quantification of the peptide products. In conclusion, digestion procedures were found to effect protein products. Hence, digestion procedures should be considered as a variation source and detailed information of these procedures and analysis parameters should be included in studies for consistent analysis studies for proteins.

REFERENCES

(1) LaPelusa, A.; Kaushik, R. Physiology, Proteins. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2023.

(2) Kavallaris, M.; Marshall, G. M. Proteomics and Disease: Opportunities and Challenges. *Med. J. Aust.* **2005**, *182* (11), 575–579. https://doi.org/10.5694/j.1326-5377.2005.tb06817.x.

(3) Rappsilber, J.; Mann, M. What Does It Mean to Identify a Protein in Proteomics? *Trends Biochem. Sci.* **2002**, *27* (2), 74–78. https://doi.org/10.1016/S0968-0004(01)02021-7.

(4) Garrels, J. I. Proteome. In *Encyclopedia of Genetics*; Brenner, S., Miller, J. H., Eds.; Academic Press: New York, 2001; pp 1575–1578. https://doi.org/10.1006/rwgn.2001.1042.

(5) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; Ge, Y.; Gunawardena, J.; Hendrickson, R. C.; Hergenrother, P. J.; Huber, C. G.; Ivanov, A. R.; Jensen, O. N.; Jewett, M. C.; Kelleher, N. L.; Kiessling, L. L.; Krogan, N. J.; Larsen, M. R.; Loo, J. A.; Ogorzalek Loo, R. R.; Lundberg, E.; MacCoss, M. J.; Mallick, P.; Mootha, V. K.; Mrksich, M.; Muir, T. W.; Patrie, S. M.; Pesavento, J. J.; Pitteri, S. J.; Rodriguez, H.; Saghatelian, A.; Sandoval, W.; Schlüter, H.; Sechi, S.; Slavoff, S. A.; Smith, L. M.; Snyder, M. P.; Thomas, P. M.; Uhlén, M.; Van Eyk, J. E.; Vidal, M.; Walt, D. R.; White, F. M.; Williams, E. R.; Wohlschlager, T.; Wysocki, V. H.; Yates, N. A.; Young, N. L.; Zhang, B. How Many Human Proteoforms Are There? *Nat. Chem. Biol.* **2018**, *14* (3), 206–214. https://doi.org/10.1038/nchembio.2576.

(6) Corthals, G. L.; Wasinger, V. C.; Hochstrasser, D. F.; Sanchez, J.-C. The Dynamic Range of Protein Expression: A Challenge for Proteomic Research. *Electrophoresis* **2000**, *21* (6), 1104–1115.

https://doi.org/10.1002/(SICI)1522-2683(20000401)21:6<1104::AID-ELPS1104>3.0.CO;2-C.

(7) Perdivara, I.; Deterding, L. J.; Przybylski, M.; Tomer, K. B. Mass Spectrometric Identification of Oxidative Modifications of Tryptophan Residues in Proteins: Chemical Artifact or Post-Translational Modification? *J. Am. Soc. Mass Spectrom.* **2010**, *21* (7), 1114–1117. https://doi.org/10.1016/j.jasms.2010.02.016.

(8) Creasy, D. M.; Cottrell, J. S. Unimod: Protein Modifications for Mass Spectrometry. *Proteomics* **2004**, *4* (6), 1534–1536. https://doi.org/10.1002/pmic.200300744.

(9) Ren, D.; Pipes, G. D.; Liu, D.; Shih, L.-Y.; Nichols, A. C.; Treuheit, M. J.; Brems, D. N.; Bondarenko, P. V. An Improved Trypsin Digestion Method Minimizes Digestion-Induced Modifications on Proteins. *Anal. Biochem.* **2009**, *392* (1), 12–21. https://doi.org/10.1016/j.ab.2009.05.018.

(10) Marier, J. R.; Rose, D. Determination of Cyanate, and a Study of Its Accumulation in Aqueous Solutions of Urea. *Anal. Biochem.* **1964**, *7* (3), 304–314. https://doi.org/10.1016/0003-2697(64)90135-6.

(11) Verbrugge, F. H.; Tang, W. H. W.; Hazen, S. L. Protein Carbamylation and Cardiovascular Disease. *Kidney Int.* **2015**, *88* (3), 474–478. https://doi.org/10.1038/ki.2015.166.

(12) Sun, S.; Zhou, J.-Y.; Yang, W.; Zhang, H. Inhibition of Protein Carbamylation in Urea Solution Using Ammonium Containing Buffers. *Anal. Biochem.* **2014**, *446*, 76. https://doi.org/10.1016/j.ab.2013.10.024.

(13) Duncan, M. W.; Aebersold, R.; Caprioli, R. M. The Pros and Cons of Peptide-Centric Proteomics. *Nat. Biotechnol.* **2010**, *28* (7), 659–664. https://doi.org/10.1038/nbt0710-659.

(14) Manzoni, C.; Kia, D. A.; Vandrovcova, J.; Hardy, J.; Wood, N. W.; Lewis, P. A.; Ferrari, R. Genome, Transcriptome and Proteome: The Rise of Omics Data

and Their Integration in Biomedical Sciences. *Brief. Bioinform.* **2016**, *19* (2), 286–302. https://doi.org/10.1093/bib/bbw114.

(15) Wasinger, V. C.; Cordwell, S. J.; Cerpa-Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I. Progress with Gene-Product Mapping of the Mollicutes: Mycoplasma Genitalium. *Electrophoresis* **1995**, *16* (1), 1090–1094. https://doi.org/10.1002/elps.11501601185.

(16) Aslam, B.; Basit, M.; Nisar, M. A.; Khurshid, M.; Rasool, M. H. Proteomics: Technologies and Their Applications. *J. Chromatogr. Sci.* **2017**, *55* (2), 182–196. https://doi.org/10.1093/chromsci/bmw167.

(17) Zengin, T.; Kılıç, S.; Yenigün, U.; Erdogmus, S.; Baysal, Ö. A Smart Way for Talking with Proteins; Proteomics. *MOJ Proteomics Bioinforma.* **2017**, *5*, 00160. https://doi.org/10.15406/mojpb.2017.05.00160.

(18) Verrills, N. M. Clinical Proteomics: Present and Future Prospects. *Clin. Biochem. Rev.* **2006**, *27* (2), 99–116.

(19) U.S. Food and Drug Administration. List of Qualified Biomarkers. *FDA* **2021**.

(20) FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and Other Tools) Resource*; Food and Drug Administration (US): Silver Spring (MD), 2016.

(21) Aronson, J. K.; Ferner, R. E. Biomarkers-A General Review. *Curr. Protoc. Pharmacol.* **2017**, *76*, 9.23.1-9.23.17. https://doi.org/10.1002/cpph.19.

(22) Rifai, N.; Gillette, M. A.; Carr, S. A. Protein Biomarker Discovery and Validation: The Long and Uncertain Path to Clinical Utility. *Nat. Biotechnol.* **2006**, *24* (8), 971–983. https://doi.org/10.1038/nbt1235.

(23) Gisbert, J. P.; Chaparro, M. Clinical Usefulness of Proteomics in Inflammatory Bowel Disease: A Comprehensive Review. *J. Crohns Colitis* **2019**, *13* (3), 374–384. https://doi.org/10.1093/ecco-jcc/jjy158.

(24) Végvári, A.; Kondo, T.; Marshall, J. G. Clinical Proteomics. *Int. J. Proteomics* **2012**, *2012*, 641491. https://doi.org/10.1155/2012/641491.

(25) Issaq, H. J.; Xiao, Z.; Veenstra, T. D. Serum and Plasma Proteomics. *Chem. Rev.* **2007**, *107* (8), 3601–3620. https://doi.org/10.1021/cr068287r.

(26) Vignoli, A.; Tenori, L.; Morsiani, C.; Turano, P.; Capri, M.; Luchinat, C. Serum or Plasma (and Which Plasma), That Is the Question. *J. Proteome Res.* **2022**, *21* (4), 1061–1072. https://doi.org/10.1021/acs.jproteome.1c00935.

(27) Geyer, P. E.; Holdt, L. M.; Teupser, D.; Mann, M. Revisiting Biomarker Discovery by Plasma Proteomics. *Mol. Syst. Biol.* **2017**, *13* (9), 942. https://doi.org/10.15252/msb.20156297.

(28) Mathew, J.; Sankar, P.; Varacallo, M. Physiology, Blood Plasma. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2023.

(29) Paul, J.; Veenstra, T. D. Separation of Serum and Plasma Proteins for In-Depth Proteomic Analysis. *Separations* **2022**, *9* (4), 89. https://doi.org/10.3390/separations9040089.

(30) The Human Protein Atlas. *The human blood proteins - secreted to blood - The Human Protein Atlas*. https://www.proteinatlas.org/humanproteome/blood+protein/secreted+to+blood (accessed 2023-12-19).

(31) Uhlén, M.; Karlsson, M. J.; Hober, A.; Svensson, A.-S.; Scheffel, J.; Kotol, D.; Zhong, W.; Tebani, A.; Strandberg, L.; Edfors, F.; Sjöstedt, E.; Mulder, J.; Mardinoglu, A.; Berling, A.; Ekblad, S.; Dannemeyer, M.; Kanje, S.; Rockberg, J.; Lundqvist, M.; Malm, M.; Volk, A.-L.; Nilsson, P.; Månberg, A.; Dodig-Crnkovic, T.; Pin, E.; Zwahlen, M.; Oksvold, P.; von Feilitzen, K.; Häussler, R. S.; Hong, M.-G.; Lindskog, C.; Ponten, F.; Katona, B.; Vuu, J.; Lindström, E.; Nielsen, J.; Robinson, J.; Ayoglu, B.; Mahdessian, D.; Sullivan, D.; Thul, P.; Danielsson, F.; Stadler, C.; Lundberg, E.; Bergström, G.; Gummesson, A.; Voldborg, B. G.; Tegel, H.; Hober, S.; Forsström, B.; Schwenk, J. M.; Fagerberg, L.; Sivertsson, Å. The Human Secretome. *Sci. Signal.* **2019**, *12* (609), eaaz0274. https://doi.org/10.1126/scisignal.aaz0274.

(32) Anderson, N. L.; Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects. *Mol. Cell. Proteomics* **2002**, *1* (11), 845–867. https://doi.org/10.1074/mcp.R200007-MCP200.

(33) Sitar, M. E.; Aydin, S.; Cakatay, U. Human Serum Albumin and Its Relation with Oxidative Stress. *Clin. Lab.* **2013**, *59* (9–10), 945–952.

(34) Goktayoglu, E.; Oztop, M. H.; Ozcan, S. Proteomics Approach to Differentiate Protein Extraction Methods in Sugar Beet Leaves. *J. Agric. FOOD Chem.* **2023**, *71* (23), 9157–9163. https://doi.org/10.1021/acs.jafc.2c09190.

(35) Carrera, M. Proteomics and Food Analysis: Principles, Techniques, and Applications. *Foods* **2021**, *10* (11), 2538. https://doi.org/10.3390/foods10112538.

(36) Romsdahl, J.; Blachowicz, A.; Chiang, A. J.; Chiang, Y.-M.; Masonjones, S.; Yaegashi, J.; Countryman, S.; Karouia, F.; Kalkum, M.; Stajich, J. E.; Venkateswaran, K.; Wang, C. C. C. International Space Station Conditions Alter Genomics, Proteomics, and Metabolomics in Aspergillus Nidulans. *Appl. Microbiol. Biotechnol.* **2019**, *103* (3), 1363–1377. https://doi.org/10.1007/s00253-018-9525-0.

(37) La Barbera, G.; Capriotti, A. L.; Michelini, E.; Piovesana, S.; Calabretta, M. M.; Zenezini Chiozzi, R.; Roda, A.; Laganà, A. Proteomic Analysis and Bioluminescent Reporter Gene Assays to Investigate Effects of Simulated Microgravity on Caco-2 Cells. *Proteomics* **2017**, *17* (15–16). https://doi.org/10.1002/pmic.201700081.

(38) Lopez, E.; Madero, L.; Melen, G. J.; Ramirez, M.; Roukos, D. H.; Cho, W. C. Clinical Proteomics in Cancer Research. *Curr. PROTEOMICS* **2013**, *10* (2), 179–191. https://doi.org/10.2174/15701646112099990001.

(39) Plymoth, A.; Hainaut, P. Proteomics beyond Proteomics: Toward Clinical Applications. *Curr. Opin. Oncol.* **2011**, *23* (1), 77–82. https://doi.org/10.1097/CCO.0b013e32834179c1.

(40) Noble, J. E.; Bailey, M. J. A. Quantitation of Protein. *Methods Enzymol.* **2009**, *463*, 73–95. https://doi.org/10.1016/S0076-6879(09)63008-1.

(41) Simonian, M. H. Spectrophotometric Determination of Protein Concentration. *Curr. Protoc. Cell Biol.* **2002**, *15* (1), A.3B.1-A.3B.7. https://doi.org/10.1002/0471143030.cba03bs15.

(42) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2006**, *1* (6), 2876–2890. https://doi.org/10.1038/nprot.2006.202.

(43) Krohn, R. I. The Colorimetric Detection and Quantitation of Total Protein. *Curr. Protoc. Cell Biol.* **2011**, *52* (1), A.3H.1-A.3H.28. https://doi.org/10.1002/0471143030.cba03hs52.

(44) Edman, P. Edman: Method for Determination of the Amino Acid Sequence in Peptides. *Acta Chem. Scand.* **1950**, No. 4, 283–293.

(45) Zimmer, M. Green Fluorescent Protein (GFP): Applications, Structure, and Related Photophysical Behavior. *Chem. Rev.* **2002**, *102* (3), 759–782. https://doi.org/10.1021/cr010142r.

(46) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. Protein Structure Determination from NMR Chemical Shifts. *Proc. Natl. Acad. Sci.* **2007**, *104* (23), 9615–9620. https://doi.org/10.1073/pnas.0610313104.

(47) Yee, A. A.; Savchenko, A.; Ignachenko, A.; Lukin, J.; Xu, X.; Skarina, T.; Evdokimova, E.; Liu, C. S.; Semesi, A.; Guido, V.; Edwards, A. M.; Arrowsmith, C. H. NMR and X-Ray Crystallography, Complementary Tools in Structural Proteomics of Small Proteins. *J. Am. Chem. Soc.* **2005**, *127* (47), 16512–16517. https://doi.org/10.1021/ja053565+.

(48) Lequin, R. M. Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA). *Clin. Chem.* **2005**, *51* (12), 2415–2418. https://doi.org/10.1373/clinchem.2005.051532.

(49) Jiang, X.; Wu, M.; Albo, J.; Rao, Q. Non-Specific Binding and Cross-Reaction of ELISA: A Case Study of Porcine Hemoglobin Detection. *Foods* **2021**, *10* (8), 1708. https://doi.org/10.3390/foods10081708.

(50) Schubert, O. T.; Röst, H. L.; Collins, B. C.; Rosenberger, G.; Aebersold, R. Quantitative Proteomics: Challenges and Opportunities in Basic and Applied Research. *Nat. Protoc.* **2017**, *12* (7), 1289–1294. https://doi.org/10.1038/nprot.2017.040.

(51) Ornberg, R. L.; Harper, T. F.; Liu, H. Western Blot Analysis with Quantum Dot Fluorescence Technology: A Sensitive and Quantitative Method for Multiplexed Proteomics, 2005. https://www.nature.com/articles/nmeth0105-79 (accessed 2023-12-20).

(52) Birhanu, A. G. Mass Spectrometry-Based Proteomics as an Emerging Tool in Clinical Laboratories. *Clin. Proteomics* **2023**, *20* (1), 32. https://doi.org/10.1186/s12014-023-09424-x.

(53) Domon, B.; Aebersold, R. Mass Spectrometry and Protein Analysis. *Science* **2006**, *312* (5771), 212–217. https://doi.org/10.1126/science.1124619.

(54) Stoscheck, C. M. Quantitation of Protein. In *Methods in Enzymology*; Deutscher, M. P., Ed.; Guide to Protein Purification; Academic Press, 1990; Vol. 182, pp 50–68. https://doi.org/10.1016/0076-6879(90)82008-P.

(55) Fornander, L. H.; Feng, B.; Beke-Somfai, T.; Nordén, B. UV Transition Moments of Tyrosine. *J. Phys. Chem. B* **2014**, *118* (31), 9247–9257. https://doi.org/10.1021/jp5065352.

(56) Antosiewicz, J. M.; Shugar, D. UV–Vis Spectroscopy of Tyrosine Side-Groups in Studies of Protein Structure. Part 2: Selected Applications. *Biophys. Rev.* **2016**, *8* (2), 163–177. https://doi.org/10.1007/s12551-016-0197-7.

(57) Prasad, S.; Mandal, I.; Singh, S.; Paul, A.; Mandal, B.; Venkatramani, R.; Swaminathan, R. Near UV-Visible Electronic Absorption Originating from Charged Amino Acids in a Monomeric Protein †Electronic Supplementary

Information (ESI) Available. See DOI: 10.1039/C7sc00880e. *Chem. Sci.* **2017**, *8* (8), 5416–5433. https://doi.org/10.1039/c7sc00880e.

