2008

Mass Spectrometric Analysis of Environmental Contaminants, Protein Structure and Expression

Ian E. Atkinson
Cleveland State University

MASS SPECTROMETRIC ANALYSIS OF ENVIRONMENTAL CONTAMINANTS, PROTEIN STRUCTURE AND EXPRESSION

IAN E. ATKINSON

Bachelor of Science in Zoology

Miami University

August 1996

Master of Science in Chemistry

Cleveland State University

May 2001

Submitted in partial fulfillment of requirement for the degree

DOCTOR OF PHILOSOPHY IN CLINICAL AND BIOANALYTICAL CHEMISTRY

at the

CLEVELAND STATE UNIVERSITY

DECEMBER, 2008

This dissertation has been approved

for the Department of CHEMISTRY

and the College of Graduate Studies by

_____

Dissertation Chairperson, Dr. Robert Wei

_____

Department & Date

_____

Dr. Lily Ng

_____

Department & Date

_____

Dr. John Masnovi

_____

Department & Date

_____

Dr. Xiang Zhou

_____

Department & Date

_____

Dr. Teresa Cutright

_____

Department & Date

Acknowledgements

This dissertation is dedicated to my loving parents, Linda & Les, without whose encouragement and support none of this would have been possible. They have taken every opportunity to enrich my life and education, and for that I am beyond thankful. They are truly the most wonderful parents a son could ask for.

I would like to deeply thank my advisor, Dr Robert Wei. His direction, guidance and patience have been fundamental in my progress and development as a scientist, a researcher, and an educator. It has been an honor and a privilege working under his tutelage, and I look forward to continuing our work for the good of society and the world.

I would like to express how very grateful I am for Dr Lily Ng's involvement in my life. Her wise counsel and insight has been fundamental in my education. Without her I would have never found my place here at Cleveland State University.

I would like to thank Dr Alan Riga and Dr Jerry Mundell, both of whom took a tremendous personal interest in my professional development and have helped prosper a wonderful outlook on science education and research within me. I owe so much of my holistic growth while at Cleveland State to these two gentlemen.

I would like to thank Dr Xiang Zhou for so many countless days of training and education on mass spectrometry. I will miss our conversations.

I would like to thank my family whose love and concern has motivated me to succeed. My brother David, sister-in-law Kristin, my niece Molly and nephew Colin, all my love, you are the light of my life. My Aunt Sondra, who helped me tremendously with editing this work, my Uncle Joe, my Aunt Wendy, my Uncle Fred, my cousins

Aaron and Elliot, I cannot express how much your endless encouragement to follow my own path has meant to me.

I would like to express my humblest thanks to all of my friends, Adam, Amanda, Ann, Bill, Brian, Carrie, Chad, Curt, Chamari, Deley, Eric, Jason, Jim, Jimmy, Jinhua, John, Kate, Kerri, Lynn, Meg, Mike, Miles, Pam, Pat, Paul, Ram, Rachel, Robert, Brooke & Tony, Isis & Ron, Kelly & Kevin, Meghan & Michael, Melissa & Michael, Nancy & Dick, Nancy & Jeff, Rachel & Ray, Ruth & David, Sheila & Johnny, Wendy & David and many more that I do not have the room to mention here. I started this project in the hopes of taking direction of my dreams to reach a point wherein I could enact change upon the world. It has been my friends support and strength over the years that has held me up, kept my life full and allowed me to see this through.

I would also like to express my deep thanks and fullest respect to Mr. Ray Fogg. My mentor, role-model and dear friend, who was very much responsible in my choice of continuing my education and this path I have taken in life. I take the next steps with your inspiration and I follow in your footsteps to ensure tikkun olam.

MASS SPECTROMETRIC ANALYSIS OF ENVIRONMENTAL CONTAMINANTS, PROTEIN STRUCTURE AND EXPRESSION

IAN E. ATKINSON

ABSTRACT

There are a wide array of sciences at our disposal to further our understanding of the roles of toxins in the environment and their effects on living organisms. Of these disciplines, the field of mass spectrometry offers a powerful analytical means to identify compounds in varying and complex sample media. The presence of altered expressions of biomolecules upon exposure to toxins, the effects that these toxins may have on biomolecule structure and the identity of toxins in the sample all can be determined by mass spectrometry.

In this study mass spectrometric analysis has been applied to environmental toxicological problems. First, the altered protein expression under various metal stress in *Helianthus annuus* (dwarf sunflower), a proven hyperaccumulator of toxic metals was analyzed by this research. Second, a model was developed for mass spectrometrically determining the location and structural effects that another class of environmental toxins, poly aromatic hydrocarbons (PAHs) can have on proteins. Third, the presence of PAHs in environmental samples taken from Lake Bolgoda and Lake Beira, Sri Lanka, was analyzed by mass spectrometry to determine the types, relative concentrations and potential sources of these toxins.

The work has demonstrated the versatility and effectiveness of the methodology, discovering novel protein expression in *H. annuus* upon metal exposure, pinpointing site

mediated adductions reactions on protein structure, and the qualitative and quantitative determination of toxins in environmental samples.

TABLE OF CONTENTS

ABSTRACT…….…… vi

TABLE OF CONTENTS……..…viii

LIST OF TABLES……..…xii

LIST OF FIGURES……...…….…..xiv

CHAPTER I LC ESI ION TRAP MS DETERMINATION OF ALTERED PROTEIN EXPRESSION IN *HELIANTHUS ANNUUS* AFTER EXPOSURE TO HEAVY METALS As, Cd, Cr, Ni AND Pb……...…….....…….1

1.1. Introduction…….……2

1.2. Methodology……..…….. 13

1.2.1. Plant growth and exposure.……..……13

1.2.2. Protein extraction and gel separation…..……...…..……14

1.2.3. Sample preparation and tryptic digest…..……..…….14

1.2.4. LC separation…….……..……16

1.2.5. MS analysis…….……..……...17

1.2.6. MASCOT bioinformatics…….……..……..17

1.3. Results…….……...……..18

1.3.1. Expression of proteins in Sundance……..……..18

1.3.1.1. Expression of proteins in response to arsenic and other metals in Sundance……18

1.3.1.2. Expression of proteins in response to arsenic only in Sundance..24

1.3.1.3. Expression of proteins in response to lead and other metals in

Sundance……...25

1.3.2. Expression of proteins in Teddy Bear…….30

1.3.2.1. Expression of proteins in response to arsenic and other metals in Teddy Bear……30

1.3.2.2. Expression of proteins in response to chromium only in Teddy Bear…….....30

1.3.2.3. Expression of proteins in response to lead and other metals in Teddy Bear……31

1.3.2.4. Expression of proteins in response to lead and other metals 2D gel separation in Teddy Bear……..34

1.4. Discussion……...42

1.5. Conclusions…….54

1.6. References……...58

CHAPTER II STUDY ON ADDUCTION REACTIONS IN CHYMOTRYPSIN'S ACTIVE SITE……...67

2.1. Background and Significance…….68

2.2. Methodology……...77

2.2.1. CHY adduct formation…….……77

2.2.1.1. CHY-TPCK adduct reaction……..…….…….77

2.2.1.2. Proposed CHY-PAH adduct reaction……...77

2.2.2. Whole protein adduct analyses…….……78

2.2.2.1. LC separation and QQQ analysis…….……78

2.2.2.2. LC separation and Ion Trap Analysis……...……79

2.2.3. Hydrazinolysis analysis…….……...80

2.2.3.1. Hydrazinolysis……..80

2.2.3.2. LC separation and QQQ analysis…….80

2.2.3.3. LC separation and Ion Trap analysis……81

2.2.4. Enzymatic digest analysis……83

2.2.5. Functional analysis of protein……..85

2.3. Results & Discussion

2.3.1. QQQ MS characterization of adduct in chymotrypsin……...…..…85

2.3.2. LC ESI-MS Ion characterization of adduct in chymotrypsin...…..…..90

2.3.3. UV spectroscopic analysis……..…….…….92

2.3.4. Tryptic digest……..……..92

2.4. Conclusions…….93

2.5. References……...95

CHAPTER III MS DETERMINATION OF POLYCYCLIC AROMATIC HYDROCARBONS (PAHS) IN SEDIMENT AND WATER SAMPLES FROM BOLGODA AND BEIRA LAKES…….……..98

3.1. Abstract……...99

3.2. Introduction……...101

3.3. Methodology…….103

3.3.1. Sample site selection & collection……...……...103

3.3.2. Sample preparation…….……105

3.3.3. HPLC Sample analysis……..…….106

3.3.4. GCMS Sample analysis……..……106

3.3.5. Evaluation of sediment toxicity……..……...107

3.4. Results & Discussion…….…..……109

3.4.1. Beira Lake……..109

3.4.2. Bolgoda Lake…….…….113

3.4.3. PAH Ratios…….……116

3.4.4. Microtox Assay…….……..119

3.5. References…….121

APPENDICES……...……..124

A. Additional tables & figures for chapter 1……...…….125

B. Additional figures for chapter 2…….....…….197

LIST OF TABLES

CHAPTER I

Table 1-I. 1-D Hydroponic Sundance Root As, Cd, Cr, Ni……21

Table 1-II. 2-D Teddy Bear Leaf Hydroponic Cd, Cr, Ni, Pb…….....37

CHAPTER II

Table 2-I. Secondary structure of chymotrypsin, individual amino acids molecular weight and the molecular weight of the expected amino acid hydrazides.89

CHAPTER III

Table 3-I. Average PAH levels in Beira and Bolgoda Lakes……...110

Table 3-II. Ratios of two-three ring to four-five ring PAHs for Water and Sediment samples from Beira and Bolgoda Lakes……..117

Table 3-III. Phenanthrene/anthracene (P/A) and fluoranthene/pyrene (F/P) ratios for water and sediment samples from Beira and Bolgoda Lakes……..118

APPENDIX A

Table A-I. 1-D Hydroponic Sundance Leaf As only…….129

Table A-II. 1-D Sundance Leaf Hydroponic Cd, Cr, Ni, Pb……..136

Table A-III. 1-D Sundance Stem Hydroponic Cd, Cr, Ni, Pb……..147

Table A-IV. 1-D Sundance Leaf Hydroponic (soluble fraction) Cd, Cr, Ni, Pb……..151

Table A-V. 1-D Sundance Leaf Soil 25-35 kDa EDTA samples Cd, Cr, Ni, Pb……156

Table A-VI. 1-D Teddy Bear Leaf Soil 25-35 kDa EDTA samples Cd, Cr, Ni, Pb…160

Table A-VII. 1-D Teddy Bear Leaf Soil Cr only……...163

Table A-VIII. 1-D Teddy Bear Leaf Hydroponic Cd, Cr, Ni, Pb……...167

Table A-IX. 1-D Teddy Bear Stem Hydroponic Cd, Cr, Ni, Pb……..172

Table A-X. 1-D Teddy Bear Leaf Hydroponic (Soluble fraction) Cd, Cr, Ni, Pb.....176