(58) Mach, H.; Middaugh, C. R. Simultaneous Monitoring of the Environment of Tryptophan, Tyrosine, and Phenylalanine Residues in Proteins by Near-Ultraviolet Second-Derivative Spectroscopy. *Anal. Biochem.* **1994**, *222* (2), 323–331. https://doi.org/10.1006/abio.1994.1499.

(59) Layne, E. Spectrophotometric and Turbidimetric Methods for Measuring Proteins. **1957**, *3*, 447–454. https://doi.org/10.1016/S0076-6879(57)03413-8.

(60) Scopes, R. K. Measurement of Protein by Spectrophotometry at 205 Nm. *Anal. Biochem.* **1974**, *59* (1), 277–282. https://doi.org/10.1016/0003-2697(74)90034-7.

(61) Rodger, A.; Marrington, R.; Roper, D.; Windsor, S. Circular Dichroism Spectroscopy for the Study of Protein-Ligand Interactions. *Methods Mol. Biol. Clifton NJ* **2005**, *305*, 343–364. https://doi.org/10.1385/1-59259-912-5:343.

(62) Siligardi, G.; Hussain, R. Circular Dichroism, Applications. In *Encyclopedia of Spectroscopy and Spectrometry (Third Edition)*; Lindon, J. C., Tranter, G. E., Koppenaal, D. W., Eds.; Academic Press: Oxford, 2017; pp 293–298. https://doi.org/10.1016/B978-0-12-803224-4.00031-5.

(63) Dodero, V.; Quirolo, Z.; Sequeira, A. Biomolecular Studies by Circular Dichroism. *Front. Biosci. J. Virtual Libr.* **2011**, *16*, 61–73. https://doi.org/10.2741/3676.

(64) Rodger, A.; Nordén, B.; Rodger, A.; Nordén, B. *Circular Dichroism and Linear Dichroism*; Oxford Chemistry Masters; Oxford University Press: Oxford, New York, 1997.

(65) *Circular Dichroism: Principles and Applications, 2nd Edition | Wiley*. Wiley.com.https://www.wiley.com/en-us/Circular+Dichroism%3A+Principles+and+Applications%2C+2nd+Edition-p-9780471330035 (accessed 2023-12-22).

(66) Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72*, 248–254. https://doi.org/10.1006/abio.1976.9999.

(67) Gornall, A. G.; Bardawill, C. J.; David, M. M. Determination of Serum Proteins By Means of The Biuret Reaction. *J. Biol. Chem.* **1949**, *177* (2), 751–766. https://doi.org/10.1016/S0021-9258(18)57021-6.

(68) Chen, L.; Yu, Z.; Lee, Y.; Wang, X.; Zhao, B.; Jung, Y. M. Quantitative Evaluation of Proteins with Bicinchoninic Acid (BCA): Resonance Raman and Surface-Enhanced Resonance Raman Scattering-Based Methods. *Analyst* **2012**, *137* (24), 5834–5838. https://doi.org/10.1039/C2AN35715A.

(69) Cortés-Ríos, J.; Zárate, A. M.; Figueroa, J. D.; Medina, J.; Fuentes-Lemus, E.; Rodríguez-Fernández, M.; Aliaga, M.; López-Alarcón, C. Protein Quantification by Bicinchoninic Acid (BCA) Assay Follows Complex Kinetics and Can Be Performed at Short Incubation Times. *Anal. Biochem.* **2020**, *608*, 113904. https://doi.org/10.1016/j.ab.2020.113904.

(70) Wiechelman, K. J.; Braun, R. D.; Fitzpatrick, J. D. Investigation of the Bicinchoninic Acid Protein Assay: Identification of the Groups Responsible for Color Formation. *Anal. Biochem.* **1988**, *175* (1), 231–237. https://doi.org/10.1016/0003-2697(88)90383-1.

(71) Strickland, R. D.; Freeman, M. L.; Gurule, F. T. Copper Binding by Proteins in Alkaline Solution. *Anal. Chem.* **1961**, *33* (4), 545–552. https://doi.org/10.1021/ac60172a019.

(72) Lowry, OliverH.; Rosebrough, NiraJ.; Farr, A. L.; Randall, RoseJ. Protein Measurement with The Folin Phenol Reagent. *J. Biol. Chem.* **1951**, *193* (1), 265–275. https://doi.org/10.1016/S0021-9258(19)52451-6.

(73) Smith, P. K.; Krohn, R. I.; Hermanson, G. T.; Mallia, A. K.; Gartner, F. H.; Provenzano, M. D.; Fujimoto, E. K.; Goeke, N. M.; Olson, B. J.; Klenk, D. C.

Measurement of Protein Using Bicinchoninic Acid. *Anal. Biochem.* **1985**, *150* (1), 76–85. https://doi.org/10.1016/0003-2697(85)90442-7.

(74) Walker, J. M. The Bicinchoninic Acid (BCA) Assay for Protein Quantitation. In *Basic Protein and Peptide Protocols*; Walker, J. M., Ed.; Methods in Molecular Biology™; Humana Press: Totowa, NJ, 1994; pp 5–8. https://doi.org/10.1385/0-89603-268-X:5.

(75) Smith, G. F.; Wilkins, D. H. New Colorimetric Reagent Specific for Cooper. *Anal. Chem.* **1953**, *25* (3), 510–511. https://doi.org/10.1021/ac60075a037.

(76) Chen, L.; Xue, X.; Jiang, D.; Yang, J.; Zhao, B.; Han, X.; Jung, Y. M. A Turn-On Resonance Raman Scattering (BCS/Cu +) Sensor for Quantitative Determination of Proteins. *Appl. Spectrosc.* **2015**, *70*, 0003702815620560. https://doi.org/10.1177/0003702815620560.

(77) Matsushita, M.; Irino, T.; Komoda, T.; Sakagishi, Y. Determination of Proteins by a Reverse Biuret Method Combined with the Copper-Bathocuproine Chelate Reaction. *Clin. Chim. Acta* **1993**, *216* (1), 103–111. https://doi.org/10.1016/0009-8981(93)90143-R.

(78) Han, X.; Aslanian, A.; Yates, J. R. Mass Spectrometry for Proteomics. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 483–490. https://doi.org/10.1016/j.cbpa.2008.07.024.

(79) Fenselau, C. Beyond Gene Sequencing: Analysis of Protein Structure with Mass Spectrometry. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20* (1), 205–220. https://doi.org/10.1146/annurev.bb.20.060191.001225.

(80) Neagu, A.-N.; Jayathirtha, M.; Baxter, E.; Donnelly, M.; Petre, B. A.; Darie, C. C. Applications of Tandem Mass Spectrometry (MS/MS) in Protein Analysis for Biomedical Research. *Molecules* **2022**, *27* (8), 2411. https://doi.org/10.3390/molecules27082411.

(81) Baldwin, M. A. Protein Identification by Mass Spectrometry: Issues to Be Considered. *Mol. Cell. Proteomics* **2004**, *3* (1), 1–9. https://doi.org/10.1074/mcp.R300012-MCP200.

(82) Silva, J. C.; Gorenstein, M. V.; Li, G.-Z.; Vissers, J. P. C.; Geromanos, S. J. Absolute Quantification of Proteins by LCMSE: A Virtue of Parallel MS Acquisition. *Mol. Cell. Proteomics* **2006**, *5* (1), 144–156. https://doi.org/10.1074/mcp.M500230-MCP200.

(83) Engelhardt, H.; Jungheim, M. Comparison and Characterization of Reversed Phases. *Chromatographia* **1990**, *29* (1), 59–68. https://doi.org/10.1007/BF02261141.

(84) Herraiz, T.; Casal, V. Evaluation of Solid-Phase Extraction Procedures in Peptide Analysis. *J. Chromatogr. A* **1995**, *708* (2), 209–221. https://doi.org/10.1016/0021-9673(95)00388-4.

(85) Liuni, P.; Wilson, D. J. Understanding and Optimizing Electrospray Ionization Techniques for Proteomic Analysis. *Expert Rev. Proteomics* **2011**, *8* (2), 197–209. https://doi.org/10.1586/epr.10.111.

(86) Mittal, R. D. Tandem Mass Spectroscopy in Diagnosis and Clinical Research. *Indian J. Clin. Biochem.* **2015**, *30* (2), 121–123. https://doi.org/10.1007/s12291-015-0498-9.

(87) Nakashima, Y.; Nahar, S.; Miyagi-Shiohira, C.; Kinjo, T.; Toyoda, Z.; Kobayashi, N.; Saitoh, I.; Watanabe, M.; Fujita, J.; Noguchi, H. A Liquid Chromatography with Tandem Mass Spectrometry-Based Proteomic Analysis of the Proteins Secreted by Human Adipose-Derived Mesenchymal Stem Cells. *Cell Transplant.* **2018**, *27* (10), 1469–1494. https://doi.org/10.1177/0963689718795096.

(88) Fox, A. Mass Spectrometry for Species or Strain Identification after Culture or without Culture: Past, Present, and Future. *J. Clin. Microbiol.* **2006**, *44* (8), 2677–2680. https://doi.org/10.1128/jcm.00971-06.

(89) Mann, M.; Kelleher, N. L. Precision Proteomics: The Case for High Resolution and High Mass Accuracy. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18132–18138. https://doi.org/10.1073/pnas.0800788105.

(90) Meng, F.; Forbes, A. J.; Miller, L. M.; Kelleher, N. L. Detection and Localization of Protein Modifications by High Resolution Tandem Mass Spectrometry. *Mass Spectrom. Rev.* **2005**, *24* (2), 126–134. https://doi.org/10.1002/mas.20009.

(91) McMaster, M. C. *LC/MS: A Practical User's Guide*; John Wiley & Sons, 2005.

(92) Murayama, C.; Kimura, Y.; Setou, M. Imaging Mass Spectrometry: Principle and Application. *Biophys. Rev.* **2009**, *1* (3), 131. https://doi.org/10.1007/s12551-009-0015-6.

(93) Jocelyn Paré, J. R.; Yaylayan, V. Chapter 7 Mass Spectrometry: Principles and Applications. In *Techniques and Instrumentation in Analytical Chemistry*; Paré, J. R. J., Bélanger, J. M. R., Eds.; Instrumental Methods in Food Analysis; Elsevier, 1997; Vol. 18, pp 239–266. https://doi.org/10.1016/S0167-9244(97)80016-9.

(94) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372. https://doi.org/10.1038/nbt.1511.

(95) Nickerson, J. L.; Baghalabadi, V.; Rajendran, S. R. C. K.; Jakubec, P. J.; Said, H.; McMillen, T. S.; Dang, Z.; Doucette, A. A. Recent Advances in Top-down Proteome Sample Processing Ahead of MS Analysis. *Mass Spectrom. Rev.* **2023**, *42* (2), 457–495. https://doi.org/10.1002/mas.21706.

(96) Dupree, E. J.; Jayathirtha, M.; Yorkey, H.; Mihasan, M.; Petre, B. A.; Darie, C. C. A Critical Review of Bottom-Up Proteomics: The Good, the Bad, and the Future of This Field. *Proteomes* **2020**, *8* (3), 14. https://doi.org/10.3390/proteomes8030014.

(97) Emmott, E.; Jovanovic, M.; Slavov, N. Approaches for Studying Ribosome Specialization. *Trends Biochem. Sci.* **2019**, *44* (5), 478–479. https://doi.org/10.1016/j.tibs.2019.01.008.

(98) Ning, Z.; Zhang, X.; Mayne, J.; Figeys, D. Peptide-Centric Approaches Provide an Alternative Perspective To Re-Examine Quantitative Proteomic Data. *Anal. Chem.* **2016**, *88* (4), 1973–1978. https://doi.org/10.1021/acs.analchem.5b04148.

(99) Switzar, L.; Giera, M.; Niessen, W. M. A. Protein Digestion: An Overview of the Available Techniques and Recent Developments. *J. Proteome Res.* **2013**, *12* (3), 1067–1077. https://doi.org/10.1021/pr301201x.

(100) Snashall, C. M.; Sutton, C. W.; Faro, L. L.; Ceresa, C.; Ploeg, R.; Shaheed, S. ul. Comparison of In-Gel and in-Solution Proteolysis in the Proteome Profiling of Organ Perfusion Solutions. *Clin. Proteomics* **2023**, *20* (1), 51. https://doi.org/10.1186/s12014-023-09440-x.

(101) Shevchenko, A.; Tomas, H.; Havli, J.; Olsen, J. V.; Mann, M. In-Gel Digestion for Mass Spectrometric Characterization of Proteins and Proteomes. *Nat. Protoc.* **2006**, *1* (6), 2856–2860. https://doi.org/10.1038/nprot.2006.468.

(102) Medzihradszky, K. F. In-Solution Digestion of Proteins for Mass Spectrometry. *Methods Enzymol.* **2005**, *405*, 50–65. https://doi.org/10.1016/S0076-6879(05)05003-2.

(103) Tanford, C. Protein Denaturation. In *Advances in Protein Chemistry*; Anfinsen, C. B., Anson, M. L., Edsall, J. T., Richards, F. M., Eds.; Academic Press, 1968; Vol. 23, pp 121–282. https://doi.org/10.1016/S0065-3233(08)60401-5.

(104) Pratt, C.; Corneyl, K. *Pratt's Essential Biochemistry, Global Edition, 4th Edition | Wiley*; Wiley, 2018.

(105) Proc, J. L.; Kuzyk, M. A.; Hardie, D. B.; Yang, J.; Smith, D. S.; Jackson, A. M.; Parker, C. E.; Borchers, C. H. A Quantitative Study of the Effects of Chaotropic Agents, Surfactants, and Solvents on the Digestion Efficiency of Human Plasma Proteins by Trypsin. *J. Proteome Res.* **2010**, *9* (10), 5422–5437. https://doi.org/10.1021/pr100656u.

(106) Shuken, S. R. An Introduction to Mass Spectrometry-Based Proteomics. *J. Proteome Res.* **2023**, *22* (7), 2151–2171. https://doi.org/10.1021/acs.jproteome.2c00838.

(107) Park, Z.-Y.; Russell, D. H. Thermal Denaturation: A Useful Technique in Peptide Mass Mapping. *Anal. Chem.* **2000**, *72* (11), 2667–2670. https://doi.org/10.1021/ac991444k.

(108) Nelson, D. L.; Cox, M. *Lehninger Principles of Biochemistry*, 8th ed.; Macmillan Learning.

(109) Rebecchi, K. R.; Go, E. P.; Xu, L.; Woodin, C. L.; Mure, M.; Desaire, H. A General Protease Digestion Procedure for Optimal Protein Sequence Coverage and PTM Analysis of Recombinant Glycoproteins: Application to the Characterization of hLOXL2 Glycosylation. *Anal. Chem.* **2011**, *83* (22), 8484–8491. https://doi.org/10.1021/ac2017037.

(110) Mthembu, S. N.; Sharma, A.; Albericio, F.; de la Torre, B. G. Breaking a Couple: Disulfide Reducing Agents. *ChemBioChem* **2020**, *21* (14), 1947–1954. https://doi.org/10.1002/cbic.202000092.

(111) Cleland, W. W. Dithiothreitol, a New Protective Reagent for SH Groups. *Biochemistry* **1964**, *3* (4), 480–482. https://doi.org/10.1021/bi00892a002.

(112) Müller, T.; Winter, D. Systematic Evaluation of Protein Reduction and Alkylation Reveals Massive Unspecific Side Effects by Iodine-Containing Reagents. *Mol. Cell. Proteomics MCP* **2017**, *16* (7), 1173–1187. https://doi.org/10.1074/mcp.M116.064048.

(113) Tsiatsiani, L.; Heck, A. J. R. Proteomics beyond Trypsin. *FEBS J.* **2015**, *282* (14), 2612–2626. https://doi.org/10.1111/febs.13287.

(114) Vandermarliere, E.; Mueller, M.; Martens, L. Getting Intimate with Trypsin, the Leading Protease in Proteomics. *Mass Spectrom. Rev.* **2013**, *32* (6), 453–465. https://doi.org/10.1002/mas.21376.

(115) Olsen, J. V.; Ong, S.-E.; Mann, M. Trypsin Cleaves Exclusively C-Terminal to Arginine and Lysine Residues. *Mol. Cell. Proteomics* **2004**, *3* (6), 608–614. https://doi.org/10.1074/mcp.T400003-MCP200.

(116) Transue, T. R.; Krahn, J. M.; Gabel, S. A.; DeRose, E. F.; London, R. E. Porcine Trypsin with No Inhibitor Bound, 2004. https://doi.org/10.2210/pdb1s81/pdb.

(117) Di Cera, E. Serine Proteases. *IUBMB Life* **2009**, *61* (5), 510–515. https://doi.org/10.1002/iub.186.