Table A-XI. 1-D Teddy Bear Leaf Hydroponic(Transmembrane fraction)Cd, Cr, Ni, Pb…….179

Table A-XII. 2-D Teddy Bear Leaf Hydroponic Cd, Cr, Ni, Pb……...181

LIST OF FIGURES

CHAPTER I

Figure 1-1. Tandem mass spectra of polypeptide from Chitinase (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……20

Figure 1-2. Tandem mass spectra of peptide from an unknown protein type with high sequence homology to pathogenesis-related protein 5-1 (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……..22

Figure 1-3. 1-D Gel Separation Sundance Hydroponic As……..23

Figure 1-4. 2-D Gel Separation Teddy Bear Leaf Hydroponic Cd, Cr, Ni, Pb……36

Figure 1-5. Tandem mass spectra of polypeptide from Chitinase (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……38

Figure 1-6. Tandem mass spectra of polypeptides from thaumatin-like protein (*Solanum tuberosum*) as indicated (bold) in the primary amino acid sequence…….40

Figure 1-7. Semi-quantitative RT-PCR of total RNA samples showing an up-regulation of Chitinase in samples containing As [11]……..43

Figure 1-8. Western Blot analysis of Sundance and Teddy Bear samples showing an up-regulation of Chitinase in samples exposed to As [11]……44

CHAPTER II

Figure 2-1. Conc. of Benzo [a] pyrene detected in air, soil, water, and plants and food……..……..68

Figure 2-2. Metabolism of Benzo [a] Pyrene……..….70

Figure 2-3. Catalytic core of Chymotrypsin, a model Serine Peptidase……...72

Figure 2-4. TPCK reaction/inhibition of CHY……..…73

Figure 2-5. Representative diagram of potential equilibrium reactions between an electrophile E and neuleophilic residues N_1 and N_2……..….75

Figure 2-6. QQQ Mass spectra of histidine and TPCK, optimization of small molecule analysis……..…….82

Figure 2-7. QQQ Mass spectra of TPCK and daughter analysis of parent molecule, optimization of small molecule analysis CID……...…….83

Figure 2-8. LC-MS QQQ Whole CHY-TPCK (top) and CHY (bottom)……..…87

Figure 2-9. LC MS QQQ Separation of hydrazine digests of CHY-TPCK (top) and CHY (bottom)……88

Figure 2-10. LC MS QQQ hydrazinolysis peak spectra at 42.6 min in CHY-TPCK digest corresponding to Histidine – Adduct……...88

Figure 2-11. LC MS Ion Trap Separation of hydrazine digests of CHY-TPCK (top) and CHY (bottom)……91

Figure 2-12. LC MS Ion Trap hydrazinolysis peak spectra at 46.4 min in CHY-TPCK digest corresponding to Histidine – Adduct…….…….91

CHAPTER III

Figure 3-1. Sampling sites in Beira Lake (A) and Bolgoda Lake (B)……109

Figure 3-2. Relative distributions of PAHs for four sites in the water column and sediments in Beira Lake……...112

Figure 3-3. Relative distributions of PAHs for six sites in the water column and sediments in Bolgoda Lake……..115

Figure 3-4. Correlation between EC50 values obtained by Microtox assays and total PAHs for the sediments……...120

APPENDIX A

Figure A-1. Tandem mass spectra of peptide from a CC-NBS-LRR RGA type protein (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……...126

Figure A-2. Tandem mass spectra of polypeptides from Lectin (*Helianthus tuberosus*) as indicated (bold) in the primary amino acid sequence……..127

Figure A-3. Tandem mass spectra of peptides from Actin (*Mesotigma viride*) as indicated (bold) in the primary amino acid sequence in comparison to primary amino acid sequence of (*H. annuus*)……..130

Figure A-4. Tandem mass spectra of peptides from Glyceraldehyde 3-phosphate dehydrogenase B subunit (*Arabidopsis thaliana*) as indicated (bold) in the primary amino acid sequence in comparison to primary amino acid sequence of Glyceraldehyde 3-phosphate dehydrogenase (*H. annuus*)...133

Figure A-5. Tandem mass spectra of peptide from Glycolate Oxidase (*Arabidopsis thaliana*) as indicated (bold) in the primary amino acid sequence……..135

Figure A-6. Tandem mass spectra of polypeptide from Chitinase (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……..137

Figure A-7. Tandem mass spectra of peptide from 33 kDa Oxygen-evolving protein (*Arabidopsis thaliana*) as indicated (bold) in the primary amino acid sequence……...138

Figure A-8. Tandem mass spectra of peptide from photosystem II protein D1 (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……...139

Figure A-9. Tandem mass spectra of peptide from Triosephosphate Isomerase (TIM) (*Arabidopsis thaliana*) as indicated (bold) in the primary amino acid sequence……...140

Figure A-10. Tandem mass spectra of hevein-like antimicrobial (*Euonymus europaeus*) as indicated (bold) in the primary amino acid sequence……..142

Figure A-11. Tandem mass spectra of ATP synthase SF1 alpha subunit (*H. annuus*) as indicated (bold) in the primary amino acid sequence……..143

Figure A-12. Tandem mass spectra of Drought-induced protein SDi-6 (*H. annuus*) as indicated (bold) in the primary amino acid sequence……..145

Figure A-13. Tandem mass spectra of FMTEVR fragment matched to Glucosyltransferase-like protein (*Arabidopsis thaliana*) as indicated (bold) in the primary amino acid sequence……146

Figure A-14. Tandem mass spectra of Quinone oxido-reductase-like protein (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……148

Figure A-15. Tandem mass spectra of Phenylcoumaran benzylic ether reductase homolog (*Tsuga heterophylla*) as indicated (bold) in the primary amino acid sequence……...149

Figure A-16. Tandem mass spectra of peptide from Triosephosphate Isomerase (TIM) (*Arabidopsis thaliana*) as indicated (bold) in the primary amino acid sequence……...152

Figure A-17. Tandem mass spectra of peptide from Triosephosphate Isomerase (*Arabidopsis thaliana*) as indicated (bold) in the primary amino acid sequence……...154

Figure A-18. Tandem mass spectra of peptide from Oxygen-evolving enhancer protein 1, chloroplast precursor (OEE1) (*Arabidopsis thaliana*) as indicated (bold) in the primary amino acid sequence…….157

Figure A-19. Tandem mass spectra of polypeptide from Chitinase (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……..159

Figure A-20. Tandem mass spectra of polypeptide from HSP70-related protein (*H. annuus*) and HSP70 (*Nicotiana benthamiana*) as indicated (bold) in the primary amino acid sequence……...……162

Figure A-21. Tandem mass spectra of polypeptide from cytosolic phosphoglycerate kinase (*H. annuus*) as indicated (bold) in the primary amino acid sequence……...164

Figure A-22. Tandem mass spectra of polypeptides from cytochrome f (*H. annuus*) as indicated (bold) in the primary amino acid sequence……..165

Figure A-23. Tandem mass spectra of polypeptides from malate dehydrogenase precursor (*Medicago sativa*) as indicated (bold) in the primary amino acid sequence……...168

Figure A-24. Tandem mass spectra of polypeptides from chloroplast ferredoxin-NADP+ oxidoreductase precursor (*Capsicum annuum*) as indicated (bold) in the primary amino acid sequence……...170

Figure A-25. Tandem mass spectra of polypeptides from NADPH oxidoreductase, putative (*Arabidopsis)* as indicated (bold) in the primary amino acid sequence……...173

Figure A-26. Tandem mass spectra of polypeptides from malate dehydrogenase precursor (*Medicago sativa*) as indicated (bold) in the primary amino acid sequence……...174

Figure A-27. Tandem mass spectra of polypeptide from Chitinase (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……..177

Figure A-28. Tandem mass spectra of polypeptide from villin 3 (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence……..178

Figure A-29. Tandem mass spectra of peptide from photosystem II protein D1 (*Helianthus annuus*) as indicated (bold) in the primary amino acid sequence……...…180

Figure A-30. Tandem mass spectra of peptides from Cytokinin-O-glucosyltransferase (*Medicago truncatula*) as indicated (bold) in the primary amino acid sequence……...…182

Figure A-31. Tandem mass spectra of peptides from ABA/WDS stress induced (*Orytza sativa*) as indicated (bold) in the primary amino acid sequence……..184

Figure A-32. Tandem mass spectra of peptides from NAC Domain protein (*Orytza sativa*) as indicated (bold) in the primary amino acid sequence……..186

Figure A-33. Tandem mass spectra of peptides from a similar to 1-aminocyclopropane-1-carboxylic acid oxidase (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence……187

Figure A-34. Tandem mass spectra of peptide from 1-aminocyclopropane-1-carboxylic acid oxidase (Rosa hybrid cultivar) as indicated (bold) in the primary amino acid sequence……189

Figure A-35. Tandem mass spectra of polypeptide from Jacalin-like protein LEM2 (*Hordeum vulgare*) as indicated (bold) in the primary amino acid sequence……..190

Figure A-36. Tandem mass spectra of polypeptide from hevein-like antimicrobial protein (*Euonymus europaeus*) as indicated (bold) in the primary amino acid sequence……...191

Figure A-37. Tandem mass spectra of polypeptide from HSP 70 protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence……..192

Figure A-38. Tandem mass spectra of polypeptide from metal ion binding protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence...196

Appendix B

Figure B-1. Chromatographic apparatus designed for whole protein separation……198

Figure B-2. Column switch diagram of offline cap-trap enrichment and online sample introduction……199

Figure B-3. Chromatographic apparatus designed for small molecule/peptide separation with post column split……200

Figure B-4. UV spectrum analysis of CHY and CHY-TPCK samples enzymatic activity by reaction with BTEE observed at 256 nm……...201

CHAPTER I

LC ESI ION TRAP MS DETERMINATION OF ALTERED PROTEIN EXPRESSION IN *HELIANTHUS ANNUUS* AFTER EXPOSURE TO HEAVY METALS As, Cd, Cr, Ni AND Pb

1.1. Introduction

Within the diverse realm of Bioanalytical Chemistry one of the most recently blossoming fields of research is Proteomics, the study of proteins, their structure, function, alterations, expression and their general role in biochemistry. The recent explosion of research ventures in the field of Proteomics over the last few decades is due largely to the evolution of modern technology that has expanded the range of focus and depth of information that can be obtained from a biological sample. The researcher now has access to a vast resource of documented and developing techniques for studying, manipulating and directing a biological sample's molecular environment, interactions, structure, and expression profiles. Extensive databases of knowledge are available for each method described and the results attained by altering gene sequences and protein expression. Coupled to all this is the exponentially expanding field of Bioinformatics, online databases that contain a comprehensive listing of nearly all published or expressed protein structures and DNA sequences.

Proteomics was defined in 1838 by Berzelius when he put forth the name "protein" (from the Greek proteios, "primary") for the complex organic nitrogen-rich substance found in the cells of all living organisms. Over the course of the nineteenth century researchers identified many of the 20 common amino acids that make up proteins. In 1864 the first protein, hemoglobin, was crystallized by Hoppe-Seyler. In 1894 a lock-and-key analogy for enzyme-substrate interactions was proposed by Fischer. Three years later, in 1897, Buchner and Buchner created enzymology by demonstrating that cell-free extracts of yeast ferment sucrose to form carbon dioxide and ethanol. In 1926 Sumner crystallized urease in pure form, demonstrating that proteins could possess the catalytic

activity of enzymes. At the same time Svedberg developed the first analytical ultracentrifuge and used it to estimate the correct molecular weight of hemoglobin. In 1933 Tiselius introduced electrophoresis for separating proteins in solution. One year later, in 1934 Bernal and Crowfoot presented the first detailed x-ray diffraction patterns of a protein, obtained from crystals of the enzyme pepsin. In 1942 Martin and Synge developed chromatography, a technique now widely used to separate proteins. In 1951 Pauling and Corey proposed the structure of a helical conformation of a chain of L-amino acids, the α helix, and the structure of the β sheet, both of which were later found in many proteins. In 1955 Sanger completed the first analysis of the amino acid sequence of a protein, insulin. In 1956 Ingram produced the first protein fingerprints, showing that the difference between sickle-cell hemoglobin and normal hemoglobin is due to a change in a single amino acid. In 1960 Kendrew described the first detailed structure of a protein (sperm whale myoglobin) to a resolution of 0.2 nm, and *Perutz* proposed a lower-resolution structure for hemoglobin. In 1963 Monod, Jacob, and Changeux recognized that many enzymes are regulated through allosteric changes in their conformation. This body of initial experimental work set the foundation for a wide range of scientific endeavors in the realm of Proteomics, limited only by the inventiveness of the researchers and the availability of precision instrumentation.

Since the 1960's a vast amount of research has been performed on proteins. Extensive data on the sequences, functions, expression patterns, and folded structures of proteins have been analyzed, documented and entered into globally accessible online databases. The largest hurdle the biochemical researcher has had to overcome has been to develop the means to more thoroughly analyze the *complex* biological sample. The

early experiments described above involved incredibly lengthy and time-consuming isolation and purification steps. Sample size had to be very large since instrument sensitivity was very poor. As Proteomics developed as a field Chromatography and Electrophoresis came to the forefront of useful analytical methodologies for the separation of the complex sample for further analysis. These techniques allow mixtures of proteins and peptides to be separated into their individual components according to their inherent characteristics; size, iso-electric point, electronic interaction, hydrophobicity, and relative pH.

The 1970's saw the development of many reliable chromatographic methods for the biochemical researcher. Pressure liquid chromatography was designed to decrease flow through time and reduce the purification times for compounds being isolated. High-pressure liquid chromatography (HPLC) was the next development, which led to the development of on-line detectors and new and improved column packing materials. Amongst these new packing materials were hydrophobic carbon chain based molecules, termed reverse phase liquid chromatography, these facilitated improved separation between very similar biological compounds. Continued method development led to improvements in separation, identification, purification and quantification far above the previous techniques. The incorporation of computers and automation in the 1980's added to the convenience and reproducibility of HPLC.

Several branches of HPLC have been developed that take advantage of specific properties by which biomolecules can be separated. Ion exchange chromatography separates molecules based on how they interact with charged groups immobilized on the

column matrix. The binding affinity of a given peptide is a result of its net charge, determined by the pH of the solution, and the presence of other ions in the sample that can compete for binding with the ion exchanger. A strong ion exchange column is ideal for peptide separation due to the wide variety of charged residues and the ability to change their charge potential by altering the pH of the mobile phase [1].

Gel filtration chromatography or size exclusion chromatography utilizes a column packing consisting of various gel beads created with pores manufactured to a preset size. As larger molecules pass around these beads they elute from the column faster than smaller molecules that are capable of interacting or passing thru the pores [1].

Affinity chromatography is designed to utilize the ability of proteins to specifically bind to a molecule or ligand in a non-covalent fashion. The column packing matrix is created to contain a protein specific ligand so that when a complex mixture is passed through the chromatographic material, the protein of interest binds to the packing and all other molecules are eluted. In this technique the isolated protein can then be eluted by changing the buffer conditions to release it from the matrix [1].

Of primary interest to our research group is reversed phase chromatography, the separation of biomolecules based upon their hydrophobic interactions with a non-polar packing material. Functional groups interact or associate with the hydrophobic groups on the column backbone. Separation is accomplished by changing the mobile phases over a gradient to vary the relative polarity. [1]

Once the biochemical researcher has separated the complex mixture then he/she has to develop a method to identify the analytes obtained. Initially this identification step was also very time intensive, it involved obtaining structural information from a variety of spectroscopic analyses and performing chemical reactions to determine the sequence information of the protein. The need for a more thorough analyte identification methodology was addressed in two distinct ways. The first involved the development of X-ray crystallography in the 1930s and nuclear magnetic resonance (NMR) spectroscopy in 1945. These methods allowed a new means of determining analyte structure and sequence but carried some limitations as both methods required high purity, relatively large sample size and the ability of the molecule to form a crystal. The development of computer technology and advanced algorithm programming provided techniques which today are extremely fundamental in understanding molecular conformation and structure, but they are an "offline" step that still depends on the somewhat difficult step of the creation of effective crystals. The second method developed for analyte identification also premiered in the 1930s, and was coined the "Mass Spectrometer".

The first clear demonstrations of mass spectrometry came at the end of World War I by Francis W. Aston and Arthur J. Dempster, in Cambridge and Chicago, respectively. Dempster used a magnetic analyzer that focused ions into an electrical collector, while Aston used both electrostatic and magnetic fields to focus ions on a photographic plate. Their work continued, along with that of Joseph H. E. Mattauch, R. F. K. Herzog, Kenneth T. Bainbridge, and Alfred O. C. Nier. Nier then incorporated developments in vacuum technologies and electronics for power supplies and ion detection, significantly improving magnetic focusing instruments. Through this technique he established that it

was not necessary to send ions along their entire path of a semicircle within an analyzer completely confined between the poles of a massive magnet (prior standard practice). Instead good results could be obtained by sending ions through a comparatively modest wedge-shaped sector magnetic field. More practical electron-bombardment ion sources, along with several other crucial aspects of construction and technique, also brought improvements in performance, convenience, and costs. Double-focusing machines were also greatly refined, attaining greater precision by adding an electrostatic analyzer. Analyzers based on the different times of flight over a set path for different accelerated ions, and others that used various types and combinations of fields also had some success.

In 1953 Wolfgang Paul and his colleagues initiated development of the quadrupole mass analyzer (and ion trap), using crossed radio frequency and electrostatic fields. The middle and later 1950s saw the first wave of efforts to couple the mass spectrometer with other significant types of instruments, creating, in this case various types of gas chromatograph mass spectrometers. In the 1960s, chemical ionization mass spectrometry (or CIMS) and field desorption MS (FDMS) emerged. In the 1970s secondary ionization MS, Fourier transform MS, plasma desorption MS, electrohydrodynamic MS, laser desorption MS, thermal desorption MS, spark source MS, and glow discharge MS were invented or developed significantly. The 1980s saw the development of resonance ionization MS, matrix-assisted laser desorption MS, fast atom bombardment MS and its continuous flow transformation. Following this was the development of ion trap MS, a dramatic advance in electrospray MS, and the subsequent development of modern liquid chromatography MS.

Our current applications of two dimensional liquid chromatography coupled with mass spectrometry involves the complete separation and subsequent identification of complex mixtures of biomolecules. In theory, samples containing 1000 or more proteins can be digested and still be resolved to complete separation and identification, limited only by the specificity of the equipment, the skills of the researcher and the extensiveness of the bioinformatics library. Virtually any sample can be analyzed with the diverse array of chromatographic techniques and mass spectrometers available to the researcher. The versatility of this method stems from coupling multiple stages of chromatographic separation to isolate many smaller fractions, from the original complex mixture, that can be more thoroughly analyzed by the multiple MS. A researcher has available the entire host of available column / gradient separation techniques of liquid chromatography. This methodology is most often coupled with digestion of the protein sample, an enzymatic process of reducing the protein into many smaller peptide fragments. These peptide fragments can then be identified using the mass spectrometer and characterized completely with tandem MS evaluations.

Protein digestion is a common technique used to evaluate the amino acid composition and to determine structural information of proteins. Information for hundreds of thousands of proteins exist within the bio-informatics databases, including information on predicted peptide cleavage patterns from varying enzymatic digests. Among the most common enzymes utilized are the Serine Peptidases, a class of digestive enzymes whose active site is composed of a catalytic triad, a unique arrangement of an aspartic acid, a histidine and a serine residue that are highly selective to specific peptide bonds due to the proteins folded structure. Located next to this catalytic triad is a hydrophobic pocket of

varying characteristics for each enzyme that alters the specificity of the cleavage patterns. Trypsin is an enzyme specific for cleaving the peptide bond on the carbonyl side of the basic residues arginine and lysine. Chymotrypsin is an enzyme specific for cleaving the peptide bond on the carbonyl side of the bulky aromatic side chains phenylalanine, tryptophan and tyrosine. Elastase is an enzyme specific for small neutral residues, particularly alanine. These enzymes used sequentially for multiple digest and sequence determination experiments on a sample can allow for detailed secondary protein structure to be determined.

The post-digest peptide fragments obtained couple excellently with reverse phase liquid chromatography. The peptide fragments possess all the properties of their constituent amino acid residue side chains in addition to their c and n-termini. The protein digest is enriched and loaded on a hydrophobic captrap, to concentrate and remove buffer contaminants by reverse phase microcolumns in order to increase signal-noise ratios and sensitivity before MS analysis. This does allow for the loss of small and hydrophyllic peptides but more importantly will wash the sample of any residues or compounds that can be damaging or interfere with the sensitive ion trap. After washing in the trap the peptides are then eluted and passed through a C-18 reversed phase column allowing separation by hydrophobic interaction. The end result is a potential reduction in the number of peptides passing through the mass spectrometer by some three orders of magnitude from that of the original complex. The coupling of the Reversed Phase chromatography step with the mass spectrometer requires a means of ionizing the peptide fragments. The mobile phase from the chromatography step is most commonly a mixture

of water and acetonitrile or water and methanol, both of which couple excellently with the electro-spray ionization technique.