(118) Kaur, J.; Singh, P. K. Trypsin Detection Strategies: A Review. *Crit. Rev. Anal. Chem.* **2022**, *52* (5), 949–967. https://doi.org/10.1080/10408347.2020.1846490.

(119) Pierce™ Microplate BCA Protein Assay Kit - Reducing Agent Compatible, 2011. https://www.thermofisher.com/order/catalog/product/23252 (accessed 2023-11-28).

(120) Pierce™ Quantitative Peptide Assays & Standards, 2014. https://www.thermofisher.com/order/catalog/product/23275 (accessed 2023-12-09).

(121) Kostenbauder, H. B.; Jawad, M. J.; Perrin, J. H.; Averhart, V. Dialysis and Circular Dichroism Study of the Binding of Sulfaethidole to Crystalline and Fraction V Bovine Serum Albumin. *J. Pharm. Sci.* **1971**, *60* (11), 1658–1660. https://doi.org/10.1002/jps.2600601114.

(122) Takeda, K.; Wada, A.; Yamamoto, K.; Moriyama, Y.; Aoki, K. Conformational Change of Bovine Serum Albumin by Heat Treatment. *J. Protein Chem.* **1989**, *8* (5), 653–659. https://doi.org/10.1007/BF01025605.

(123) Bourassa, P.; Kanakis, C. D.; Tarantilis, P.; Pollissiou, M. G.; Tajmir-Riahi, H. A. Resveratrol, Genistein, and Curcumin Bind Bovine Serum Albumin. *J. Phys. Chem. B* **2010**, *114* (9), 3348–3354. https://doi.org/10.1021/jp9115996.

(124) Takeda, K.; Shigeta, M.; Aoki, K. Secondary Structures of Bovine Serum Albumin in Anionic and Cationic Surfactant Solutions. *J. Colloid Interface Sci.* **1987**, *117* (1), 120–126. https://doi.org/10.1016/0021-9797(87)90174-3.

(125) Li, D.; Wang, Y.; Chen, J.; Ji, B. Characterization of the Interaction between Farrerol and Bovine Serum Albumin by Fluorescence and Circular Dichroism. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2011**, *79* (3), 680–686. https://doi.org/10.1016/j.saa.2011.04.005.

(126) Ottervald, J.; Franzén, B.; Nilsson, K.; Andersson, L. I.; Khademi, M.; Eriksson, B.; Kjellström, S.; Marko-Varga, G.; Végvári, Á.; Harris, R. A.; Laurell, T.; Miliotis, T.; Matusevicius, D.; Salter, H.; Ferm, M.; Olsson, T. Multiple Sclerosis: Identification and Clinical Evaluation of Novel CSF Biomarkers. *J. Proteomics* **2010**, *73* (6), 1117–1132. https://doi.org/10.1016/j.jprot.2010.01.004.

(127) Rao, P. V.; Reddy, A. P.; Lu, X.; Dasari, S.; Krishnaprasad, A.; Biggs, E.; Roberts, C. T. Jr.; Nagalla, S. R. Proteomic Identification of Salivary Biomarkers of Type-2 Diabetes. *J. Proteome Res.* **2009**, *8* (1), 239–245. https://doi.org/10.1021/pr8003776.

(128) Dutt, M.; Hartel, G.; Richards, R. S.; Shah, A. K.; Mohamed, A.; Apostolidou, S.; Gentry-Maharaj, A.; Group, A. O. C. S.; Hooper, J. D.; Perrin, L. C.; Menon, U.; Hill, M. M. Discovery and Validation of Serum Glycoprotein Biomarkers for High Grade Serous Ovarian Cancer. *Proteomics – Clin. Appl.* **2023**, *17* (4), 2200114. https://doi.org/10.1002/prca.202200114.

(129) Facchiano, F.; D'Arcangelo, D.; Facchiano, A. Alpha-2-Macroglobulin Is a Novel Anticancer Agent. *Oncology* **2023**, 1–5. https://doi.org/10.1159/000536033.

(130) Deutsch, E. W.; Eng, J. K.; Zhang, H.; King, N. L.; Nesvizhskii, A. I.; Lin, B.; Lee, H.; Yi, E. C.; Ossola, R.; Aebersold, R. Human Plasma PeptideAtlas. *Proteomics* **2005**, *5* (13), 3497–3500. https://doi.org/10.1002/pmic.200500160.

(131) Rios, A. G.; Ribeiro, A. M.; Rodrigues, A. E.; Ferreira, A. F. P. Bovine Serum Albumin and Myoglobin Separation by Ion-Exchange SMB. *Ind. Eng. Chem. Res.* **2023**, *62* (35), 13932–13942. https://doi.org/10.1021/acs.iecr.3c01613.

(132) The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51* (D1), D523–D531. https://doi.org/10.1093/nar/gkac1052.

(133) Albumin, Bovine, 1997. https://www.sigmaaldrich.com/TR/en/product/sigma/a4919 (accessed 2024-11-02).

(134) Li, R.; Wu, Z.; Wangb, Y.; Ding, L.; Wang, Y. Role of pH-Induced Structural Change in Protein Aggregation in Foam Fractionation of Bovine Serum Albumin. *Biotechnol. Rep.* **2016**, *9*, 46–52. https://doi.org/10.1016/j.btre.2016.01.002.

(135) Doumas, B. T. Standards for Total Serum Protein Assays--a Collaborative Study. *Clin. Chem.* **1975**, *21* (8), 1159–1166.

(136) Hsieh, S.-R.; Reddy, P. M.; Chang, C.-J.; Kumar, A.; Wu, W.-C.; Lin, H.-Y. Exploring the Behavior of Bovine Serum Albumin in Response to Changes in the Chemical Composition of Responsive Polymers: Experimental and Simulation Studies. *Polymers* **2016**, *8* (6), 238. https://doi.org/10.3390/polym8060238.

(137) Roy, S.; Nandi, R.; Ganai, S.; Majumdar, K.; Das, T. Binding Interaction of Phosphorus Heterocycles with Bovine Serum Albumin: A Biochemical Study. *J. Pharm. Anal.* **2016**, *7*, 19–26. https://doi.org/10.1016/j.jpha.2016.05.009.

(138) Mitaki, S.; Wada, Y.; Sheikh, A. Md.; Yamaguchi, S.; Nagai, A. Proteomic Analysis of Extracellular Vesicles Enriched Serum Associated with Future Ischemic Stroke. *Sci. Rep.* **2021**, *11*, 24024. https://doi.org/10.1038/s41598-021-03497-0.

(139) Cater, J. H.; Wilson, M. R.; Wyatt, A. R. Alpha-2-Macroglobulin, a Hypochlorite-Regulated Chaperone and Immune System Modulator. *Oxid.*

Med. Cell. Longev. **2019**, *2019*, 5410657. https://doi.org/10.1155/2019/5410657.

(140) Kabasakal, S. O. Determination of Alpha-2-Macroglubulin in Serum Samples. *J. Fac. Pharm. Ank. Univ.* **2022**, *46* (3), 967–978. https://doi.org/10.33483/jfpau.1139157.

(141) Birkenmeier, G.; Müller, R.; Huse, K.; Forberg, J.; Gläser, C.; Hedrich, H.; Nicklisch, S.; Reichenbach, A. Human A2-Macroglobulin: Genotype–Phenotype Relation. *Exp. Neurol.* **2003**, *184* (1), 153–161. https://doi.org/10.1016/S0014-4886(03)00110-9.

(142) Thieme, R.; Kurz, S.; Kolb, M.; Debebe, T.; Holtze, S.; Morhart, M.; Huse, K.; Szafranski, K.; Platzer, M.; Hildebrandt, T. B.; Birkenmeier, G. Analysis of Alpha-2 Macroglobulin from the Long-Lived and Cancer-Resistant Naked Mole-Rat and Human Plasma. *PloS One* **2015**, *10* (6), e0130470. https://doi.org/10.1371/journal.pone.0130470.

(143) Varma, V. R.; Varma, S.; An, Y.; Hohman, T. J.; Seddighi, S.; Casanova, R.; Beri, A.; Dammer, E. B.; Seyfried, N. T.; Pletnikova, O.; Moghekar, A.; Wilson, M. R.; Lah, J. J.; O'Brien, R. J.; Levey, A. I.; Troncoso, J. C.; Albert, M. S.; Thambisetty, M. Alpha-2 Macroglobulin in Alzheimer's Disease: A Marker of Neuronal Injury through the RCAN1 Pathway. *Mol. Psychiatry* **2017**, *22* (1), 13–23. https://doi.org/10.1038/mp.2016.206.

(144) Selvaraju, V.; Babu, J. R.; Geetha, T. Salivary C-Reactive Protein as a Biomarker and Implications for Diabetes. In *Biomarkers in Diabetes*; Patel, V. B., Preedy, V. R., Eds.; Biomarkers in Disease: Methods, Discoveries and Applications; Springer International Publishing: Cham, 2023; pp 329–347. https://doi.org/10.1007/978-3-031-08014-2_17.

(145) Sigma-Aldrich. *Bovine Serum Albumin heat shock fraction, protease free, low endotoxin, cell culture mammalian, 7, = 98 9048-46-8*. http://www.sigmaaldrich.com/ (accessed 2023-11-28).

(146) Han, S. Y. S.; Cooper, J. D.; Ozcan, S.; Rustogi, N.; Penninx, B. W. J. H.; Bahn, S. Integrating Proteomic, Sociodemographic and Clinical Data to Predict Future Depression Diagnosis in Subthreshold Symptomatic Individuals. *Transl. Psychiatry* **2019**, *9* (277). https://doi.org/10.1038/s41398-019-0623-2.

(147) Cooper, J. D.; Ozcan, S.; Gardner, R. M.; Rustogi, N.; Wicks, S.; van Rees, G. F.; Leweke, F. M.; Dalman, C.; Karlsson, H.; Bahn, S. Schizophrenia-Risk and Urban Birth Are Associated with Proteomic Changes in Neonatal Dried Blood Spots. *Transl. Psychiatry* **2017**, *7* (12), 1290. https://doi.org/10.1038/s41398-017-0027-0.

(148) Cooper, J. D.; Han, S. Y. S.; Tomasik, J.; Ozcan, S.; Rustogi, N.; van Beveren, N. J. M.; Leweke, F. M.; Bahn, S. Multimodel Inference for Biomarker Development: An Application to Schizophrenia. *Transl. Psychiatry* **2019**, *9* (1), 1–10. https://doi.org/10.1038/s41398-019-0419-4.

(149) Ozcan, S.; Cooper, J. D.; Lago, S. G.; Kenny, D.; Rustogi, N.; Stocki, P.; Bahn, S. Towards Reproducible MRM Based Biomarker Discovery Using Dried Blood Spots. *Sci. Rep.* **2017**, *7* (1), 45178. https://doi.org/10.1038/srep45178.

(150) Addona, T. A.; Abbatiello, S. E.; Schilling, B.; Skates, S. J.; Mani, D. R.; Bunk, D. M.; Spiegelman, C. H.; Zimmerman, L. J.; Ham, A.-J. L.; Keshishian, H.; Hall, S. C.; Allen, S.; Blackman, R. K.; Borchers, C. H.; Buck, C.; Cardasis, H. L.; Cusack, M. P.; Dodder, N. G.; Gibson, B. W.; Held, J. M.; Hiltke, T.; Jackson, A.; Johansen, E. B.; Kinsinger, C. R.; Li, J.; Mesri, M.; Neubert, T. A.; Niles, R. K.; Pulsipher, T. C.; Ransohoff, D.; Rodriguez, H.; Rudnick, P. A.; Smith, D.; Tabb, D. L.; Tegeler, T. J.; Variyath, A. M.; Vega-Montoto, L. J.; Wahlander, Å.; Waldemarson, S.; Wang, M.; Whiteaker, J. R.; Zhao, L.; Anderson, N. L.; Fisher, S. J.; Liebler, D. C.; Paulovich, A. G.; Regnier, F. E.; Tempst, P.; Carr, S. A. Multi-Site Assessment of the Precision and Reproducibility of Multiple Reaction Monitoring–Based Measurements

of Proteins in Plasma. *Nat. Biotechnol.* **2009**, *27* (7), 633–641. https://doi.org/10.1038/nbt.1546.

(151) Geyer, P. E.; Arend, F. M.; Doll, S.; Louiset, M.-L.; Virreira Winter, S.; Müller-Reif, J. B.; Torun, F. M.; Weigand, M.; Eichhorn, P.; Bruegel, M.; Strauss, M. T.; Holdt, L. M.; Mann, M.; Teupser, D. High-Resolution Serum Proteome Trajectories in COVID-19 Reveal Patient-Specific Seroconversion. *EMBO Mol. Med.* **2021**, *13* (8), e14167. https://doi.org/10.15252/emmm.202114167.

(152) Geyer, P. E.; Kulak, N. A.; Pichler, G.; Holdt, L. M.; Teupser, D.; Mann, M. Plasma Proteome Profiling to Assess Human Health and Disease. *Cell Syst.* **2016**, *2* (3), 185–195. https://doi.org/10.1016/j.cels.2016.02.015.

(153) Geyer, P. E.; Voytik, E.; Treit, P. V.; Doll, S.; Kleinhempel, A.; Niu, L.; Müller, J. B.; Buchholtz, M.-L.; Bader, J. M.; Teupser, D.; Holdt, L. M.; Mann, M. Plasma Proteome Profiling to Detect and Avoid Sample-Related Biases in Biomarker Studies. *EMBO Mol. Med.* **2019**, *11* (11), e10427. https://doi.org/10.15252/emmm.201910427.

(154) Geyer, P. E.; Wewer Albrechtsen, N. J.; Tyanova, S.; Grassl, N.; Iepsen, E. W.; Lundgren, J.; Madsbad, S.; Holst, J. J.; Torekov, S. S.; Mann, M. Proteomics Reveals the Effects of Sustained Weight Loss on the Human Plasma Proteome. *Mol. Syst. Biol.* **2016**, *12* (12), 901. https://doi.org/10.15252/msb.20167357.

(155) Wewer Albrechtsen, N. J.; Geyer, P. E.; Doll, S.; Treit, P. V.; Bojsen-Møller, K. N.; Martinussen, C.; Jørgensen, N. B.; Torekov, S. S.; Meier, F.; Niu, L.; Santos, A.; Keilhauer, E. C.; Holst, J. J.; Madsbad, S.; Mann, M. Plasma Proteome Profiling Reveals Dynamics of Inflammatory and Lipid Homeostasis Markers after Roux-En-Y Gastric Bypass Surgery. *Cell Syst.* **2018**, *7* (6), 601-612.e3. https://doi.org/10.1016/j.cels.2018.10.012.

(156) Kuzyk, M. A.; Smith, D.; Yang, J.; Cross, T. J.; Jackson, A. M.; Hardie, D. B.; Anderson, N. L.; Borchers, C. H. Multiple Reaction Monitoring-Based,

Multiplexed, Absolute Quantitation of 45 Proteins in Human Plasma*. *Mol. Cell. Proteomics* **2009**, *8* (8), 1860–1877. https://doi.org/10.1074/mcp.M800540-MCP200.

(157) Demichev, V.; Tober-Lau, P.; Lemke, O.; Nazarenko, T.; Thibeault, C.; Whitwell, H.; Röhl, A.; Freiwald, A.; Szyrwiel, L.; Ludwig, D.; Correia-Melo, C.; Aulakh, S. K.; Helbig, E. T.; Stubbemann, P.; Lippert, L. J.; Grüning, N.-M.; Blyuss, O.; Vernardis, S.; White, M.; Messner, C. B.; Joannidis, M.; Sonnweber, T.; Klein, S. J.; Pizzini, A.; Wohlfarter, Y.; Sahanic, S.; Hilbe, R.; Schaefer, B.; Wagner, S.; Mittermaier, M.; Machleidt, F.; Garcia, C.; Ruwwe-Glösenkamp, C.; Lingscheid, T.; Bosquillon de Jarcy, L.; Stegemann, M. S.; Pfeiffer, M.; Jürgens, L.; Denker, S.; Zickler, D.; Enghard, P.; Zelezniak, A.; Campbell, A.; Hayward, C.; Porteous, D. J.; Marioni, R. E.; Uhrig, A.; Müller-Redetzky, H.; Zoller, H.; Löffler-Ragg, J.; Keller, M. A.; Tancevski, I.; Timms, J. F.; Zaikin, A.; Hippenstiel, S.; Ramharter, M.; Witzenrath, M.; Suttorp, N.; Lilley, K.; Mülleder, M.; Sander, L. E.; Kleinschmidt, M.; Heim, K. M.; Millet, B.; Meyer-Arndt, L.; Hübner, R. H.; Andermann, T.; Doehn, J. M.; Opitz, B.; Sawitzki, B.; Grund, D.; Radünzel, P.; Schürmann, M.; Zoller, T.; Alius, F.; Knape, P.; Breitbart, A.; Li, Y.; Bremer, F.; Pergantis, P.; Schürmann, D.; Temmesfeld-Wollbrück, B.; Wendisch, D.; Brumhard, S.; Haenel, S. S.; Conrad, C.; Georg, P.; Eckardt, K.-U.; Lehner, L.; Kruse, J. M.; Ferse, C.; Körner, R.; Spies, C.; Edel, A.; Weber-Carstens, S.; Krannich, A.; Zvorc, S.; Li, L.; Behrens, U.; Schmidt, S.; Rönnefarth, M.; Dang-Heine, C.; Röhle, R.; Lieker, E.; Kretzler, L.; Wirsching, I.; Wollboldt, C.; Wu, Y.; Schwanitz, G.; Hillus, D.; Kasper, S.; Olk, N.; Horn, A.; Briesemeister, D.; Treue, D.; Hummel, M.; Corman, V. M.; Drosten, C.; von Kalle, C.; Ralser, M.; Kurth, F. A Time-Resolved Proteomic and Prognostic Map of COVID-19. *Cell Syst.* **2021**, *12* (8), 780-794.e7. https://doi.org/10.1016/j.cels.2021.05.005.

(158) Fortin, T.; Salvador, A.; Charrier, J. P.; Lenz, C.; Bettsworth, F.; Lacoux, X.; Choquet-Kastylevsky, G.; Lemoine, J. Multiple Reaction Monitoring Cubed for Protein Quantification at the Low Nanogram/Milliliter Level in Nondepleted Human Serum. *Anal. Chem.* **2009**, *81* (22), 9343–9352. https://doi.org/10.1021/ac901447h.

(159) Oberbach, A.; Schlichting, N.; Neuhaus, J.; Kullnick, Y.; Lehmann, S.; Heinrich, M.; Dietrich, A.; Mohr, F. W.; von Bergen, M.; Baumann, S. Establishing a Reliable Multiple Reaction Monitoring-Based Method for the Quantification of Obesity-Associated Comorbidities in Serum and Adipose Tissue Requires Intensive Clinical Validation. *J. Proteome Res.* **2014**, *13* (12), 5784–5800. https://doi.org/10.1021/pr500722k.

(160) Hober, A.; Edfors, F.; Ryaboshapkina, M.; Malmqvist, J.; Rosengren, L.; Percy, A. J.; Lind, L.; Forsström, B.; Uhlén, M.; Oscarsson, J.; Miliotis, T. Absolute Quantification of Apolipoproteins Following Treatment with Omega-3 Carboxylic Acids and Fenofibrate Using a High Precision Stable Isotope-Labeled Recombinant Protein Fragments Based SRM Assay. *Mol. Cell. Proteomics MCP* **2019**, *18* (12), 2433–2446. https://doi.org/10.1074/mcp.RA119.001765.

(161) Gaither, C.; Popp, R.; Mohammed, Y.; H. Borchers, C. Determination of the Concentration Range for 267 Proteins from 21 Lots of Commercial Human Plasma Using Highly Multiplexed Multiple Reaction Monitoring Mass Spectrometry. *Analyst* **2020**, *145* (10), 3634–3644. https://doi.org/10.1039/C9AN01893J.

(162) Wesseling, H.; Mair, W.; Kumar, M.; Schlaffner, C. N.; Tang, S.; Beerepoot, P.; Fatou, B.; Guise, A. J.; Cheng, L.; Takeda, S.; Muntel, J.; Rotunno, M. S.; Dujardin, S.; Davies, P.; Kosik, K. S.; Miller, B. L.; Berretta, S.; Hedreen, J. C.; Grinberg, L. T.; Seeley, W. W.; Hyman, B. T.; Steen, H.; Steen, J. A. Tau PTM Profiles Identify Patient Heterogeneity and Stages of Alzheimer's

Disease. *Cell* **2020**, *183* (6), 1699-1713.e13. https://doi.org/10.1016/j.cell.2020.10.029.

(163) Wiśniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal Sample Preparation Method for Proteome Analysis. *Nat. Methods* **2009**, *6* (5), 359–362. https://doi.org/10.1038/nmeth.1322.

(164) Frin, K.; Nascimento, V. Rhenium(I) Polypyridine Complexes as Luminescence-Based Sensors for the BSA Protein. *J. Braz. Chem. Soc.* **2015**, *27*. https://doi.org/10.5935/0103-5053.20150268.

(165) Stewart, I. I.; Thomson, T.; Figeys, D. 18O Labeling: A Tool for Proteomics. *Rapid Commun. Mass Spectrom.* **2001**, *15* (24), 2456–2465. https://doi.org/10.1002/rcm.525.

(166) Prasopdee, S.; Yingchutrakul, Y.; Krobthong, S.; Pholhelm, M.; Wongtrakoongate, P.; Butthongkomvong, K.; Kulsantiwong, J.; Phanaksri, T.; Kunjantarachot, A.; Sathavornmanee, T.; Tesana, S.; Thitapakorn, V. Differential Plasma Proteomes of the Patients with Opisthorchiasis Viverrini and Cholangiocarcinoma Identify a Polymeric Immunoglobulin Receptor as a Potential Biomarker. *Heliyon* **2022**, *8* (10), e10965. https://doi.org/10.1016/j.heliyon.2022.e10965.

(167) Qian, W.-J.; Jacobs, J. M.; Camp II, D. G.; Monroe, M. E.; Moore, R. J.; Gritsenko, M. A.; Calvano, S. E.; Lowry, S. F.; Xiao, W.; Moldawer, L. L.; Davis, R. W.; Tompkins, R. G.; Smith, R. D. Comparative proteome analyses of human plasma following in vivo lipopolysaccharide administration using multidimensional separations coupled with tandem mass spectrometry. *Proteomics* **2005**, *5* (2), 572–584. https://doi.org/10.1002/pmic.200400942.

(168) Karuso, P.; Crawford, A. S.; Veal, D. A.; Scott, G. B. I.; Choi, H.-Y. Real-Time Fluorescence Monitoring of Tryptic Digestion in Proteomics. *J. Proteome Res.* **2008**, *7* (1), 361–366. https://doi.org/10.1021/pr0704480.

(169) Dubois, C.; Payen, D.; Simon, S.; Junot, C.; Fenaille, F.; Morel, N.; Becher, F. Top-Down and Bottom-Up Proteomics of Circulating S100A8/S100A9 in

Plasma of Septic Shock Patients. *J. Proteome Res.* **2020**, *19* (2), 914–925. https://doi.org/10.1021/acs.jproteome.9b00690.

(170) Bennike, T. B.; Bellin, M. D.; Xuan, Y.; Stensballe, A.; Møller, F. T.; Beilman, G. J.; Levy, O.; Cruz-Monserrate, Z.; Andersen, V.; Steen, J.; Conwell, D. L.; Steen, H. A Cost-Effective High-Throughput Plasma and Serum Proteomics Workflow Enables Mapping of the Molecular Impact of Total Pancreatectomy with Islet Autotransplantation. *J. Proteome Res.* **2018**, *17* (5), 1983–1992. https://doi.org/10.1021/acs.jproteome.8b00111.

(171) Xiao, Z.; Conrads, T. P.; Lucas, D. A.; Janini, G. M.; Schaefer, C. F.; Buetow, K. H.; Issaq, H. J.; Veenstra, T. D. Direct Ampholyte-Free Liquid-Phase Isoelectric Peptide Focusing: Application to the Human Serum Proteome. *Electrophoresis* **2004**, *25* (1), 128–133. https://doi.org/10.1002/elps.200305700.

(172) Chan, K. C.; Lucas, D. A.; Hise, D.; Schaefer, C. F.; Xiao, Z.; Janini, G. M.; Buetow, K. H.; Issaq, H. J.; Veenstra, T. D.; Conrads, T. P. Analysis of the Human Serum Proteome. *Clin. Proteomics* **2004**, *1* (2), 101–225. https://doi.org/10.1385/CP:1:2:101.

(173) Molloy, M. P.; Hill, C.; O'Rourke, M. B.; Chandra, J.; Steffen, P.; McKay, M. J.; Pascovici, D.; Herbert, B. R. Proteomic Analysis of Whole Blood Using Volumetric Absorptive Microsampling for Precision Medicine Biomarker Studies. *J. Proteome Res.* **2022**, *21* (4), 1196–1203. https://doi.org/10.1021/acs.jproteome.1c00971.

(174) Rosting, C.; Yu, J.; Cooper, H. J. High Field Asymmetric Waveform Ion Mobility Spectrometry in Nontargeted Bottom-up Proteomics of Dried Blood Spots. *J. Proteome Res.* **2018**, *17* (6), 1997–2004. https://doi.org/10.1021/acs.jproteome.7b00746.

(175) Niu, L.; Thiele, M.; Geyer, P. E.; Rasmussen, D. N.; Webel, H. E.; Santos, A.; Gupta, R.; Meier, F.; Strauss, M.; Kjaergaard, M.; Lindvig, K.; Jacobsen, S.; Rasmussen, S.; Hansen, T.; Krag, A.; Mann, M. Noninvasive Proteomic

Biomarkers for Alcohol-Related Liver Disease. *Nat. Med.* **2022**, *28* (6), 1277–1287. https://doi.org/10.1038/s41591-022-01850-y.

(176) R. Abou-Shakra, F.; P. Rayman, M.; Ward, N. I.; Hotton, V.; Bastian, G. Enzymatic Digestion for the Determination of Trace Elements in Blood Serum by Inductively Coupled Plasma Mass Spectrometry. *J. Anal. At. Spectrom.* **1997**, *12* (4), 429–433. https://doi.org/10.1039/A607972E.

(177) Lu, Y.; Kippler, M.; Harari, F.; Grandér, M.; Palm, B.; Nordqvist, H.; Vahter, M. Alkali Dilution of Blood Samples for High Throughput ICP-MS Analysis—Comparison with Acid Digestion. *Clin. Biochem.* **2015**, *48* (3), 140–147. https://doi.org/10.1016/j.clinbiochem.2014.12.003.

(178) Harrington, J. M.; Young, D. J.; Essader, A. S.; Sumner, S. J.; Levine, K. E. Analysis of Human Serum and Whole Blood for Mineral Content by ICP-MS and ICP-OES: Development of a Mineralomics Method. *Biol. Trace Elem. Res.* **2014**, *160* (1), 132–142.

(179) Integrated proteomic sample preparation with combination of on-line high-abundance protein depletion, denaturation, reduction, desalting and digestion to achieve high throughput plasma proteome quantification | *Elsevier Enhanced Reader*. https://doi.org/10.1016/j.aca.2021.338343.

(180) Amrani, M. E.; Gerencser, L.; Huitema, A. D. R.; Hack, C. E.; van Luin, M.; van der Elst, K. C. M. A Generic Sample Preparation Method for the Multiplex Analysis of Seven Therapeutic Monoclonal Antibodies in Human Plasma or Serum with Liquid Chromatography-Tandem Mass Spectrometry. *J. Chromatogr. A* **2021**, *1655*, 462489. https://doi.org/10.1016/j.chroma.2021.462489.

(181) Grassin-Delyle, S.; Martin, M.; Hamzaoui, O.; Lamy, E.; Jayle, C.; Sage, E.; Etting, I.; Devillier, P.; Alvarez, J.-C. A High-Resolution ICP-MS Method for the Determination of 38 Inorganic Elements in Human Whole Blood, Urine, Hair and Tissues after Microwave Digestion. *Talanta* **2019**, *199*, 228–237. https://doi.org/10.1016/j.talanta.2019.02.068.

(182) Carr, S. A.; Abbatiello, S. E.; Ackermann, B. L.; Borchers, C.; Domon, B.; Deutsch, E. W.; Grant, R. P.; Hoofnagle, A. N.; Hüttenhain, R.; Koomen, J. M.; Liebler, D. C.; Liu, T.; MacLean, B.; Mani, D.; Mansfield, E.; Neubert, H.; Paulovich, A. G.; Reiter, L.; Vitek, O.; Aebersold, R.; Anderson, L.; Bethem, R.; Blonder, J.; Boja, E.; Botelho, J.; Boyne, M.; Bradshaw, R. A.; Burlingame, A. L.; Chan, D.; Keshishian, H.; Kuhn, E.; Kinsinger, C.; Lee, J. S. H.; Lee, S.-W.; Moritz, R.; Oses-Prieto, J.; Rifai, N.; Ritchie, J.; Rodriguez, H.; Srinivas, P. R.; Townsend, R. R.; Van Eyk, J.; Whiteley, G.; Wiita, A.; Weintraub, S. Targeted Peptide Measurements in Biology and Medicine: Best Practices for Mass Spectrometry-Based Assay Development Using a Fit-for-Purpose Approach*. *Mol. Cell. Proteomics* **2014**, *13* (3), 907–917. https://doi.org/10.1074/mcp.M113.036095.

(183) UVic, G. Human PlasmaAssay Validation on Agilent 6490., 2016. https://assays.cancer.gov/non-CPTAC-1068.

(184) Li, X.; Hayward, C.; Fong, P.-Y.; Dominguez, M.; Hunsucker, S. W.; Lee, L. W.; McLean, M.; Law, S.; Butler, H.; Schirm, M.; Gingras, O.; Lamontagne, J.; Allard, R.; Chelsky, D.; Price, N. D.; Lam, S.; Massion, P. P.; Pass, H.; Rom, W. N.; Vachani, A.; Fang, K. C.; Hood, L.; Kearney, P. A Blood-Based Proteomic Classifier for the Molecular Characterization of Pulmonary Nodules. *Sci. Transl. Med.* **2013**, *5* (207), 207ra142-207ra142. https://doi.org/10.1126/scitranslmed.3007013.

(185) Zhou, Y.; Qin, S.; Sun, M.; Tang, L.; Yan, X.; Kim, T.-K.; Caballero, J.; Glusman, G.; Brunkow, M. E.; Soloski, M. J.; Rebman, A. W.; Scavarda, C.; Cooper, D.; Omenn, G. S.; Moritz, R. L.; Wormser, G. P.; Price, N. D.; Aucott, J. N.; Hood, L. Measurement of Organ-Specific and Acute-Phase Blood Protein Levels in Early Lyme Disease. *J. Proteome Res.* **2020**, *19* (1), 346–359. https://doi.org/10.1021/acs.jproteome.9b00569.

(186) Klont, F.; Pouwels, S. D.; Hermans, J.; van de Merbel, N. C.; Horvatovich, P.; ten Hacken, N. H. T.; Bischoff, R. A Fully Validated Liquid Chromatography-

Mass Spectrometry Method for the Quantification of the Soluble Receptor of Advanced Glycation End-Products (sRAGE) in Serum Using Immunopurification in a 96-Well Plate Format. *Talanta* **2018**, *182*, 414–421. https://doi.org/10.1016/j.talanta.2018.02.015.

(187) De Marchi, T.; Kuhn, E.; Dekker, L. J.; Stingl, C.; Braakman, R. B. H.; Opdam, M.; Linn, S. C.; Sweep, F. C. G. J.; Span, P. N.; Luider, T. M.; Foekens, J. A.; Martens, J. W. M.; Carr, S. A.; Umar, A. Targeted MS Assay Predicting Tamoxifen Resistance in Estrogen-Receptor-Positive Breast Cancer Tissues and Sera. *J. Proteome Res.* **2016**, *15* (4), 1230–1242. https://doi.org/10.1021/acs.jproteome.5b01119.

(188) Kuzyk, M. A.; Smith, D.; Yang, J.; Cross, T. J.; Jackson, A. M.; Hardie, D. B.; Anderson, N. L.; Borchers, C. H. Multiple Reaction Monitoring-Based, Multiplexed, Absolute Quantitation of 45 Proteins in Human Plasma. *Mol. Cell. Proteomics MCP* **2009**, *8* (8), 1860–1877. https://doi.org/10.1074/mcp.M800540-MCP200.

(189) Fernández, A. S.; Rodríguez-González, P.; Álvarez, L.; García, M.; Iglesias, H. G.; García Alonso, J. I. Multiple Heart-Cutting Two Dimensional Liquid Chromatography and Isotope Dilution Tandem Mass Spectrometry for the Absolute Quantification of Proteins in Human Serum. *Anal. Chim. Acta* **2021**, *1184*, 339022. https://doi.org/10.1016/j.aca.2021.339022.

(190) van Rees, G. F.; Lago, S. G.; Cox, D. A.; Tomasik, J.; Rustogi, N.; Weigelt, K.; Ozcan, S.; Cooper, J.; Drexhage, H.; Leweke, F. M.; Bahn, S. Evidence of Microglial Activation Following Exposure to Serum from First-Onset Drug-Naïve Schizophrenia Patients. *Brain. Behav. Immun.* **2018**, *67*, 364–373. https://doi.org/10.1016/j.bbi.2017.10.003.

(191) Qian, W.-J.; Liu, T.; Petyuk, V. A.; Gritsenko, M. A.; Petritis, B. O.; Polpitiya, A. D.; Kaushal, A.; Xiao, W.; Finnerty, C. C.; Jeschke, M. G.; Jaitly, N.; Monroe, M. E.; Moore, R. J.; Moldawer, L. L.; Davis, R. W.; Tompkins, R. G.; Herndon, D. N.; Camp, D. G. I.; Smith, R. D. Large-Scale Multiplexed

Quantitative Discovery Proteomics Enabled by the Use of an 18O-Labeled "Universal" Reference Sample. *J. Proteome Res.* **2009**, *8* (1), 290–299. https://doi.org/10.1021/pr800467r.