Electro-Spray Ionization (ESI) is a technique that facilitates the transfer of ions from the liquid phase to the gas phase. In ESI, charged droplets are produced from electrolytes dissolved in the solvent and nebulized by use of a gas stream upon entry to the MS. The charged droplets are then dried and reduced in size by solvent evaporation and repeated splitting, eventually creating very small highly charged droplets capable of producing gas phase ions. This ionization technique allows the formation of analyte ions in the gas phase and analysis of the charged molecules by their interaction with the mass spectrometers detector [2]. An Ion Trap MS can allow for multiple tandem MS analysis, currently up to 12+ times daughter analysis of a single biomolecule. 2-D LC separation with MS^n ion trap analysis can in theory allow tens of thousands of proteins to be simultaneously analyzed, and this couples with high throughput, an excellent dynamic range (able to analyze 10^6:1 relative protein abundance) and a very flexible methodology with variance available in LC conditions and columns, as well as a variety of enzymatic digests available.

Within the field of Proteomics extensive work has been undertaken using the technique of multidimensional liquid chromatography coupled mass spectrometry (LC-MS). LC-MS techniques have been combined with 1 dimensional and 2 dimensional electrophoretic separation for direct and systematic identification of expressed proteins from prepared biological media. Proteins resolved by 2D gel electrophoresis can be enzymatically digested in–gel, or digested during blotting onto membranes containing immobilized trypsin. Trypsin cleaves after the basic residues arginine and lysine, if not

followed by proline along the polypeptide chain. Each of the tryptically produced peptides will therefore have a basic arginine or lysine c-terminus residue in addition to its n-terminus, both of which are capable of retaining a positive charge. This digest method couples perfectly with ESI ion trap analysis since each peptide created can retain at minimum a 2+ charged state allowing for in depth fragment detection in MS/MS ion trap analysis providing in depth structural definition.

Phytoremediation, the use of plants to restore soil and water contaminated with pollutants, is considered a promising approach due to its low cost and the minimal impact the procedure has on the environment. Interest in the use of plants for environmental restoration has grown following the discovery that some plants are capable of accumulating toxic metals at levels many times greater than common nonaccumulator plants. These plants are described as hyperaccumulators or hyperacumulating plants. There are over 400 plant species that are considered metal hyperaccumulators of one or more toxic metals [3]. *Helianthus annuus* (Dwarf Sunflower) is capable of hyperaccumulating several metals including Ni, Cd, Cu, and Cr [4]. As metal accumulation is of serious interest and since accumulation of metals is toxic to plants, hyperaccumulating plants must have developed mechanisms of tolerating metal accumulation. Mechanisms underlying hyperaccumulation of toxic metals are complex. One major mechanism might be to counteract the oxidative stress initiated by heavy metals by induction of antioxidative enzymes [5]. Another defense mechanism appears to involve induction of metal binding proteins, e.g., phytochelatins (PC) and metallothioneins (MT) [6].

These small sulfur-rich peptides have been the focus of recent studies. For instance, he synthesis of PC has been shown to be induced by a variety of different metals [7,8] Rendering the plants tolerant to accumulation of such metals as copper and cadmium [9,10]. On the other hand, some proteins that are expressed in response to the metal stress are not considered as metal chelators. Notable examples are chitinases and heat shock proteins [11,12], although the significance of their expression in the metal hyperaccumulation is not understood.

Proteomics has recently been applied to study the regulation of proteins as related to metal stress in plant species. To date few studies have observed the specific protein expression in *H. annuus* or any hyperaccumulator plants to specific metal stress [11,13]. The purpose of the present study is to identify and characterize proteins that are observed in increased spot intensity in response to four metals (Cd, Cr, Ni, and Pb) that hyperaccumulate in *H. annuus*. The proteins may not play a direct role in the transport or sequestration of toxic metals but could be involved in mediating metal tolerance in other ways. Recent studies of metal accumulation have induced oxidative stress responses in plants [14] in addition to a variety of other observed proteins [15-18]. These responses leave open the question of what possible response stress mechanisms are being initiated on metal accumulation. Many of the observed proteins in these studies are not considered to be metal binding, although their biological function may in some way contribute to the ability of the plants to deal with metal stress. Many of these studies leave unresolved questions as to the detailed proteomic effects of metal exposure [17,18]. This study aims to further expand some of the proteomic findings of the exposure to heavy metals on a known hyper-accumulator species, *H. annuus*.

1.2. Methodology

1.2.1. Plant Growth and Exposure

Sunflowers were exposed to heavy metals in solutions containing single metals (As, Cd, Cr, Ni or Pb) and combinations of Cd, Cr and Ni, with or without As or Pb. Each heavy metal was added at 30 mg/L. The 30 mg/L refers to the concentration of the individual metal, not the compound added. The metals were applied as As^{5-}, ($Na_2HAsO_4 \cdot 7H_2O$), Cd^{2+} ($CdSO_4 \cdot 8H_2O$), Cr^{3+} ($CrCl_3 \cdot 6H_2O$), Ni^{2+} ($NiSO_4 \cdot 6H_2O$), and Pb^{2+} ($Pb(NO_3)_2$.

Experiments were conducted with two different *Helianthus annuus* cultivars (Sundance and Teddy Bear). Seeds of each cultivar were grown hydroponically in Rockwool, a non-reactive, nonabsorbent support system and sustained with a standard hydroponic nutrient solution (Cropking Corp, Seville, OH). The seeds were grown in a greenhouse illuminated with natural light. The average temperature of the greenhouse was 28^{o}C (winter) and 35^{o}C (summer) in the day and 20^{o}C at night. After a four week growth period, seven sunflowers were transferred to a PVC trough. One set of plants was then subjected to a 4 L solution containing 30 mg/L of the toxins as described above. Solutions were introduced at a complete recycle rate of 6 gallons per hour. After 17 days of exposure, the plants were harvested and the leaves were then sectioned and weighed. The tissues were used for analysis of metal levels, and for protein and RNA isolation. Plant growth and exposure was conducted by Dr. Teresa Cutright at the University of Akron, Department of Civil Engineering [11].

1.2.2. Protein Extraction and Gel Separation

The plant tissues were mashed with mortar and pestle in liquid nitrogen, and 100-200 mg and were placed in extraction buffer. The plant extraction solution Focus^{TM} Plant Proteome (G-Bioscience, St. Louis, MO) was used following the manufacturer's instructions, then centrifuged for 20 min at 15,000 x g. The supernatants were collected and the protein concentrations were measured using the Bradford method with albumin as the reference protein solution. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed to establish distribution of proteins, molecular mass of the proteins present in the control and the metal-exposed plant biomass. Later experiments first separated proteins by isoelectric focusing prior to electrophoretic separation. Gels were stained with Coomassie Brilliant Blue and protein bands (1 dimensional) or spots (2 dimensional) of interest gels were excised for further digestion and analysis. Protein extraction and gel separation was performed in Dr Harry vanKeulen's laboratory by Chamari Hasintha, Cleveland State University [11].

1.2.3. Sample Preparation and Tryptic Digest

Trypsin was obtained from Promega or Pierce. Protein digestion was performed on gel bands (1-D) and gel spots (2-D) of interest that were excised, cut into approximately 1mm^3 pieces and stored at -20° C until digestion. In preparation for digestion gel pieces were washed/destained with 200 µL wash reagent (50% Et-OH / 5% acetic acid / 45% H_2O), for 1 h. The wash was discarded and the wash step repeated for an additional hour. Samples were then reduced by adding 30 µL 100 mM dithiothreitol and incubating

the sample at 60° C for 10 min, reducing buffer was then discarded. Samples were then alkylated by adding 30 µL 100 mM iodoacetamide and incubating in the dark at room temperature for 1 h, alkylation buffer was then discarded. Gel pieces were washed with 200 µL wash reagent for 1 h. The wash was discarded and the wash step repeated for an additional hour. Gels were then dehydrated by adding 200 µL acetonitrile for 10 min. The acetonitrile was discarded and samples were rehydrated in 200 µL 100 mM ammonium bicarbonate for 5 min. Then the ammonium bicarbonate was discarded and samples were again dehydrated with 200 µL acetonitrile for 10 min. Next, the acetonitrile was discarded and the samples were dried in SpeedVac for approximately 3-5 min. Finally, samples were rehydrated in 50 µL trypsin (20 ng/µL in 50 mM ammonium bicarbonate) for 10 min on ice. Excess Trypsin solution was then removed and 5 µL of 50 mM ammonium was added to submerge the gel pieces. Samples were incubated at 37° C for 4 h or alternatively held at room temperature (23-25° C) overnight.

After the digestion step was completed, samples were microcentrifuged and the supernatant digest mixture collected. 30 µL extraction reagent (50% ACN / 5% formic acid / 45% water) was added to the remaining gel pieces. The sample was briefly vortexed and left for 10 min. The extraction mixture was removed and added to the digest mixture, the extraction step was repeated once. This additional extraction step can slightly increase digest yield and also serves to terminate any continuing trypsin activity. Combined supernatants were placed in a SpeedVac to reduce volume to approximately 10 µL, about 45 – 60 min. Sample volumes were reconstituted to approximately 25 µL total volume with 1% acetic acid. Samples were now ready for LC-MS analysis.

HPLC ESI-Ion Trap MS was carried out by injecting the sample offline into a secondary LC loop using a loading phase of 2% ACN in H_2O with 0.2 % HCOOH, the molecules were enriched on a C-18 CapTrap (MicromResources). The sample was washed for 10 minutes in the CapTrap and then the main column switch brought the CapTrap online into the primary Agilent LC mobile phase ramp as described below. Ion Trap conditions were set initially to those determined for small molecule analysis and optimized as needed.

1.2.4. LC Separation

HPLC analysis was carried out on an Agilent 1100 series pump (Agilent, Palo Alto, CA). Samples were loaded offline in 10 μl aliquots onto a Peptide C-18 CAPTRAP (Michrom Bioresources, Auburn, CA) for enrichment and washing and were subsequently separated on a reverse-phase C-18 MS column (100 x 0.3 mm, 5 μm, 300 Å; Vydac, Hesperia, CA). Initial Peptide Cap Trap loading was performed using the following conditions: 10 μl/min using a 2% acetonitrile (ACN), 1% formic acid solvent mixture in water for 10 min on a LV-10 ADVP pump (Shimadzu, Colombia, MD). At 10 min loading solvent flow switched to primary LC flow. Peptides were resolved using the following conditions: flow rate of 5 μl/min, linear gradient from 5 - 40% buffer B (0.1% formic acid, 99.9 % ACN) in solvent A (0.1 % ACN, 0.1 % formic acid in water) for 70 min, then 40 - 80% B for 30 min.

1.2.5. MS Analysis

The samples were analyzed using an EsquireHCT Brucker Daltonics mass spectrometer equipped with a standard Agilent electrospray (ESI) ion source (Brucker Daltonics, Billerica, MA). The instrument settings were as follows: nebulizer-11.0 psi, flow rate of dry gas 5.0 Liter/min; temperature 310^{o}C; ESI voltage 3.0 kV; scan mode-standard-enhanced with range 50 - 3000 m/z, speed 8, 100 m/z/s; trap settings-positive polarity, ICC Smart Target-105; maximum accumulation time 2 x 10^4 ms, scan-350-1700 m/z; and target mass-800 m/z. MS/MS data were acquired using an Auto MSn feature that functions in the "Triple Play' type data-dependent mode, a total scan that included three scan events: a full range MS scan; a narrow-range, high resolution zoom scan on a selected ion from the MS scan; and a MS/MS scan of the selected ion.