(192) Domanski, D.; Percy, A. J.; Yang, J.; Chambers, A. G.; Hill, J. S.; Freue, G. V. C.; Borchers, C. H. MRM-Based Multiplexed Quantitation of 67 Putative Cardiovascular Disease Biomarkers in Human Plasma. *Proteomics* **2012**, *12* (8), 1222–1243. https://doi.org/10.1002/pmic.201100568.

(193) Kuzyk, M. A.; Smith, D.; Yang, J.; Cross, T. J.; Jackson, A. M.; Hardie, D. B.; Anderson, N. L.; Borchers, C. H. Multiple Reaction Monitoring-Based, Multiplexed, Absolute Quantitation of 45 Proteins in Human Plasma *. *Mol. Cell. Proteomics* **2009**, *8* (8), 1860–1877. https://doi.org/10.1074/mcp.M800540-MCP200.

(194) Sleddering, M. A.; Markvoort, A. J.; Dharuri, H. K.; Jeyakar, S.; Snel, M.; Juhasz, P.; Lynch, M.; Hines, W.; Li, X.; Jazet, I. M.; Adourian, A.; Hilbers, P. A. J.; Smit, J. W. A.; Dijk, K. W. V. Proteomic Analysis in Type 2 Diabetes Patients before and after a Very Low Calorie Diet Reveals Potential Disease State and Intervention Specific Biomarkers. *PLOS ONE* **2014**, *9* (11), e112835. https://doi.org/10.1371/journal.pone.0112835.

(195) Anwar, M. A.; Dai, D. L.; Wilson-McManus, J.; Smith, D.; Francis, G. A.; Borchers, C. H.; McManus, B. M.; Hill, J. S.; Cohen Freue, G. V. Multiplexed LC–ESI–MRM-MS-Based Assay for Identification of Coronary Artery Disease Biomarkers in Human Plasma. *Proteomics – Clin. Appl.* **2019**, *13* (4), 1700111. https://doi.org/10.1002/prca.201700111.

(196) Dupin, M.; Fortin, T.; Larue-Triolet, A.; Surault, I.; Beaulieu, C.; Gouel-Chéron, A.; Allaouchiche, B.; Asehnoune, K.; Roquilly, A.; Venet, F.; Monneret, G.; Lacoux, X.; Roitsch, C. A.; Pachot, A.; Charrier, J.-P.; Pons, S. Impact of Serum and Plasma Matrices on the Titration of Human Inflammatory Biomarkers Using Analytically Validated SRM Assays. *J.*

Proteome Res. **2016**, *15* (8), 2366–2378. https://doi.org/10.1021/acs.jproteome.5b00803.

(197) DelGuidice, C. E.; Ismaiel, O. A.; Mylott, W. R.; Halquist, M. S. Optimization and Method Validation for the Quantitative Analysis of a Monoclonal Antibody and Its Related Fab Fragment in Human Plasma after Intravitreal Administration, Using LC–MS/MS. *J. Chromatogr. B* **2021**, *1164*, 122474. https://doi.org/10.1016/j.jchromb.2020.122474.

(198) Lebert, D.; Louwagie, M.; Goetze, S.; Picard, G.; Ossola, R.; Duquesne, C.; Basler, K.; Ferro, M.; Rinner, O.; Aebersold, R.; Garin, J.; Mouz, N.; Brunner, E.; Brun, V. DIGESTIF: A Universal Quality Standard for the Control of Bottom-Up Proteomics Experiments. *J. Proteome Res.* **2015**, *14* (2), 787–803. https://doi.org/10.1021/pr500834z.

(199) Schmidt, A.; Farine, H.; Keller, M. P.; Sebastian, A.; Kozera, L.; Welford, R. W. D.; Strasser, D. S. Immunoaffinity Targeted Mass Spectrometry Analysis of Human Plasma Samples Reveals an Imbalance of Active and Inactive CXCL10 in Primary Sjögren's Syndrome Disease Patients. *J. Proteome Res.* **2020**, *19* (10), 4196–4209. https://doi.org/10.1021/acs.jproteome.0c00494.

(200) Yu, K. H.; Barry, C. G.; Austin, D.; Busch, C. M.; Sangar, V.; Rustgi, A. K.; Blair, I. A. Stable Isotope Dilution Multidimensional Liquid Chromatography-Tandem Mass Spectrometry for Pancreatic Cancer Serum Biomarker Discovery. *J. Proteome Res.* **2009**, *8* (3), 1565–1576. https://doi.org/10.1021/pr800904z.

(201) Alayi, T. D.; Tawalbeh, S. M.; Ogundele, M.; Smith, H. R.; Samsel, A. M.; Barbieri, M. L.; Hathout, Y. Tandem Mass Tag-Based Serum Proteome Profiling for Biomarker Discovery in Young Duchenne Muscular Dystrophy Boys. *ACS Omega* **2020**, *5* (41), 26504–26517. https://doi.org/10.1021/acsomega.0c03206.

(202) Kim, K.; Kim, S. J.; Yu, H. G.; Yu, J.; Park, K. S.; Jang, I.-J.; Kim, Y. Verification of Biomarkers for Diabetic Retinopathy by Multiple Reaction

Monitoring. *J. Proteome Res.* **2010**, *9* (2), 689–699. https://doi.org/10.1021/pr901013d.

(203) Han, S. Y. S.; Tomasik, J.; Rustogi, N.; Lago, S. G.; Barton-Owen, G.; Eljasz, P.; Cooper, J. D.; Ozcan, S.; Olmert, T.; Farrag, L. P.; Friend, L. V.; Bell, E.; Cowell, D.; Thomas, G.; Tuytten, R.; Bahn, S. Diagnostic Prediction Model Development Using Data from Dried Blood Spot Proteomics and a Digital Mental Health Assessment to Identify Major Depressive Disorder among Individuals Presenting with Low Mood. *Brain. Behav. Immun.* **2020**, *90*, 184–195. https://doi.org/10.1016/j.bbi.2020.08.011.

(204) Breen, M. S.; Uhlmann, A.; Ozcan, S.; Chan, M.; Pinto, D.; Bahn, S.; Stein, D. J. Parallel Changes in Serum Proteins and Diffusion Tensor Imaging in Methamphetamine-Associated Psychosis. *Sci. Rep.* **2017**, *7* (1), 43777. https://doi.org/10.1038/srep43777.

(205) Keshishian, H.; Burgess, M. W.; Gillette, M. A.; Mertins, P.; Clauser, K. R.; Mani, D. R.; Kuhn, E. W.; Farrell, L. A.; Gerszten, R. E.; Carr, S. A. Multiplexed, Quantitative Workflow for Sensitive Biomarker Discovery in Plasma Yields Novel Candidates for Early Myocardial Injury. *Mol. Cell. Proteomics MCP* **2015**, *14* (9), 2375–2393. https://doi.org/10.1074/mcp.M114.046813.

(206) Wiśniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal Sample Preparation Method for Proteome Analysis. *Nat. Methods* **2009**, *6* (5), 359–362. https://doi.org/10.1038/nmeth.1322.

(207) Wu, Z.; Huang, J.; Huang, J.; Li, Q.; Zhang, X. Lys-C/Arg-C, a More Specific and Efficient Digestion Approach for Proteomics Studies. *Anal. Chem.* **2018**, *90* (16), 9700–9707. https://doi.org/10.1021/acs.analchem.8b02448.

(208) Shuford, C. M.; Sederoff, R. R.; Chiang, V. L.; Muddiman, D. C. Peptide Production and Decay Rates Affect the Quantitative Accuracy of Protein Cleavage Isotope Dilution Mass Spectrometry (PC-IDMS). *Mol. Cell.*

Proteomics **2012**, *11* (9), 814–823. https://doi.org/10.1074/mcp.O112.017145.

(209) Shuford, C. M.; Li, Q.; Sun, Y.-H.; Chen, H.-C.; Wang, J.; Shi, R.; Sederoff, Ronald. R.; Chiang, V. L.; Muddiman, D. C. Comprehensive Quantification of Monolignol-Pathway Enzymes in Populus Trichocarpa by Protein Cleavage Isotope Dilution Mass Spectrometry. *J. Proteome Res.* **2012**, *11* (6), 3390–3404. https://doi.org/10.1021/pr300205a.

(210) Kiseleva, O. I.; Romashova, Y. A.; Moskaleva, N. E.; Petushkova, N. A.; Teryaeva, N. B.; Belyaev, A. Yu.; Lisitsa, A. V. Plasma Preparation to Measure FDA-Approved Protein Markers by Selected Reaction Monitoring. *Clin. Transl. Med.* **2015**, *4* (1), e32. https://doi.org/10.1186/s40169-015-0071-4.

(211) Bagby, S.; Tong, K. I.; Ikura, M. Optimization of Protein Solubility and Stability for Protein Nuclear Magnetic Resonance. *Methods Enzymol.* **2001**, *339*, 20–41. https://doi.org/10.1016/s0076-6879(01)39307-2.

(212) Skoog, D. A.; West, D. M.; Holler, F. J.; Crouch, S. R. Statistical Data Treatment and Evaluation. In *Fundamentals of Analytical Chemistry*; Brooks/Cole: Belmont, CA, 2014; pp 127–135.

(213) Skoog, D. A.; Holler, F. J.; Crouch, S. R. *Principles of Instrumental Analysis*; Cengage Learning, 2017.

(214) Neurath, H. Evolution of Proteolytic Enzymes. *Science* **1984**, *224* (4647), 350–357. https://doi.org/10.1126/science.6369538.

(215) Rej, R. Accurate Enzyme Activity Measurements. *Arch Pathol Lab Med* **1993**, *117* (352), 64.

(216) Jänsch, N.; Colin, F.; Schröder, M.; Meyer-Almes, F.-J. Using Design of Experiment to Optimize Enzyme Activity Assays. *ChemTexts* **2019**, *5*. https://doi.org/10.1007/s40828-019-0095-2.

(217) Promega. *Sequencing Grade Modified Trypsin*. https://worldwide.promega.com/products/mass-

spectrometry/trypsin/sequencing-grade-modified-trypsin/ (accessed 2023-10-28).

(218) Núñez, O.; Lucci, P. Applications and Uses of Formic Acid in Liquid Chromatography-Mass Spectrometry Analysis; 2014; pp 71–86.

(219) Milac, T. I.; Randolph, T. W.; Wang, P. Analyzing LC-MS/MS Data by Spectral Count and Ion Abundance: Two Case Studies. *Stat. Interface* **2012**, *5* (1), 75–87. https://doi.org/10.4310/SII.2012.v5.n1.a7.

(220) Resing, K. A.; Meyer-Arendt, K.; Mendoza, A. M.; Aveline-Wolf, L. D.; Jonscher, K. R.; Pierce, K. G.; Old, W. M.; Cheung, H. T.; Russell, S.; Wattawa, J. L.; Goehle, G. R.; Knight, R. D.; Ahn, N. G. Improving Reproducibility and Sensitivity in Identifying Human Proteins by Shotgun Proteomics. *Anal. Chem.* **2004**, *76* (13), 3556–3568. https://doi.org/10.1021/ac035229m.

(221) Desiere, F.; Deutsch, E. W.; King, N. L.; Nesvizhskii, A. I.; Mallick, P.; Eng, J.; Chen, S.; Eddes, J.; Loevenich, S. N.; Aebersold, R. The PeptideAtlas Project. *Nucleic Acids Res.* **2006**, *34* (suppl_1), D655–D658. https://doi.org/10.1093/nar/gkj040.

(222) Marquioni, V.; Nunes, F. M. F.; Novo-Mansur, M. T. M. Protein Identification by Database Searching of Mass Spectrometry Data in the Teaching of Proteomics. *J. Chem. Educ.* **2021**, *98* (3), 812–823. https://doi.org/10.1021/acs.jchemed.0c00853.

APPENDICES

A. Calibration Curves for BCA Assay

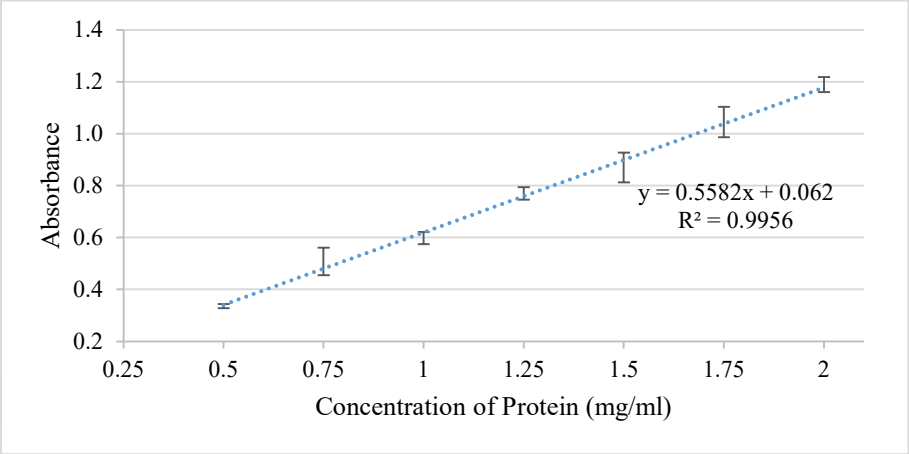


Figure A.1. BCA Assay Calibration Curve (QC-DI-BCA) for Native BSA in Deionized Water. Absorption values are measured at 562 nm.

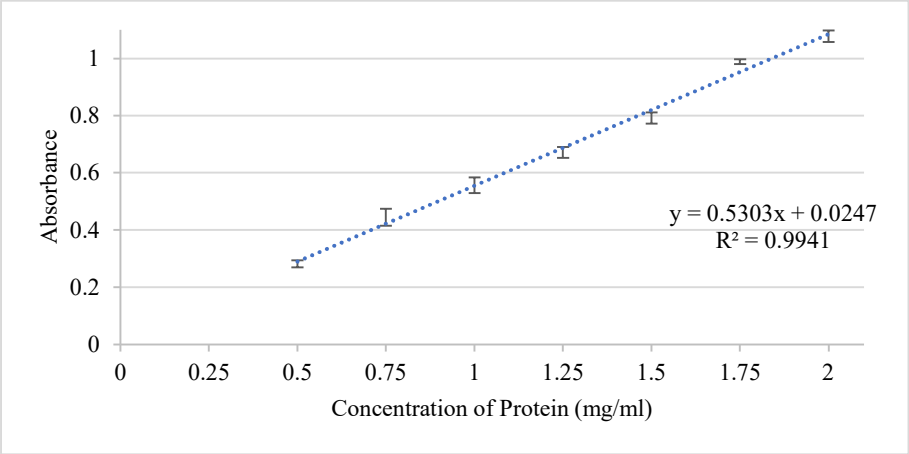


Figure A.2. BCA Assay Calibration Curve for Method A in 50 mM ABC Buffer Absorption values are measured at 562 nm.

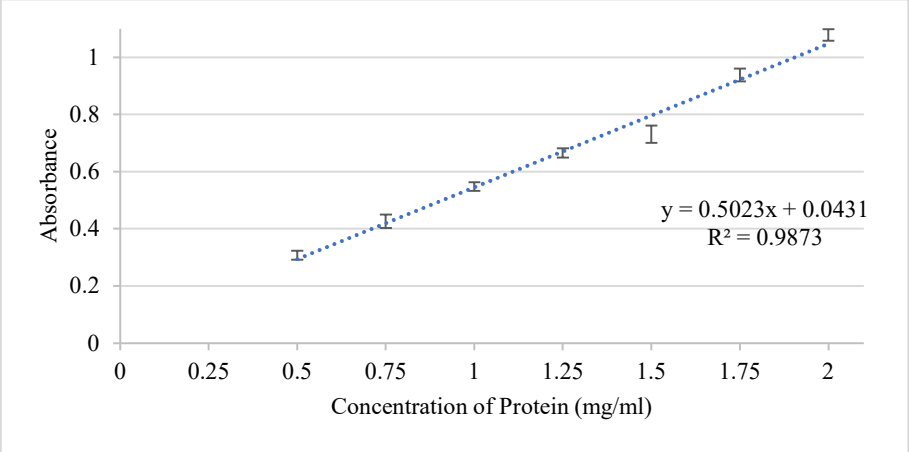


Figure A.3. BCA Assay Calibration Curve for Method B, Method C, and Method J in100 mM Tris Buffer. Absorption values are measured at 562 nm.