1.2.6. MASCOT Bioinformatics

The data were searched against NCBI using the MASCOT soft ware. Additional BLAST search algorithms were exhaustively performed against the NCBI database using individual peptide strings and whole protein sequences for potential matches.

Two sets of Mascot MS/MS ion search criteria were applied to both the Veridiplantae and the Arabidopsis thaliana taxonomy filters. The initial search parameters for trypsin digestion allowing one miscleavage, variable modification—oxidation of methionine, MS/MS tolerance of 0.5 Da, and an ion score cut-off of 20. A second intentionally broad search was applied, allowing a wider range of possible peptide matches from many species to compliment the somewhat sparse *H. annuus* data base, results from this search are excluded from the tables.

1.3. Results

1.3.1. Expression of proteins in Sundance

1.3.1.1. Expression of proteins in response to arsenic and other metals in Sundance

Hydroponic Sundance root samples 1-D PAGE analysis revealed three distinct spots around markers at 30 and 66 kDa (Fig. 1-1). The mass spectra of tryptic digests of these spots indicated a positive match to two proteins in the first spot, and one protein in each of the other spots (Table 1-I). The first peptide from the first spot matched the peptide YGGMVMoIWDR, Mass 1095.52 Da (549.27, 2+) to Chitinase *H. annuus* a 32 kDa protein (Fig. 1-2). MS/MS fragmentation yielded twelve fragment ions of the nineteen most intense peaks matching the expected masses of b_2, b_3, b_4, b_6, b_7, y_1, y_3, y_4, y_5, y_6, y_7, y_8. The second peptide from the first spot matched the peptide LENGQSWQITVAPGTAQAR, mass 2026.02 Da (1014.05, 2+), (676.37, 3+) to an unknown protein sharing close homology to pathogenesis-related protein 5-1 *H. annuus* a 25 kDa protein (Fig. 1-3). MS/MS fragmentation of the 2+ charged peptide yielded thirteen fragment ions of the fourteen most intense peaks matching the expected masses of b_{11}, b_{11}^{0}, b_{12}, b_{18}^{*}, y_7, y_8, y_9, y_{10}, y_{11}, y_{12}, y_{12}^{0}, y_{13}^{++}, y_{14}. MS/MS fragmentation of the 3+ charged peptide yielded forty-one fragment ions of the forty-four most intense peaks matching the expected masses of b_5, b_6, b_6^{*}, b_6^{0}, b_8^{0}, b_9, b_{10}^{*}, b_{10}^{*++}, b_{10}^{0++}, b_{11}, b_{11}^{++}, b_{11}^{*++}, b_{11}^{0}, b_{11}^{0++}, b_{12}, b_{12}^{++}, b_{11}^{*++}, b_{12}^{0++}, b_{13}^{*++}, b_{15}^{++}, b_{16}^{++}, b_{16}^{*++}, b_{16}^{0++}, b_{18}^{++}, y_2, y_6, y_7, y_7^{*}, y_8, y_8^{++}, y_8^{0}, y_9, y_9^{*}, y_9^{0}, y_{11}^{++}, y_{13}^{++}, y_{14}^{++}, y_{16}^{*++}, y_{16}^{0++}, y_{17}^{*++}, y_{17}^{0++}. The second spot, below the 66 kDa marker, matched the peptide DEKPEGTSR, mass 1016.53 Da (340.06, 3+) to CC-NBS-LRR RGA protein, *H. annuus*, a 19 kDa protein (Fig. A-1). MS/MS fragmentation yielded seven fragment ions of the nine most

intense peaks matching the expected masses of b_4*, b_4^0, y_3++, y_4++, y_5++, y_6++, y_8++. The third spot, around the 66 kDa marker, matched multiple peptides VVGFFGR, mass 780.43 (391.21, 2+), QTYGPYGTNGGTDFSCPIAK, mass 2075.93 (692.93, 3+), NGSLSTIDIADDEEIIEINGK , mass 2245.10 (1123.59, 2+) corresponding to a 15 kDa subunit of Lectin *Helianthus tuberosus* (Fig. A-2). MS/MS fragmentation of VVGFFGR yielded five fragment ions of the six most intense peaks matching expected masses of b_2, y_3, y_4, y_5, y_6. MS/MS fragmentation of QTYGPYGTNGGTDFSCPIAK yielded seventeen fragment ions of the thirty-five most intense peaks matching expected masses of b_4^{++}, b_9^*, b_{10}^{++}, b_{13}^{++}, b_{15}^{*++}, b_{15}^{0++}, b_{18}^{*++}, b_{18}^{0++}, b_{19}^{++}, y_6, y_7, y_8, y_{10}^{++}, y_{11}, y_{13}^{++}, y_{15}^{++}, y_{19}^{0++}. MS/MS fragmentation of NGSLSTIDIADDEEIIEINGK yielded twenty-six fragment ions of the twenty-two most intense peaks matching expected masses of b_6, b_6^0, b_7, b_9, b_{10}^0, b_{15}, b_{15}^0, b_{16}, b_{17}, b_{18}, y_3, y_4, y_5, y_6, y_7, y_9, y_{10}, y_{11}, y_{12}, y_{12}^{++}, y_{13}, y_{14}, y_{14}^0, y_{15}, y_{17}.

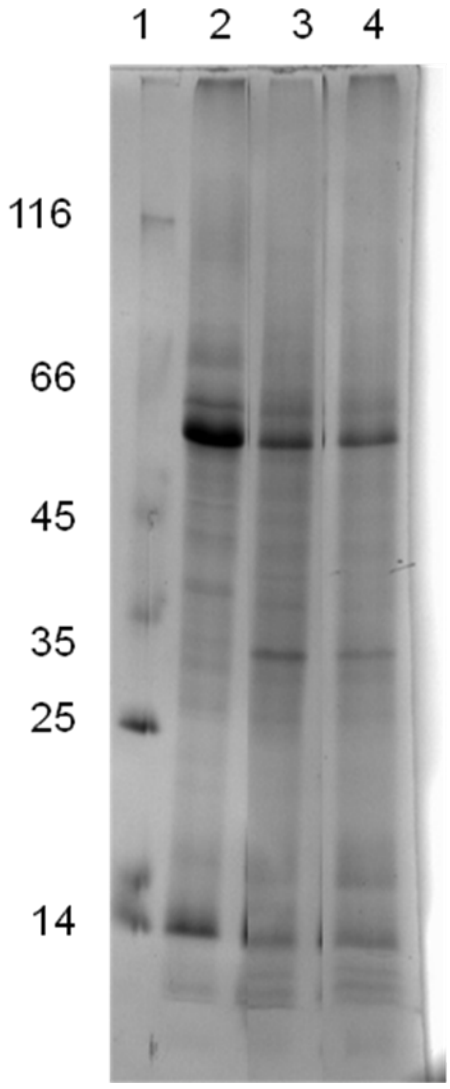
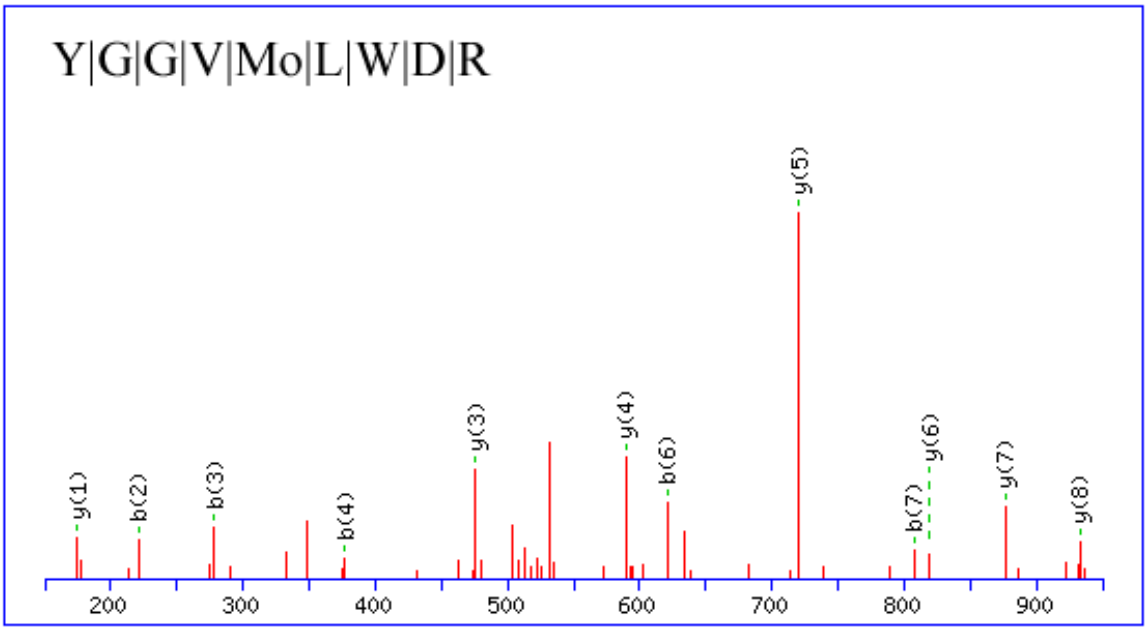


Figure 1-1. 1-D Gel Separation Sundance Hydroponic As [11]

Table 1-I. 1-D Hydroponic Sundance Root As, Cd, Cr, Ni

Marker Region kDa	Theoretical Mass kDa	Accession no.	sp.	Protein Name	Score	Peptide (Coverage)	Peptides
30	32	gi|126723930	*H. annuus*	Chitinase	72	1 (2%)	YGGVMoLWDR
30	25	gi|13938981	*H. annuus*	Unnamed protein	64	1(9%)	LENGQSWQITVAPGTAQAR
60	19	gi|49425431	*H. annuus*	CC-NBS-LRR RGA	31	1 (3%)	DEKPKEGTSR
66	15*	gi|11559262	*H. tuberosus*	Lectin	211	3 (18%)	VVGFFGR, QTYGPYGTNGGTDFSCPIAK, NGSLSTIDIADDEEIIEINGK



(A)