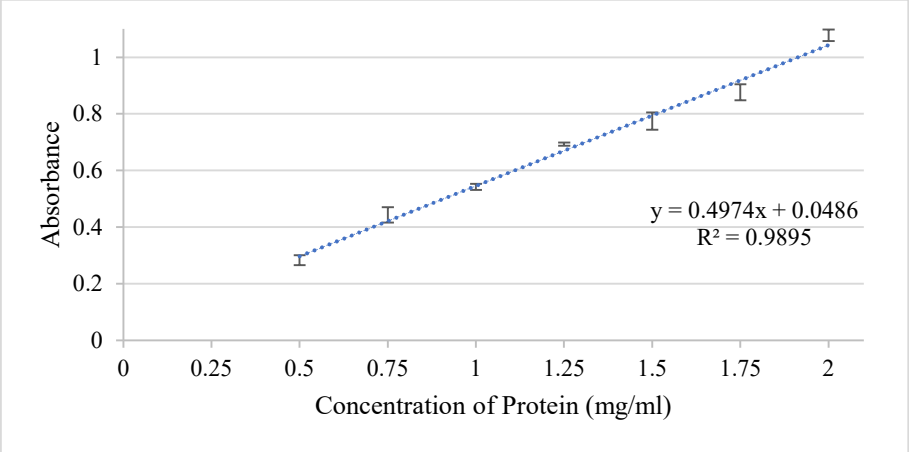


Figure A.4. BCA Assay Calibration Curve for Method D and Method G in 25 mM ABC Buffer. Absorption values are measured at 562 nm.

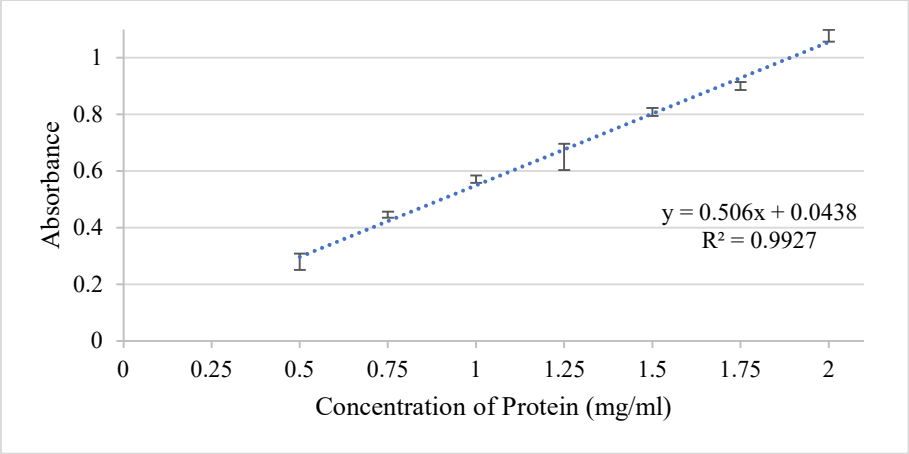


Figure A.5. BCA Assay Calibration Curve for Method E in 100 mM ABC Buffer Absorption values are measured at 562 nm.

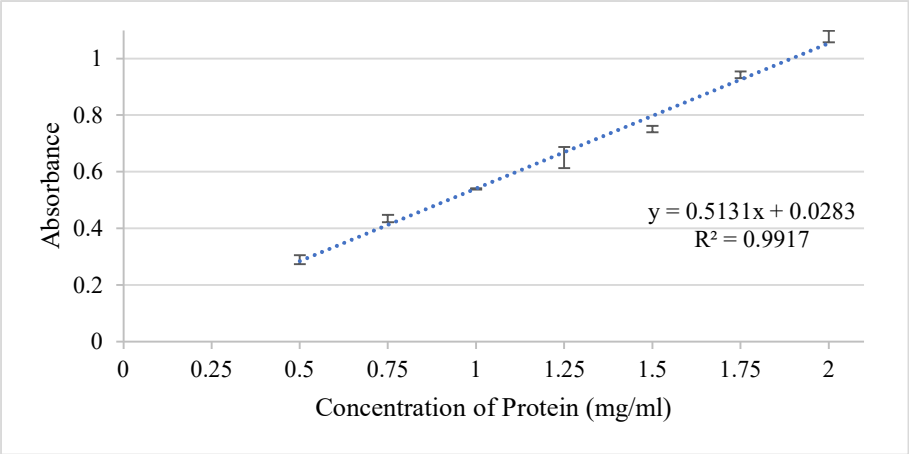


Figure A.6. BCA Assay Calibration Curve for Method F with 10 mM Tris Buffer. Absorption values are measured at 562 nm.

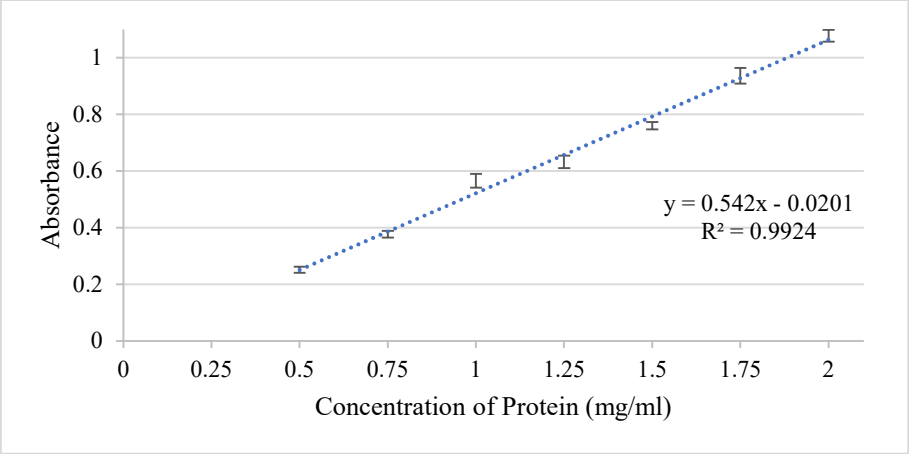


Figure A.7. BCA Assay Calibration Curve for Method H with 300 mM Tris Buffer. Absorption values are measured at 562 nm.

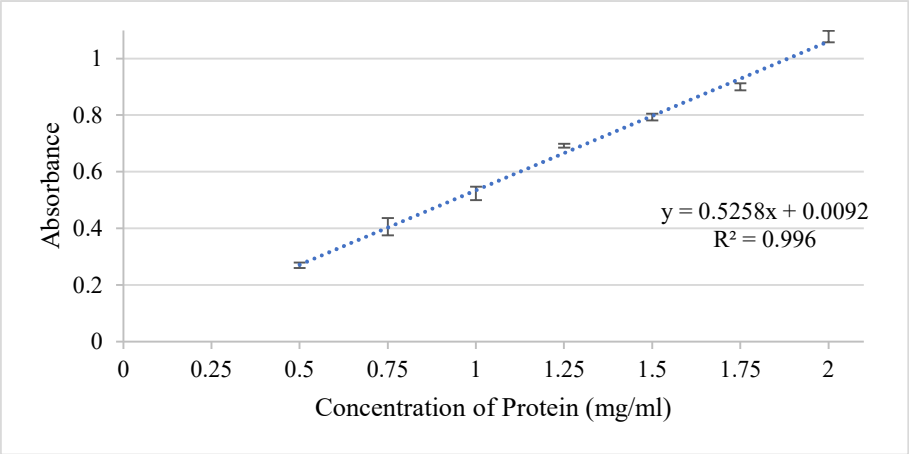


Figure A.8. BCA Assay Calibration Curve for Method I with 200 mM Tris Buffer. Absorption values are measured at 562 nm.

B. Total Ion Chromatogram of A2MG Sample Digests

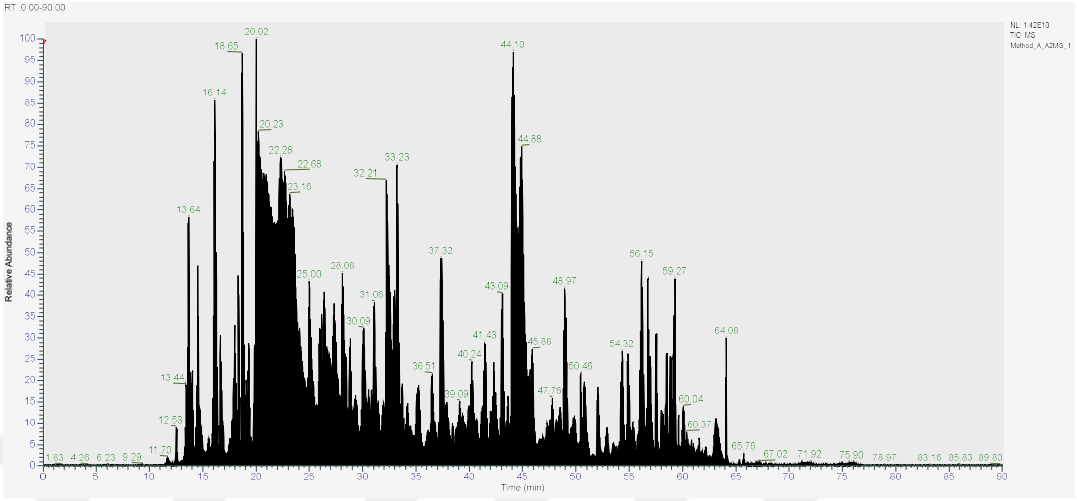


Figure B.9. Total Ion Chromatography of A2MG Protein Digested with Method A – Replicate 1

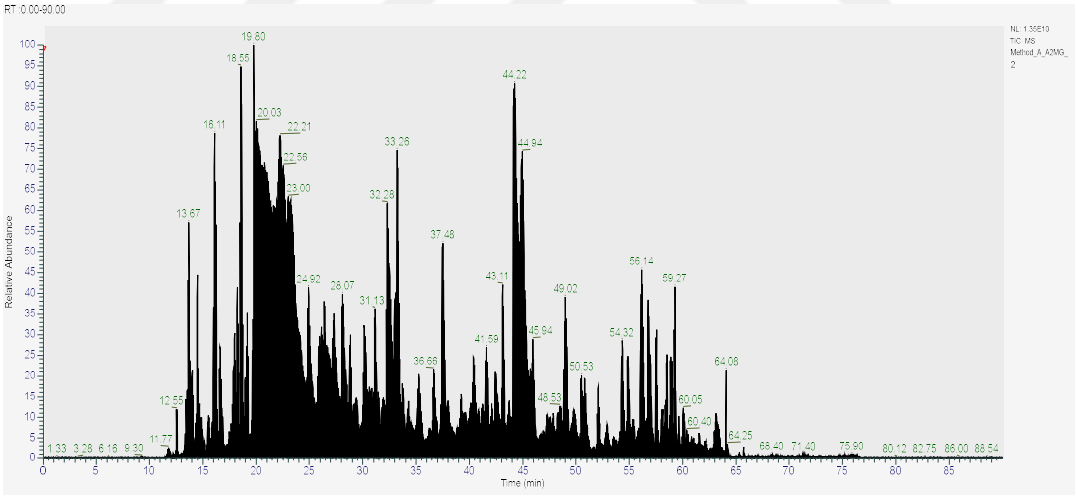


Figure B.10. Total Ion Chromatography of A2MG Protein Digested with Method A – Replicate 2

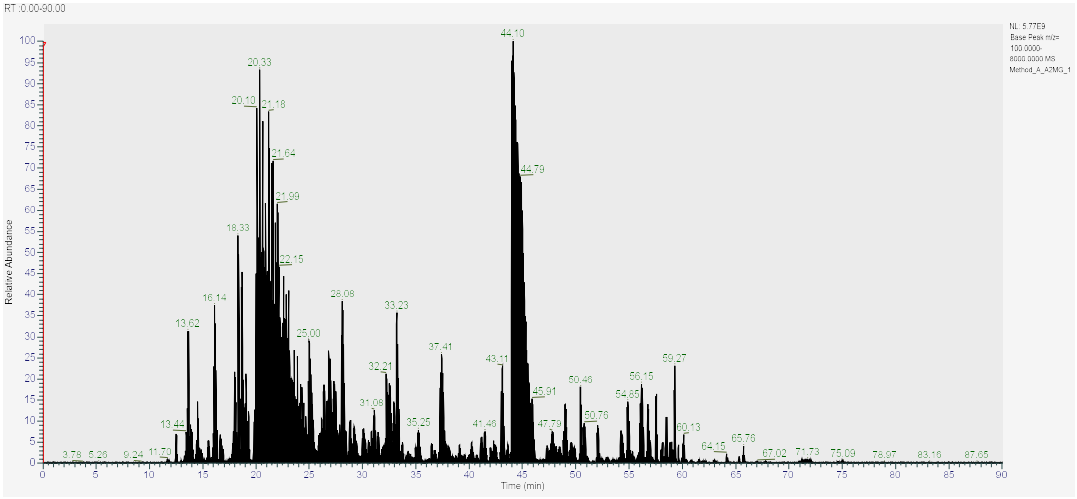


Figure B.11. Base Peak Spectra of A2MG Protein Digested with Method A – Replicate 1

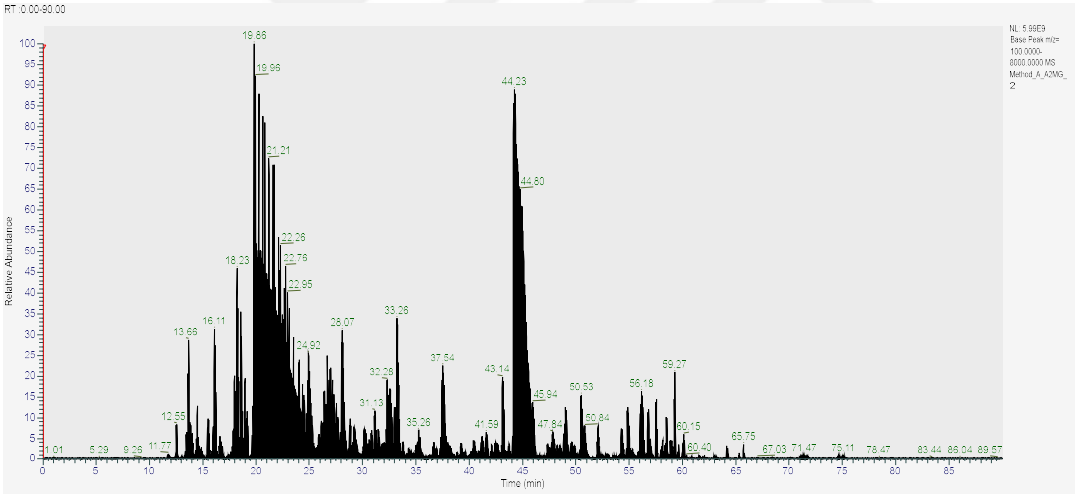


Figure B.12. Base Peak Spectra of A2MG Protein Digested with Method A – Replicate 2

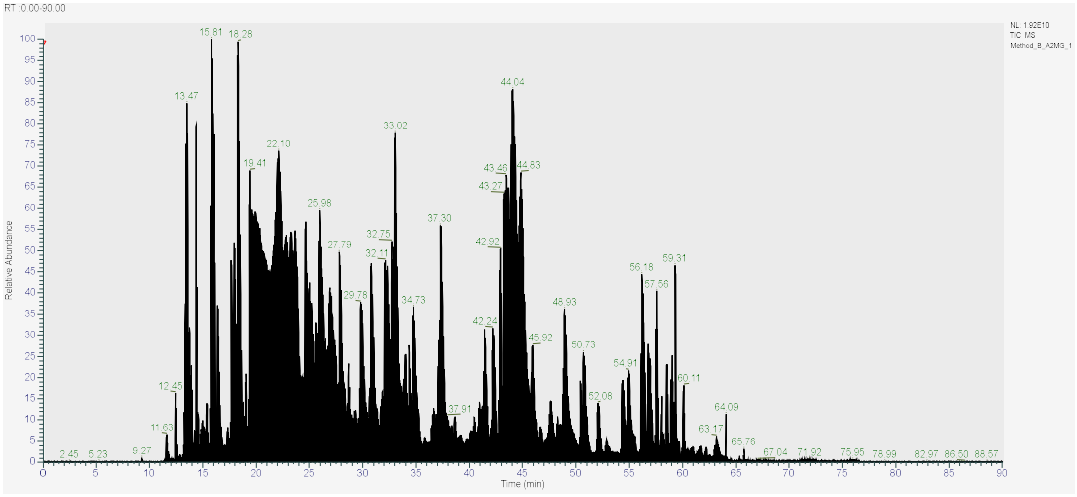


Figure B.13. Total Ion Chromatography of A2MG Protein Digested with Method B – Replicate 1