MEFTHPALLLLLFITVFSFLKPSTAAGIATYWGQQSDDTEGTLEAACATGNYQFV

NIAFLSTFGNNQQPVLNLAAHCDPASTCSRYSSQIKACQAQNVKVFLSIGGQRGS

YSLSSPQDAQQVADFLWNTYLGGQPATRPLGDAVLDGIDFDIEQGTDQFWSDLA

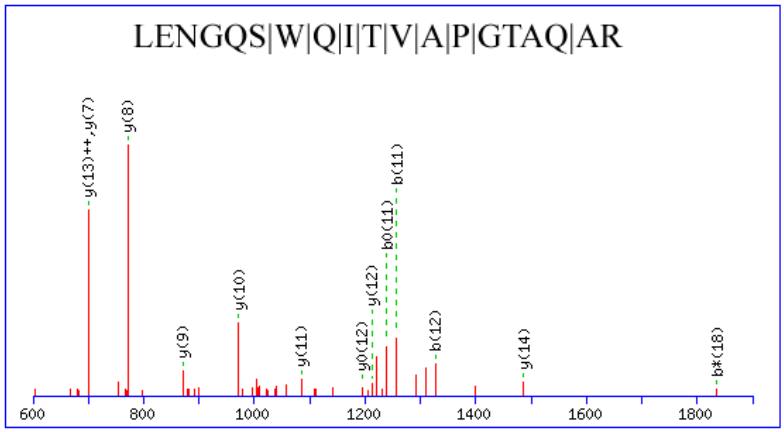
KAPAAYSSQKKVYLSAAPQCPFPSGDVRNQLLPAIREGLFDYVWVQFYNNEQCQ

YGANADALLARWNEWTQVTTNTIFLGLPAAASGAAPSGGYIPPDILTSQILPSIKS

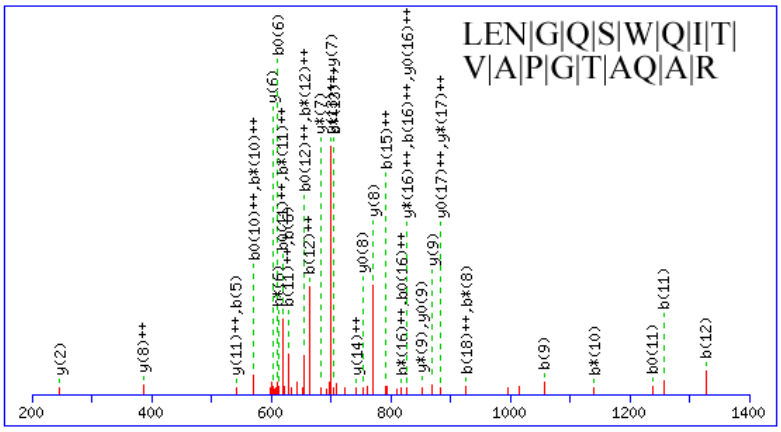
SPK**YGGVMLWDR**FYDKQSGYSDAIKGSIN

(B)

Figure 1-2. Tandem mass spectra of peptide (A) from Chitinase (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).



(A)



(B)

MTCAKNLLLSITLLSIGCFTLTRGATFDVINQCQYPVWAAWASTTPGGGKR**LEN**

GQSWQITVAPGTAQARIWGRTGCNFDANGRGRCDTGDCNGMLECQGYGAPP

NTLAEFALNQDNNNDFVDISLVDGFNIPMEFSPVGASCKTMRCAGNLNGECPNE

LRTQGGCNNPCTVYKTNEYCCTNGPGSCGPTPLSRFFKDRCPDAYSYPQDDPTSL

FTCPGGTNYKVVFCP

(C)

Figure 1-3. Tandem mass spectra of +2 (A) and +3 (B) peptides from an unknown protein type with high sequence homology to pathogenesis-related protein 5-1 (*H. annuus*) as indicated (bold) in the primary amino acid sequence (C).

1.3.1.2. Expression of proteins in response to arsenic only in Sundance

Hydroponic Sundance leaf samples 1-D PAGE analysis revealed one very distinct spot around the marker region at 45 kDa. The mass spectra of the tryptic digest of this spot indicated positive matches to three proteins (Table A-I). The first set of peptides, HQGVMVGMGQK mass 1170.56 (391.22, 3+), AVFPSIVGRPR mass 1197.70 (400.28, 3+), SYELPDGQVITIGNER mass 1789.88 (895.93, 2+), VAPEEHPILLTEAPLNPK mass 1967.07 (656.81, 3+), match to Actin, *Mesostigma viride*, a 42 kDa protein with high homology to Actin, *H. annuus* (Fig. A-3). MS/MS fragmentation of HQGVMVGMGQK yielded five fragment ions of the fifteen most intense peaks matching the expected masses of y_3^{++}, y_4^{++}, y_5^{++}, y_7^{++}, y_7^{*++}. MS/MS fragmentation of AVFPSIVGRPR yielded four fragment ions of the five most intense peaks matching the expected masses of b_2, b_3, y_8^{++}, y_9^{++}. MS/MS fragmentation of SYELPDGQVITIGNER yielded twenty-six fragment ions of the thirty-four most intense peaks matching the expected masses of b_2, b_3, b_4, b_4^0, b_8^0, b_9, b_9^0, b_{10}, b_{10}^0, b_{12}, b_{12}^{*++}, b_{12}^0, b_{12}^{0++}, y_4, y_5, y_6, y_7, y_9, y_{10}, y_{11}, y_{12}, y_{12}^{++}, y_{12}^0, y_{13}, y_{13}^{++}, y_{14}^{++}. MS/MS fragmentation of VAPEEHPILLTEAPLNPK yielded seventeen fragment ions of the thirty-one most intense peaks matching the expected masses of b_8^0, b_9, b_{10}, b_{10}^{++}, b_{11}^{++}, b_{12}^{++}, b_{13}^{++}, b_{16}^{*++}, y_3, y_5, y_5^* y_6, y_7, y_8, y_9, y_{12}, y_{16}^{++}. The second set of peptides, HIQAGASK mass 810.43 (271.35, 3+), AVSLVLPQLK mass 1066.67 (534.35, 2+), match to glyceraldehyde 3-phosphate dehydrogenase B subunit, *Arabidopsis*, a 44 kDa protein (Fig. A-4). MS/MS fragmentation of HIQAGASK yielded seven fragment ions of the fifteen most intense peaks matching the expected masses of b_2, b_2^{++}, b_4^{++}, b_4^{*++}, y_2^{++}, y_2^*, y_5^{++}. MS/MS fragmentation of AVSLVLPQLK yielded nine fragment ions of the

nine most intense peaks matching the expected masses of b_2, b_4^0, b_5, b_6, y_4, y_5, y_6, y_7, y_8. The final peptide, IAIQAGAAGIIVSNHGAR mass 1717.96 (573.81), matched glycolate oxidase, *H. annuus*, a 40 kDa protein (Fig. A-5). MS/MS fragmentation of IAIQAGAAGIIVSNHGAR yielded eighteen fragment ions of the thirty-nine most intense peaks matching the expected masses of b_3, b_6^*, b_6^{*++}, b_{13}^{++}, y_5^*, y_7, y_7^{0++}, y_9^{++}, y_{10}, y_{10}^{++}, y_{12}^{++}, y_{13}^{++}, y_{14}^{++}, y_{14}^{*++}, y_{14}^{0++}, y_{15}^{++}, y_{15}^{*++}, y_{16}^{++}.

1.3.1.3. Expression of proteins in response to lead and other metals in Sundance

Hydroponic Sundance leaf samples 1-D PAGE analysis revealed two distinct spots, one around the marker region at 37 kDa and the other just below it. The mass spectra of tryptic digest of these spots indicated positive matches to two and six proteins respectively (Table A-II). The first peptide from the first spot, matched the peptide YGGMVMoIWDR, mass 1095.52 Da (548.75, 2+) to Chitinase *H. annuus* a 32 kDa protein (Fig. A-6). MS/MS fragmentation yielded twelve fragment ions of the nineteen most intense peaks matching the expected masses of b_2, b_3, b_4, b_6, y_1, y_3, y_4, y_5, y_6, y_7, y_8. The second peptide from the first spot, matched the peptide GGSTGYDNAVALPAGGR, mass 1561.75 (781.43, 2+), to a 33 kDa oxygen evolving protein *Arabidopsis* (Fig. A-7). MS/MS fragmentation yielded twenty-two fragment ions of the thirty-eight most intense peaks matching the expected masses of b_5, b_7, b_9^0, b_{10}^0, b_{11}, b_{11}^0, b_{12}, b_{12}^0, b_{13}, b_{14}, b_{14}^{++}, b_{15}^{++}, y_4, y_5, y_5^*, y_6, y_7, y_8, y_9, y_{10}, y_{11}, y_{12}, y_{13}. The first peptide from the second spot, matched VINTWQDIINR, mass 1313.71 (657.93, 2+), to photosystem II protein D1 *H. annuus*, a 38 kDa protein (Fig. A-8). MS/MS fragmentation yielded seven fragment ions of the eleven most intense peaks matching the

expected masses y_3, y_3^*, y_4, y_5, y_6, y_8, y_{10}. The second peptides from the second spot, matched DEFIGKK, mass 835.44 (418.04 2+), VASPQQAQEVHVAVR, mass 1532.81 (511.84, 3+), IIYGGSVNGGNSAELAK, mass 1648.84 (825.33, 2+), to TIM Triosephosphate Isomerase *Arabidopsis*, a 33 kDa protein (Fig. A-9). MS/MS fragmentation of DEFIGKK yielded six fragment ions of the twenty most intense peaks matching the expected masses b_1, b_3, b_3^0, b_4, y_2, y_3^*. MS/MS fragmentation of VASPQQAQEVHVAVR yielded twenty-four fragment ions of the thirty-one most intense peaks matching the expected masses b_2, b_5, b_6, b_6^{++}, b_9^0, b_{10}^{++}, y_4, y_5, y_5^{++}, y_6^{++}, y_7, y_7^{++}, y_8, y_8^{++}, y_9^{++}, y_{10}^{++}, y_{10}^{*++}, y_{11}^{++}, y_{12}^{++}, y_{13}^{++}, y_{14}^{++}, y_{14}^{*++}, y_{14}^{0++}. MS/MS fragmentation of IIYGGSVNGGNSAELAK yielded twenty fragment ions of the fifty-six most intense peaks matching the expected masses b_6, b_7, b_7^0, b_8, b_{14}, b_{15}, b_{15}^*, b_{15}^0, b_{16}, b_{16}^0, y_3, y_8^{++}, y_9, y_{10}, y_{10}^*, y_{11}, y_{12}, y_{13}, y_{13}^*, y_{14}, y_{14}^*, y_{15}, y_{15}^{++}, y_{15}^{*++}, y_{15}^{0++}, y_{16}^{++}. The third peptide from the second spot, matched GPLQLTWNYNYAQAGR, mass 1850.91 (926.45, 2+), to hevein-like antimicrobial peptide *Eunonymus europaeus*, a 34 kDa protein (Fig. A-10). MS/MS fragmentation yielded thirteen fragment ions of the fifteen most intense peaks matching the expected masses b_4, b_6, b_7^*, b_9, b_9^*, y_5, y_6, y_7, y_8, y_9, y_{10}, y_{11}, y_{12}. The fourth peptides from the second spot, matched RPPGR, mass 581.34 (292.09, 2+), LIESPAPGIISR, mass 1251.72 (626.87, 2+), to ATP synthase CF1 alpha subunit *H. annuus*, a 55 kDa protein (Fig. A-11). MS/MS fragmentation of RPPGR yielded six fragment ions of the thirty-four most intense peaks matching the expected masses y_1, y_2, y_2^*, y_2^{*++}, y_3, y_3^*. MS/MS fragmentation of LIESPAPGIISR yielded nine fragment ions of the twenty-nine most intense peaks matching the expected masses y_2, y_3, y_6, y_7^{++}, y_8, y_9, y_{10}, y_{10}^0. The fifth peptide from the second spot, matched

SGATGATPPAEAPLVTEEK, mass 1824.91 (913.44, 2+), drought-induced protein SDi-6 *H. annuus* a 18 kDa protein (Fig. A-12). MS/MS fragmentation yielded fifteen of the thirty-five most intense peaks matching the expected masses b_7, b_7^0, b_{12}^{++}, b_{13}, b_{15}^{0++}, y_4, y_7, y_8, y_9, y_{10}, y_{11}, y_{12}, y_{12}^{++}, y_{13}, y_{17}^*. The sixth peptide from the second spot, matched FMTEVR, mass 781.38 (390.90, 2+), glucosyltransferase-like protein *Arabidopsis* a 37 kDa protein (Fig. A-13). MS/MS fragmentation yielded four of the five most intense peaks matching the expected masses b_1, b_2, b_3, b_5.

Hydroponic Sundance stem samples 1-D PAGE analysis revealed 1 distinct spots around the marker region at 39 kDa. The mass spectra of tryptic digest of the spot indicated positive matches to two proteins (Table A-III). The first peptide matched SLGADVAIDYTK, mass 1251.63 (626.87, 2+), to Quinone oxido-reductase-like *H. annuus*, a 33 kDa protein (Fig. A-14). MS/MS fragmentation yielded twelve of the twenty most intense peaks matching the expected masses b_2, b_6^0, b_{11}, b_{11}^0, y_3, y_4, y_5, y_6,y_7, y_8, y_9^*, y_{10}. The second set of peptides matched VVILGDGNAK, mass 984.56 (493.31, 2+), DKVVILGDGNAK, mass 1227.68 (614.38, 2+), to phenylcoumaran benzylic ether reductase homolog TH3 *Tsuga heterophylla*, a 33 kDa protein (Fig. A-15). MS/MS fragmentation of VVILGDGNAK yielded nine of the thirteen most intense peaks matching the expected masses b_2, b_3, b_4, b_6, b_6^*, b_9, y_6, y_7, y_8. MS/MS fragmentation of DKVVILGDGNAK yielded five of the nineteen most intense peaks matching the expected masses b_2^0, y_6, y_7, y_9, y_{10}.

Hydroponic Sundance leaf samples (soluble fraction only) 1-D PAGE analysis revealed one distinct spot around the marker region at 25-35 kDa. The mass spectra of tryptic digest of the spot indicated positive matches to two homogenous proteins (Table

A-IV). The first protein match from the spot, matched peptides VASPQQAQEVHVAVR, mass 1617.86 (540.61, 3+), IIYGGSVNGGNSAELAK, mass 1648.84 (825.33, 2+), TRIIYGGSVNGGNSAELAK, mass 1905.99 (636.61, 3+) to TIM Triosephosphate Isomerase *Arabidopsis*, a 33 kDa protein (Fig. A-16). MS/MS fragmentation of VASPQQAQEVHVAVR yielded twenty-five fragment ions of the forty-five most intense peaks matching the expected masses b_2, b_3, b_6^{*++}, b_6^{0++}, b_7, b_8^{++}, b_{12}^{++}, y_2, y_3, y_3^{*++}, y_4, y_5, y_5^{++}, y_6, y_7, y_7^{++}, y_8, y_{10}^{++}, y_{10}^{*++}, y_{10}^{0++}, y_{11}^{*++}, y_{11}^{0++}, y_{12}^{++}, y_{13}^{++}, y_{13}^{*++}. MS/MS fragmentation of IIYGGSVNGGNSAELAK yielded twenty fragment ions of the fifty-six most intense peaks matching the expected masses b_6, b_7, b_7^0, b_8, b_{14}, b_{15}, b_{15}^*, b_{15}^0, b_{16}, b_{16}^0, y_3, y_8^{++}, y_9, y_{10}, y_{10}^*, y_{11}, y_{12}, y_{13}, y_{13}^*, y_{14}, y_{14}^*, y_{15}, y_{15}^{++}, y_{15}^{*++}, y_{15}^{0++}, y_{16}^{++}. MS/MS fragmentation of TRIIYGGSVNGGNSAELAK yielded eight fragment ions of the thirty-four most intense peaks matching the expected masses y_6, y_6^{++}, y_6^*, y_7^{++}, y_{10}, y_{12}^{++}, y_{14}^{*++}, y_{14}^{0++}. The second protein match from the spot, matched peptides VASPQQAQEVHVAVR, mass 1616.86 (540.61, 3+), VAYALSQGLKVIACVGETLEQR, mass 2347.26 (783.64, 3+), to Triosephosphate Isomerase *Lactuca sativa*, a 21kDa protein (Fig. A-17). MS/MS fragmentation of VASPQQAQEVHVAVR yielded thirty-four fragment ions of the fifty-three most intense peaks matching the expected masses b_2, b_3, b_5, b_6^*, b_6^{*++}, b_6^{0++}, b_7, b_8^{++}, b_{12}^{++}, b_{13}^{++},y_2, y_2^*, y_3, y_3^{*++}, y_4, y_5, y_5^{++}, y_6, y_7, y_7^{++}, y_8, y_8^{++}, y_{10}^{++}, y_{10}^{*++}, y_{10}^{0++}, y_{11}^{++}, y_{11}^{*++}, y_{11}^{0++}, y_{12}^{++}, y_{12}^{*++}, y_{12}^{0++}, y_{13}^{++}, y_{13}^{*++}, y_{14}^{++}. MS/MS fragmentation of VAYALSQGLKVIACVGETLEQR yielded six fragment ions of the eleven most intense peaks matching the expected masses b_6^{++}, b_{10}^{++}, b_{14}^{++}, y_2, y_{17}^{++}, y_{19}^{++}.

Soil based Sundance leaf samples 1-D PAGE analysis revealed one distinct spot around the marker region at 25-35 kDa in the control sample (no EDTA) and one spot in the EDTA exposed samples. The mass spectra of tryptic digest of the spot indicated positive matches to two distinct proteins (Table A-V). The peptides observed in the control spot, SSQSVKSK, mass 721.36 (362.24, 2+), GGSTGYDNAVALPAGGR, mass 1561.75 (781.96, 2+) matched to Oxygen-evolving enhancer protein 1, chloroplast precursor (OEE1) *Arabidopsis*, a 35 kDa protein (Fig. A-18). MS/MS fragmentation of SSQSVKSK yielded eleven fragment ions of the nineteen most intense peaks matching the expected masses b_3, b_3^0, b_4^{++}, b_4^{*++}, b_4^{0++}, b_6, b_6^{++}, y_1^*, y_2^{++}, y_3^0, y_6^{++}. MS/MS fragmentation of GGSTGYDNAVALPAGGR yielded sixteen fragment ions of the thirty-four most intense peaks matching the expected masses b_7, b_{11}^{*++}, b_{11}^0, b_{11}^{0++}, b_{12}, b_{12}^*, b_{12}^0, b_{13}, b_{14}, b_{14}^{++}, y_5, y_6, y_7, y_8, y_9, y_{12}. The peptide observed in the EDTA exposed samples matched YGGMVMoIWDR, mass 1095.52 Da (548.75, 2+) to Chitinase *H. annuus* a 32 kDa protein (Fig. A-19). MS/MS fragmentation yielded twelve and thirteen fragment ions respectively of the sixteen most intense peaks matching the expected masses of b_2, b_3, b_4, b_5, b_6, b_7, y_1, y_3, y_4, y_5, y_6, y_7, y_8.

1.3.2. Expression of proteins in Teddy Bear

1.3.2.1. Expression of proteins in response to arsenic and other metals in Teddy Bear

Soil–based Teddy Bear leaf samples 1-D PAGE analysis revealed one distinct spot around the marker region at 25-35 kDa. The mass spectra of the tryptic digest of the spots indicated a positive match to one protein (Table A-VI). The peptides in the spot, GVMTK mass 534.28 (268.03, 2+), VPADVK mass 627.36 (314.20, 2+) matched to HSP70 related protein *H. annuus*, a 28 kDa protein (Fig. A-20). MS/MS fragmentation of GVMTK yielded two fragment ions of the two most intense peaks matching the expected masses y_2^{++}, y_4^{++}. MS/MS fragmentation of VPADVK yielded two fragment ions of the eight most intense peaks matching the expected masses b_2, b_3.

1.3.2.2. Expression of proteins in response to chromium only in Teddy Bear

Soil–based Teddy Bear leaf samples 1-D PAGE analysis revealed two distinct spots, one around the marker region at 40 kDa and the other above 200 kDa. The mass spectra of the tryptic digest of the spots indicated positive matches to one distinct protein in each spot (Table A-VII). The peptide in the first spot, GVTTIIGGGDSVAAVEK, mass 721.36 (362.24, 2+) matched to cytosolic phosphoglycerate kinase *H. annuus*, a 42 kDa protein (Fig. A-21). MS/MS fragmentation yielded eight fragment ions of the eleven most intense peaks matching the expected masses b_{13}, y_5, y_8, y_9, y_{10}, y_{11}, y_{12}, y_{15}^{++}. The peptides in the second spot located past the final marker in the 200+ kDa range, NILVIGPVPGQK 1233.74 Da (617.93, 2+), QVVDIIPPGPELLVSEGESIK 2218.21 Da (740.44, 3+) matched to cytochrome f *H. annuus* a 32 kDa protein (Fig. A-22). MS/MS fragmentation of NILVIGPVPGQK yielded fourteen fragment ions of the sixteen most

intense peaks matching the expected masses b_3, b_4, b_5, b_6^*, b_8, b_8^*, b_{11}, y_4, y_6, y_7, y_7^{++}, y_8, y_9, y_{10}. MS/MS fragmentation yielded nineteen fragment ions of the twenty-nine most intense peaks matching the expected masses b_5^*, b_6, b_6^*, b_{12}, b_{14}, b_{15}^{++}, b_{17}^{++}, b_{20}^{*++}, b_{20}^{0++}, y_3, y_4, y_5, y_6, y_6^{*++}, y_{15}^{++}, y_{16}^{++}, y_{19}^{*++}.