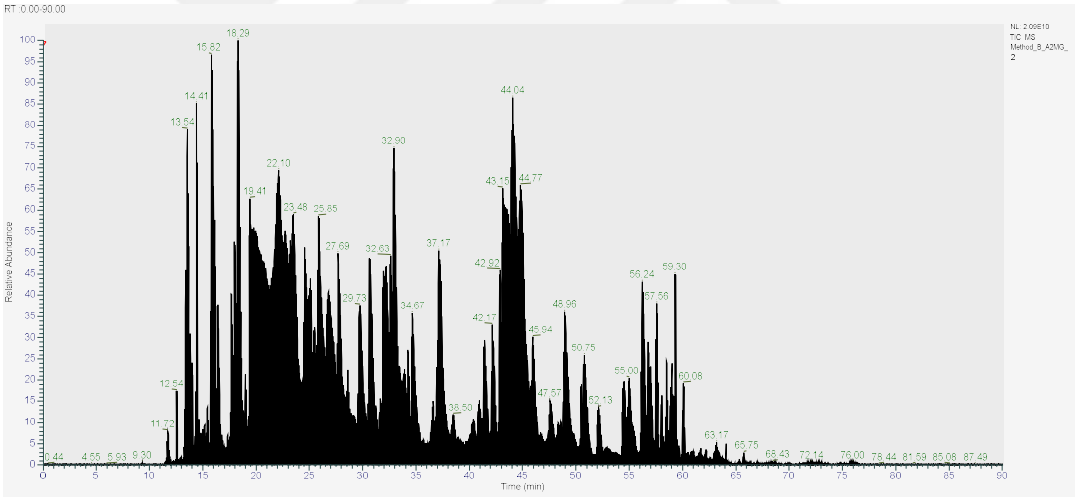


Figure B.14. Total Ion Chromatography of A2MG Protein Digested with Method B – Replicate 2

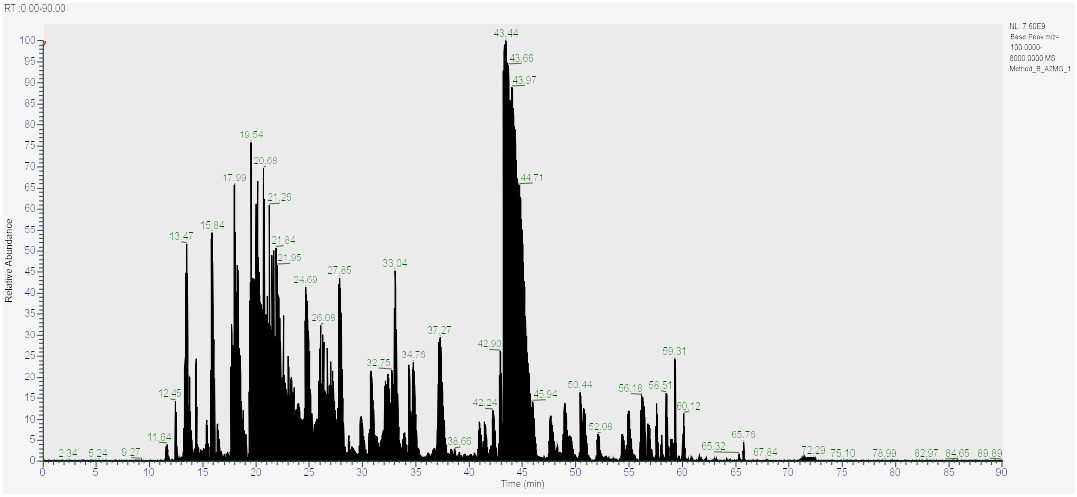


Figure B.15. Base Peak Spectra of A2MG Protein Digested with Method B – Replicate 1

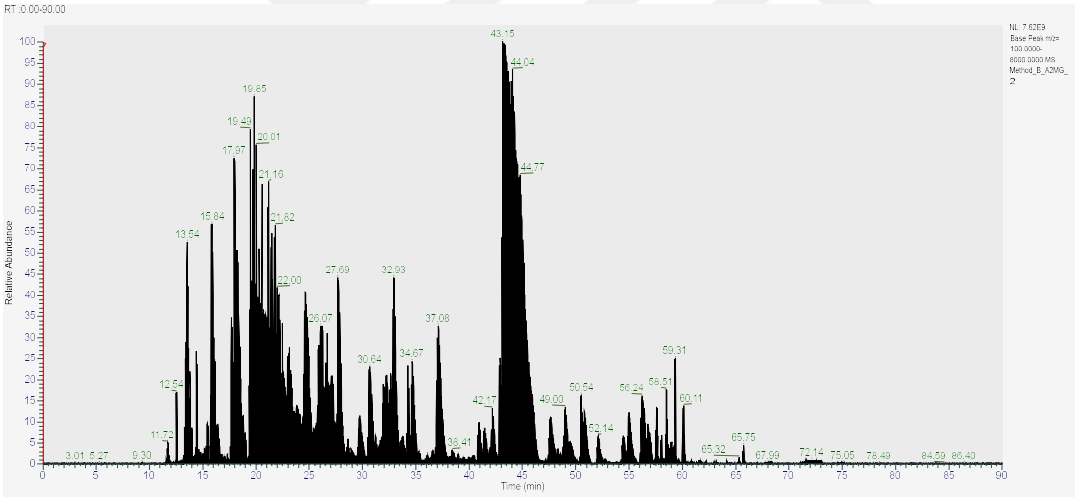


Figure B.16. Base Peak Spectra of A2MG Protein Digested with Method B – Replicate 2

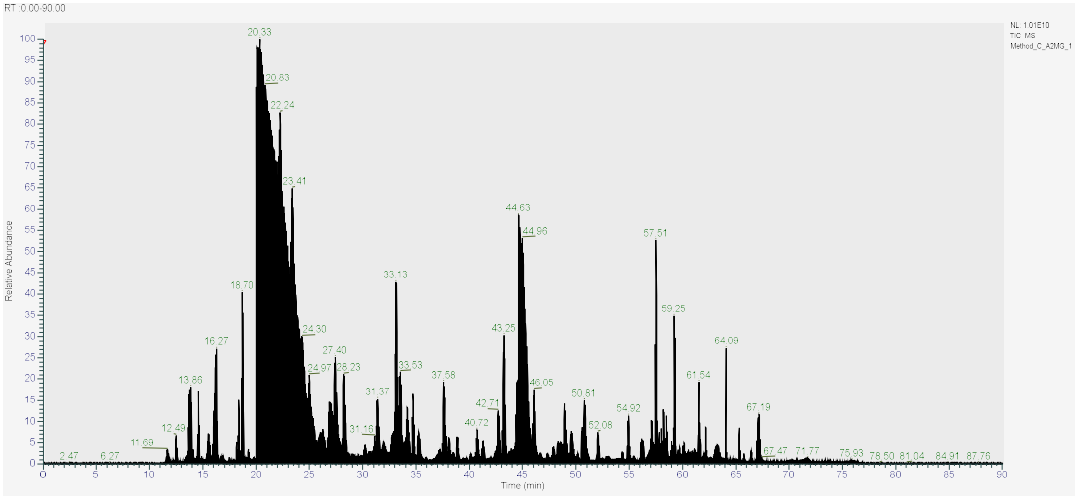


Figure B.17. Total Ion Chromatography of A2MG Protein Digested with Method C – Replicate 1

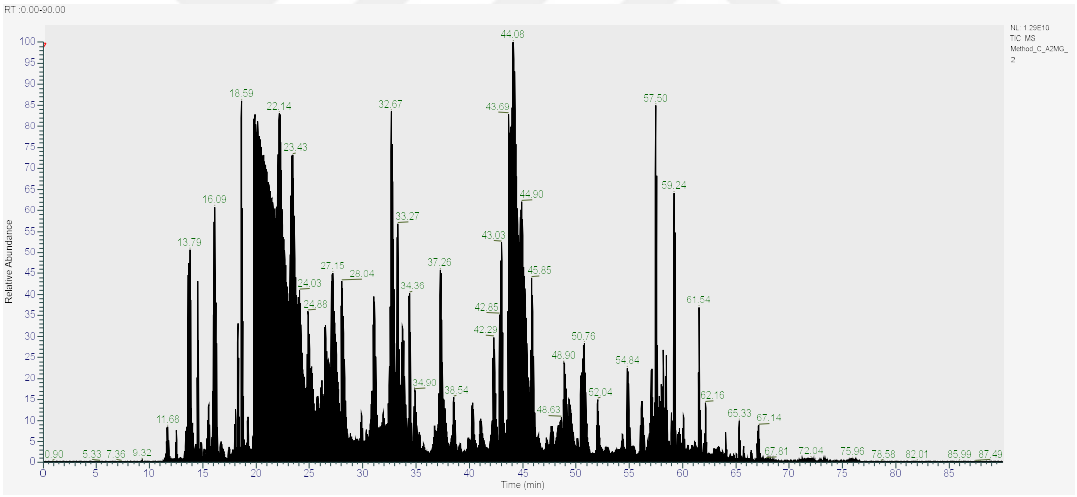


Figure B.18. Total Ion Chromatography of A2MG Protein Digested with Method C – Replicate 2

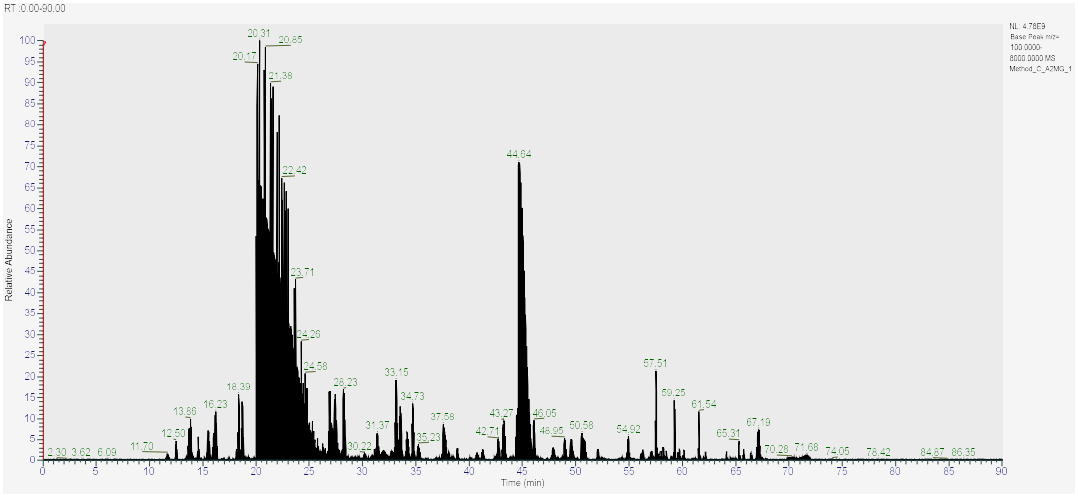


Figure B.19. Base Peak Spectra of A2MG Protein Digested with Method C – Replicate 1

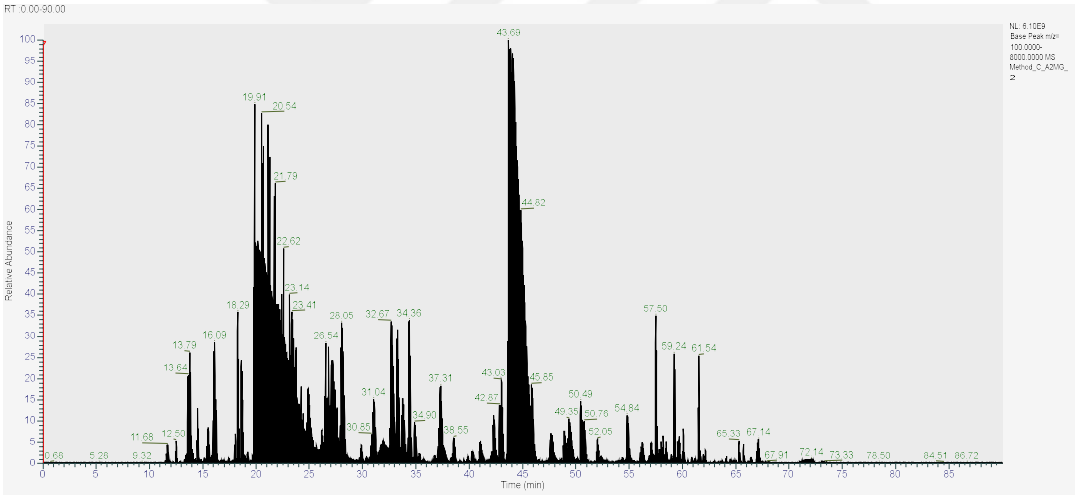


Figure B.20. Base Peak Spectra of A2MG Protein Digested with Method C – Replicate 2

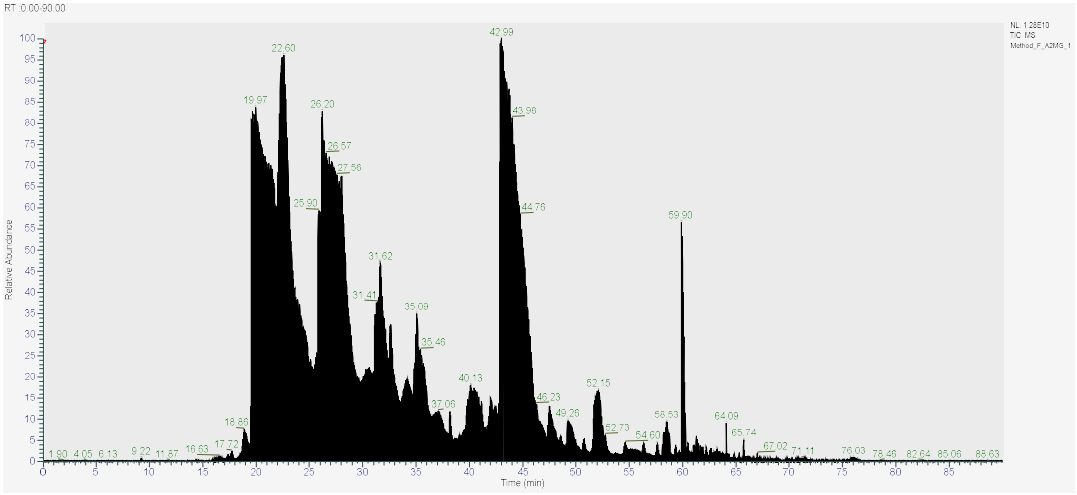


Figure B.21. Total Ion Chromatography of A2MG Protein Digested with Method F – Replicate 1

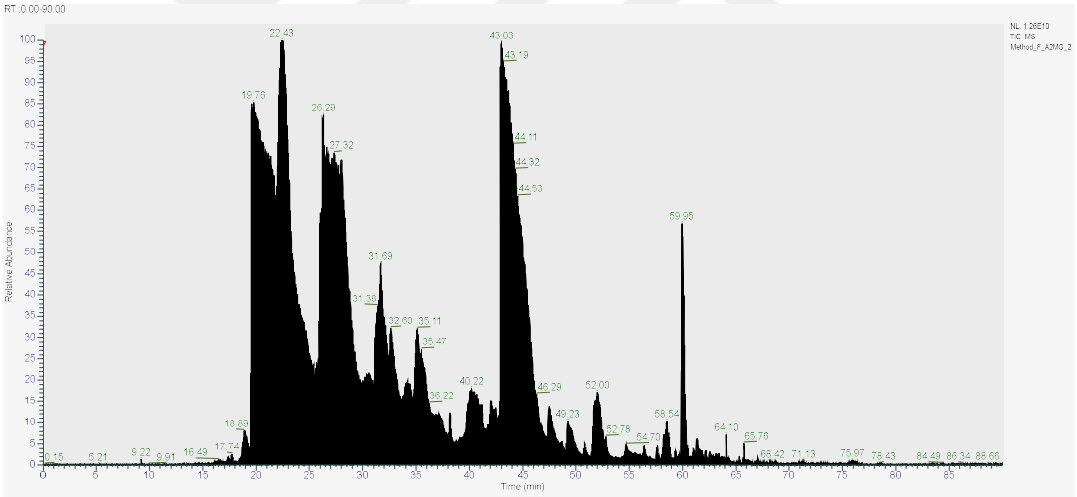


Figure B.22. Total Ion Chromatography of A2MG Protein Digested with Method F – Replicate 2

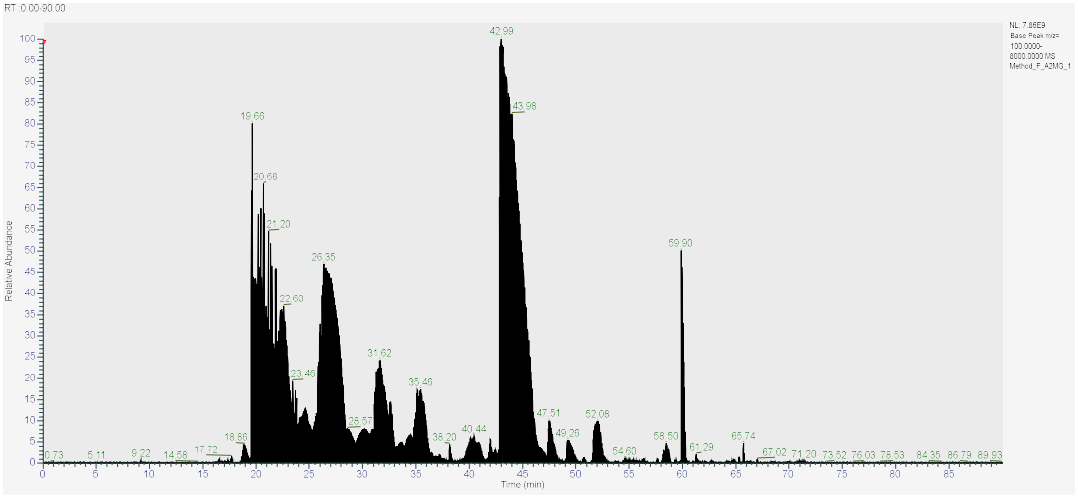


Figure B.23. Base Peak Spectra of A2MG Protein Digested with Method F – Replicate 1

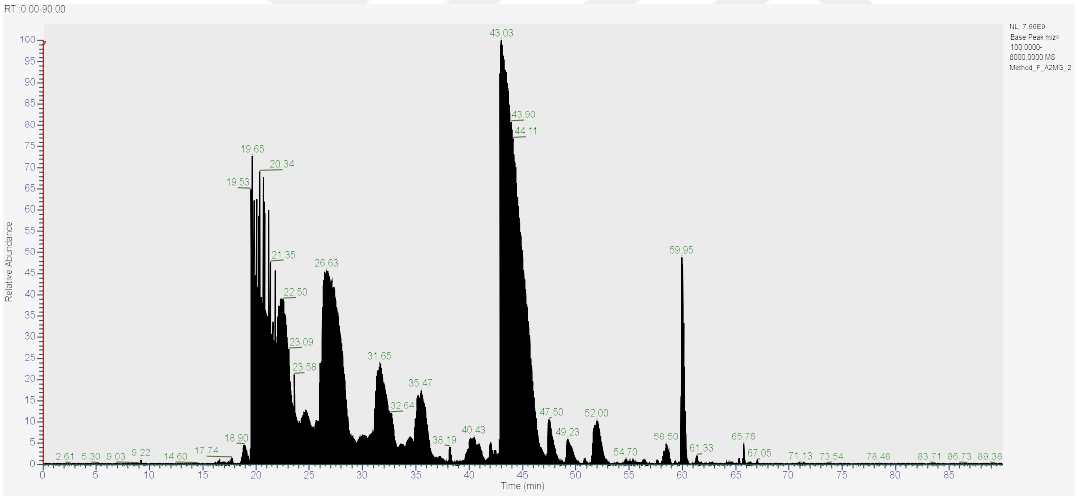


Figure B.24. Base Peak Spectra of A2MG Protein Digested with Method F – Replicate 2

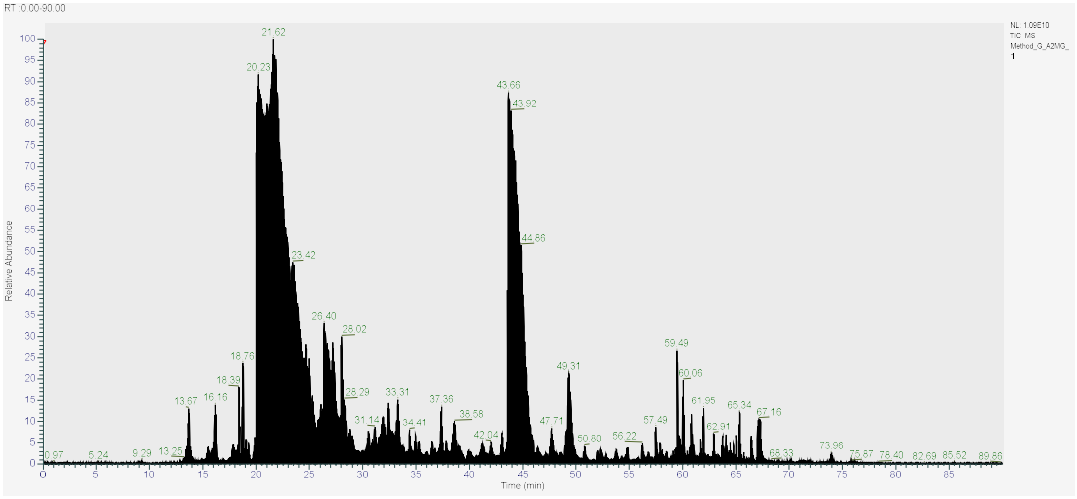


Figure B.25. Total Ion Chromatography of A2MG Protein Digested with Method G – Replicate 1

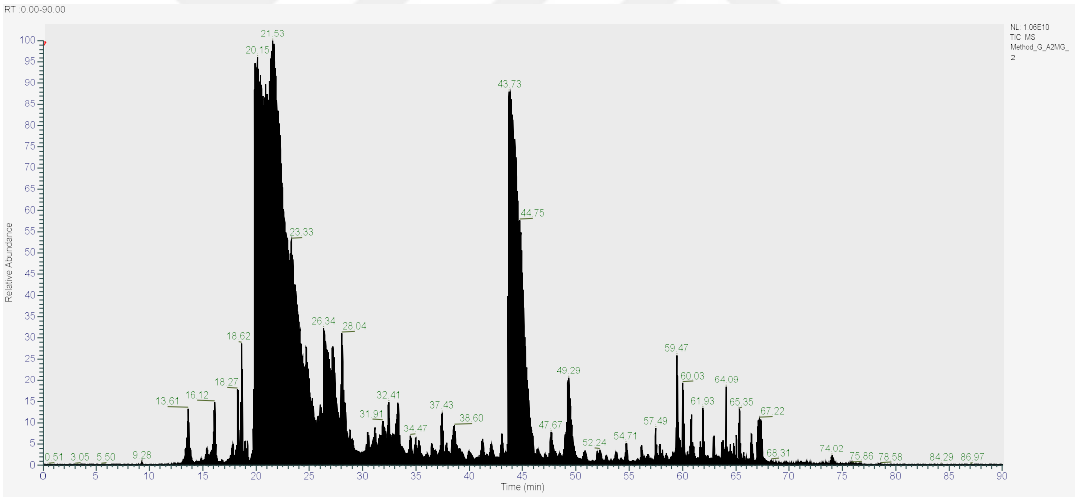


Figure B.26. Total Ion Chromatography of A2MG Protein Digested with Method G – Replicate 2

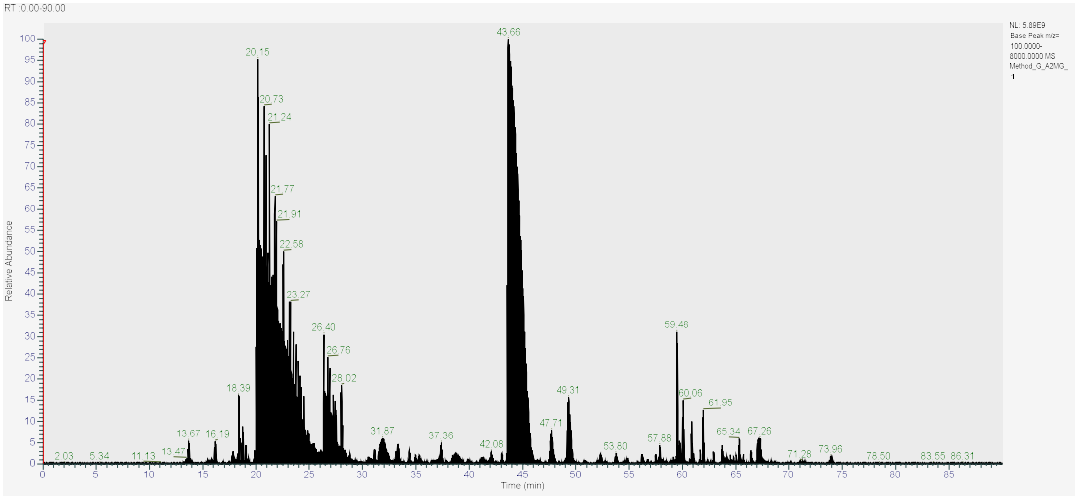


Figure B.27. Base Peak Spectra of A2MG Protein Digested with Method G – Replicate 1

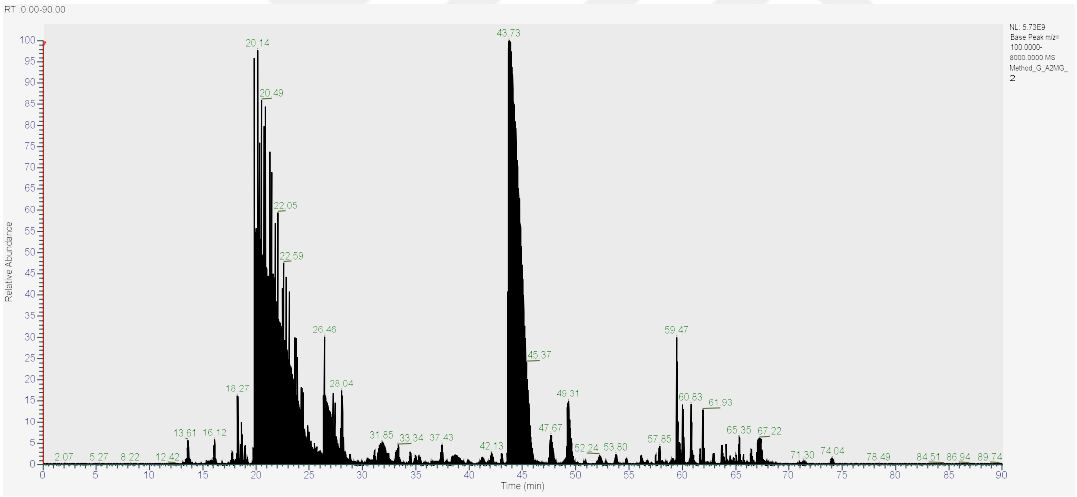


Figure B.28. Base Peak Spectra of A2MG Protein Digested with Method G – Replicate 2

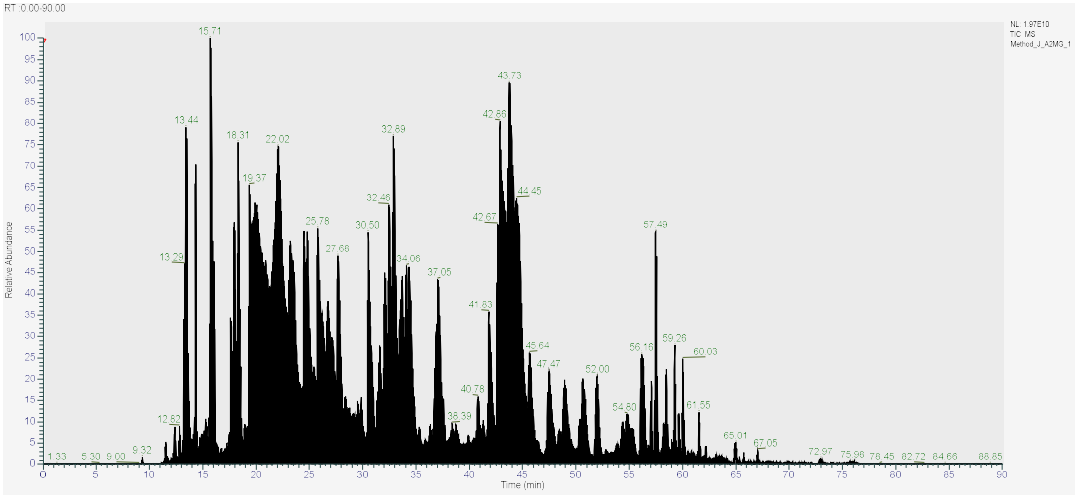


Figure B.29. Total Ion Chromatography of A2MG Protein Digested with Method J – Replicate 1

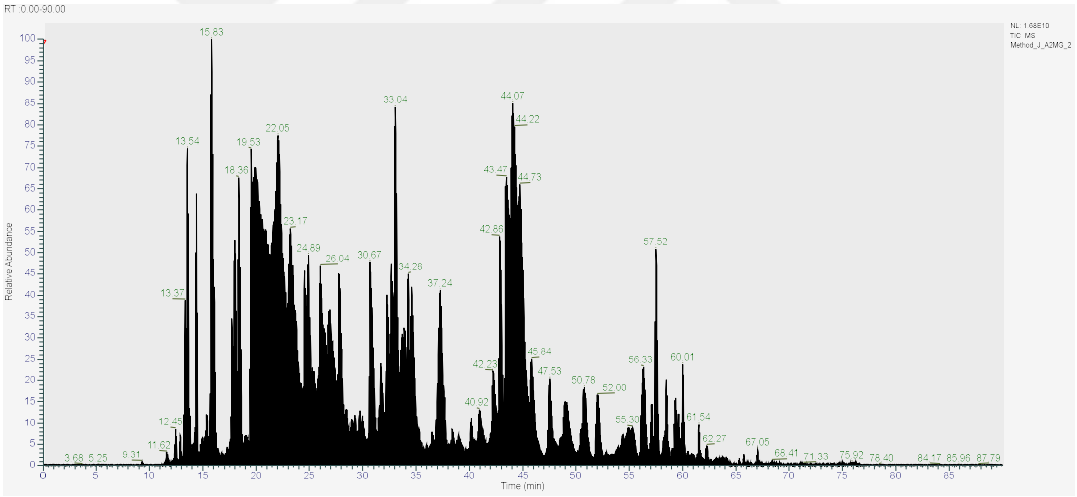


Figure B.30. Total Ion Chromatography of A2MG Protein Digested with Method J – Replicate 2

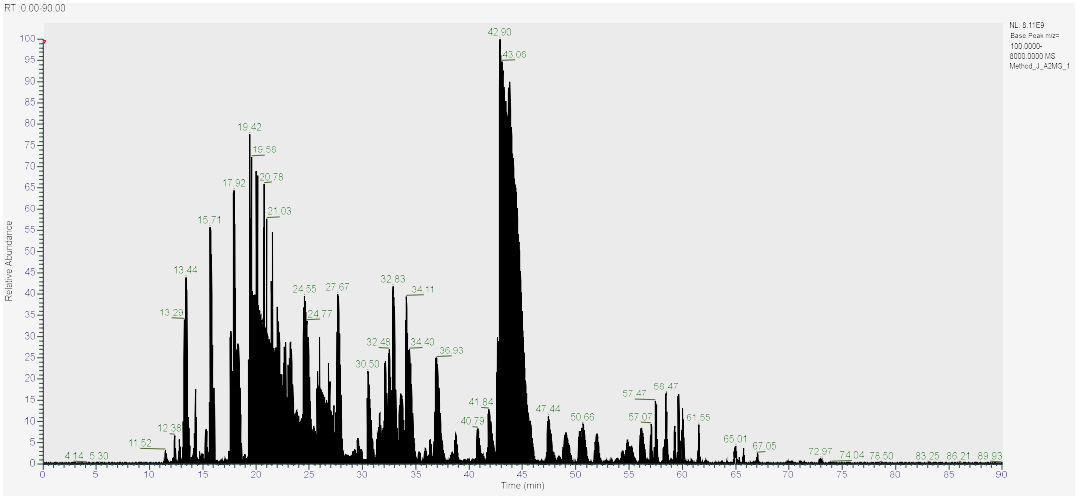


Figure B.31. Base Peak Spectra of A2MG Protein Digested with Method J – Replicate 1

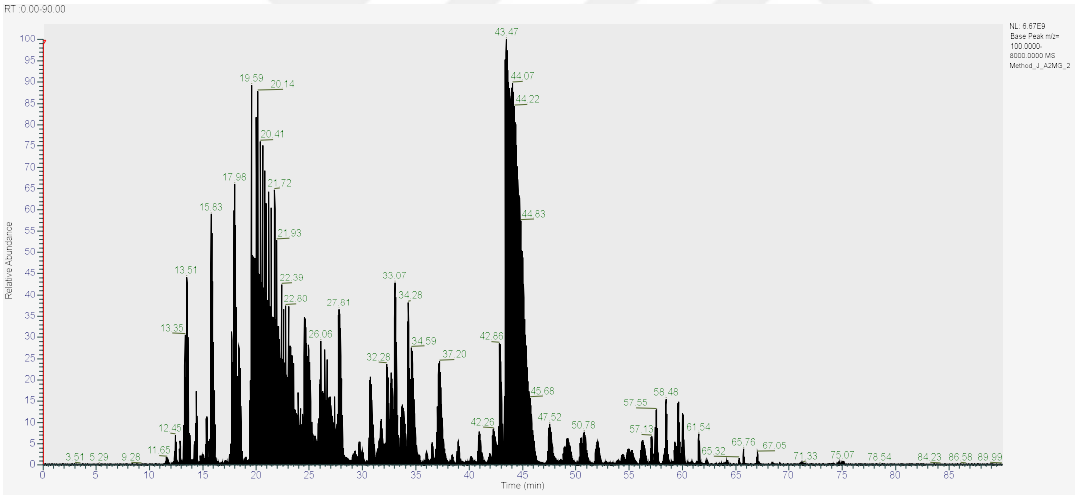


Figure B.32. Base Peak Spectra of A2MG Protein Digested with Method J – Replicate 2