1.3.2.3. Expression of proteins in response to lead and other metals in Teddy Bear

Hydroponic Teddy Bear leaf samples 1-D PAGE analysis revealed one distinct spot around the marker region at 40 kDa. The mass spectra of the tryptic digest of the spot indicated positive matches to two distinct proteins (Table A-VIII). The first set of peptides, VAILGAAGGIGQPLSLLMK mass 18085.06 (905.08, 2+), SSIEKGIKFASQ mass 1293.69 (646.91, 2+), matched to malate dehydrogenase precursor *Medicago sativa*, a 36 kDa protein (Fig. A-23). MS/MS fragmentation of VAILGAAGGIGQPLSLLMK yielded twenty-one fragment ions of the twenty-six most intense peaks matching the expected masses b_4, b_8, b_{10}, b_{12}, b_{12}^{++}, b_{15}^{++}, b_{17}, b_{18}^{*++}, b_{18}^{0++}, y_2, y_3, y_4, y_7, y_8, y_8^*, y_9, y_{11}, y_{11}^{++}, y_{12}, y_{13}, y_{15}. MS/MS fragmentation of SSIEKGIKFASQ yielded five fragment ions of the ten most intense peaks matching the expected masses b_4^0, b_5, b_6, b_6^0, b_8^{0++}. The second set of peptides, EPYIGR mass 733.38 (245.02, 3+), LYSIASSALGDFGDSK mass 1629.79 (815.85, 2+) matched to chloroplast ferredoxin-NADP+ oxidoreductase precursor *Capsicum annuum*, a 40 kDa protein (Fig. A-24). MS/MS fragmentation of EPYIGR yielded two fragment ions of the six most intense peaks matching the expected masses b_2, b_3. MS/MS fragmentation of LYSIASSALGDFGDSK yielded twenty-five fragment ions of the thirty-seven most

intense peaks matching the expected masses b_5^{++}, b_5^{0}, b_8^{0}, b_9^{0},b_{10}^{0}, b_{11}^{0++}, b_{12}, b_{13}, b_{14}, y_2, y_3, y_5, y_5^{*}, y_5^{0}, y_6, y_7, y_8, y_9, y_{10}, y_{10}^{++}, y_{11}, y_{12}, y_{14}, y_{14}^{*++}, y_{14}^{0++}.

Hydroponic Teddy Bear stem samples 1-D PAGE analysis revealed one distinct spot below the marker region at 40 kDa. The mass spectra of the tryptic digest of the spot indicated positive matches to two distinct proteins (Table A-IX). The first peptide, ILIIGGTGYLGK mass 1203.72 (602.95, 2+), matched to NADPH oxidoreductase putative *Arabidopsis* (Fig. A-25). MS/MS fragmentation of ILIIGGTGYLGK yielded seventeen fragment ions of the nineteen most intense peaks matching the expected masses of b_2, b_3, b_3^{++}, b_9^{0}, b_{10}, b_{10}^{0}, y_4, y_5, y_6, y_7, y_7^{*++}, y_7^{0++}, y_8, y_8^{0}, y_9, y_{10}, y_{11}^{0++}. The second set of peptides, VAILGAAGGIGQPLSLLMK mass 1808.06 (904.92, 2+), ALEGADVVIIPAGVPR mass 1575.90 (526.32, 3+), DDLFNINAGIVK mass 1317.69 (659.88, 2+), matched to malate dehydrogenase precursor *M. sativa* a 36 kDa protein (Fig. A-26). MS/MS fragmentation of VAILGAAGGIGQPLSLLMK yielded twenty-four fragment ions of the fifty-eight most intense peaks matching the expected masses b_7, b_{10}, b_{11}, b_{12}, b_{12}^{*}, b_{15}^{0}, b_{16}, b_{16}^{++}, b_{18}^{*++}, b_{18}^{0++}, y_5, y_7, y_8^{*}, y_9, y_9^{0}, y_{10}, y_{11}, y_{11}^{0}, y_{12}^{++}, y_{15}^{*++}, y_{15}^{0++}, y_{17}^{*++}, y_{17}^{0++}, y_{18}^{++}. MS/MS fragmentation of ALEGADVVIIPAGVPR yielded fourteen fragment ions of the forty-three most intense peaks matching the expected masses b_3, b_7, b_8, b_9, b_9^{++}, b_9^{0++}, b_{10}, b_{10}^{++}, b_{15}^{++}, y_4, y_6, y_6^{++}, y_7, y_8. MS/MS fragmentation of DDLFNINAGIVK yielded twelve fragment ions of the twelve most intense peaks matching the expected masses b_4^{++}, b_7^{++}, b_{10}, b_{11}, y_2, y_4, y_6, y_7, y_8, y_9, y_{10}, y_{10}^{++}.

Hydroponic Teddy Bear leaf samples soluble fraction 1-D PAGE analysis revealed two distinct spots, one around the marker region at 25-35 kDa and one around the marker region at 115-125 kDa. The mass spectra of the tryptic digest of the spots indicated positive matches to one distinct protein in each spot (Table A1-X). The peptide identified in the first spot, YGGMVMoLWDR, mass 1095.52 Da (548.28, 2+) to Chitinase *H. annuus* a 32 kDa protein (Fig. A-27). MS/MS fragmentation yielded thirteen fragment ions of the twenty-four most intense peaks matching the expected masses of b_2, b_3, b_4, b_5, b_6, b_7, b_8, y_3, y_4, y_5, y_6, y_7, y_8. The peptide identified in the second spot, AAALAALTSAFNSSSGR mass 1593.81 (797.92, 2+) to villin 3 *Arabidopsis* a 106 kDa protein (Fig. A-28). MS/MS fragmentation yielded four fragment ions of the five most intense peaks matching the expected masses of y_8^{++}, y_{10}^{++}, y_{15}^{++}, y_{16}^{++}.

Hydroponic Teddy Bear leaf samples transmembrane protein fraction 1-D PAGE analysis revealed one distinct spot around the marker region at 35 kDa. The mass spectra of tryptic digest of the spot indicated a positive match to one protein (Table A-XI). The peptide identified, VINTWQDIINR, mass 1313.71 (657.85, 2+), matched to photosystem II protein D1 *H. annuus*, a 38 kDa protein (Fig. A-29). MS/MS fragmentation yielded twenty fragment ions of the forty-two most intense peaks matching the expected masses b_5, b_7, b_7^*, b_8^*, b_8^0, b_9, b_9^*, y_4, y_5, y_5^*, y_6, y_7, y_8, y_9, y_9^{++}, y_9^{*++}, y_9^0, y_9^{0++}, y_{10}^{*++}, y_{10}^{0++}.

1.3.2.4. Expression of proteins in response to lead and other metals 2D Gel Separation in Teddy Bear

Hydroponic Teddy Bear leaf samples 2-D PAGE analysis revealed three distinct spots, two around the marker region at 25-35 kDa, 4 – 5 pH, and one around the marker region at 45 kDa, 5-7 pH. The mass spectra of the tryptic digest of the spots indicated positive matches to one protein in the first spot, two distinct proteins in the second spot and one protein type in the third spot (Table A-XII). The peptides identified in the first spot, RPHAVLINSISSSR, mass 1535.85 (769.78, 2+), NFRLKDLPRLIK, mass 1511.93 (757.49, 3+), NPNDLTIRFNTEVADKCLR, mass 2218.12. (740.27, 3+) matched to Cytokinin-O-glucosyltransferase *M. trunculata*, a 31 kDa protein (Fig. A-30). MS/MS fragmentation of RPHAVLINSISSSR yielded seven fragment ions of the seventeen most intense peaks matching the expected masses b_9^{++}, b_{12}^{++}, y_3, y_6, y_6^{*++}, y_6^{0++}, y_7. MS/MS fragmentation of NFRLKDLPRLIK yielded six fragment ions of the ten most intense peaks matching the expected masses y_4, y_5^{*}, y_6, y_7, y_7^{++}, y_{10}. MS/MS fragmentation of NPNDLTIRFNTEVADKCLR yielded seven fragment ions of the seven most intense peaks matching the expected masses b_3^{*}, b_{10}^{++}, b_{17}^{++}, y_5^{*++}, y_5^{0++}, y_8^{++}, y_{15}^{++}. The first set of peptides identified in the second spot, KHHLFG, mass 737.40 (737.59, 1+), AQKPVHDASKR, mass 1235.67 (619.31, 2+) matched to ABA/WDS stress induced protein *O. sativa*, a 24.5 kDa protein (Fig. A-31). MS/MS fragmentation of KHHLFG yielded three fragment ions of the thirteen most intense peaks matching the expected masses b_3, b_4, b_5. MS/MS fragmentation of AQKPVHDASKR yielded four fragment ions of the nineteen most intense peaks matching the expected masses y_5^{++}, y_6^{++}, y_8^{++}, y_{10}^{++}. The second set of peptides identified in the second spot, GRGLLVGMK, mass

929.55 (466.22, 2+), GEKTPWMHEYR, mass 1531.72 (766.64, 2+) matched to NAC domain protein *O. sativa*, a 34 kDa protein (Fig. A-32). MS/MS fragmentation of GRGLLVGMK yielded two fragment ions of the three most intense peaks matching the expected masses b_2, b_6. MS/MS fragmentation of GEKTPWMHEYR yielded four fragment ions of the nineteen most intense peaks matching the expected masses y_5^{++}, y_6^{++}, y_8^{++}, y_{10}^{++}. The peptides in the third spot matched 1-aminocyclopropane-1-carboxylic acid oxidase from two different species, DFQDAIK, 835.41 mass (179.08, 3+), VYGPIK, mass 675.40 (226.11, 3+), GLGGTSALLDFRV, mass 1107.59 (639.30, 2+), matched to Arabidopsis (Fig. A-33) and AIANFMSKAR, mass 1276.70 (370.23, 3+), matched to Rosa hybrid cultivar (Fig. A-34). MS/MS fragmentation of DFQAIK yielded three fragment ions of the three most intense peaks matching the expected masses b_2, b_3, y_3. MS/MS fragmentation of VYGPIK yielded five fragment ions of the twelve most intense peaks matching the expected masses b_3^{++}, b_4^{++}, y_2^{*}, y_3^{++}, y_4^{++}. MS/MS fragmentation of GLGGTSALLDFRV yielded eleven fragment ions of the twenty-nine most intense peaks matching the expected masses b_5^{0}, b_7^{0}, b_{12}^{0++}, y_3^{++}, y_4^{++}, y_5, y_6^{++} y_8^{++}, y_8^{*++}, y_{10}^{++}, y_{12}^{++} . MS/MS fragmentation of AIANFMSKAR yielded three fragment ions of the four most intense peaks matching the expected masses b_3^{++}, b_5^{++}, b_6^{++}.

Hydroponic Teddy Bear leaf samples 2-D PAGE analysis revealed four distinct spots, two around the marker region at 25-35 kDa, 4 – 5 pH, one around the marker region at 80-90 kDa, 5-7 pH and the last around the marker region at 25-35 kDa , 8-10 pH (Fig. 1-4). The mass spectra of the tryptic digest of the spots indicated positive matches to two

distinct proteins in three spots, and one distinct protein in the fourth spot (Table 1-II). The first peptide identified in the first spot, YGGMVMoLWDR, mass 1095.52 Da

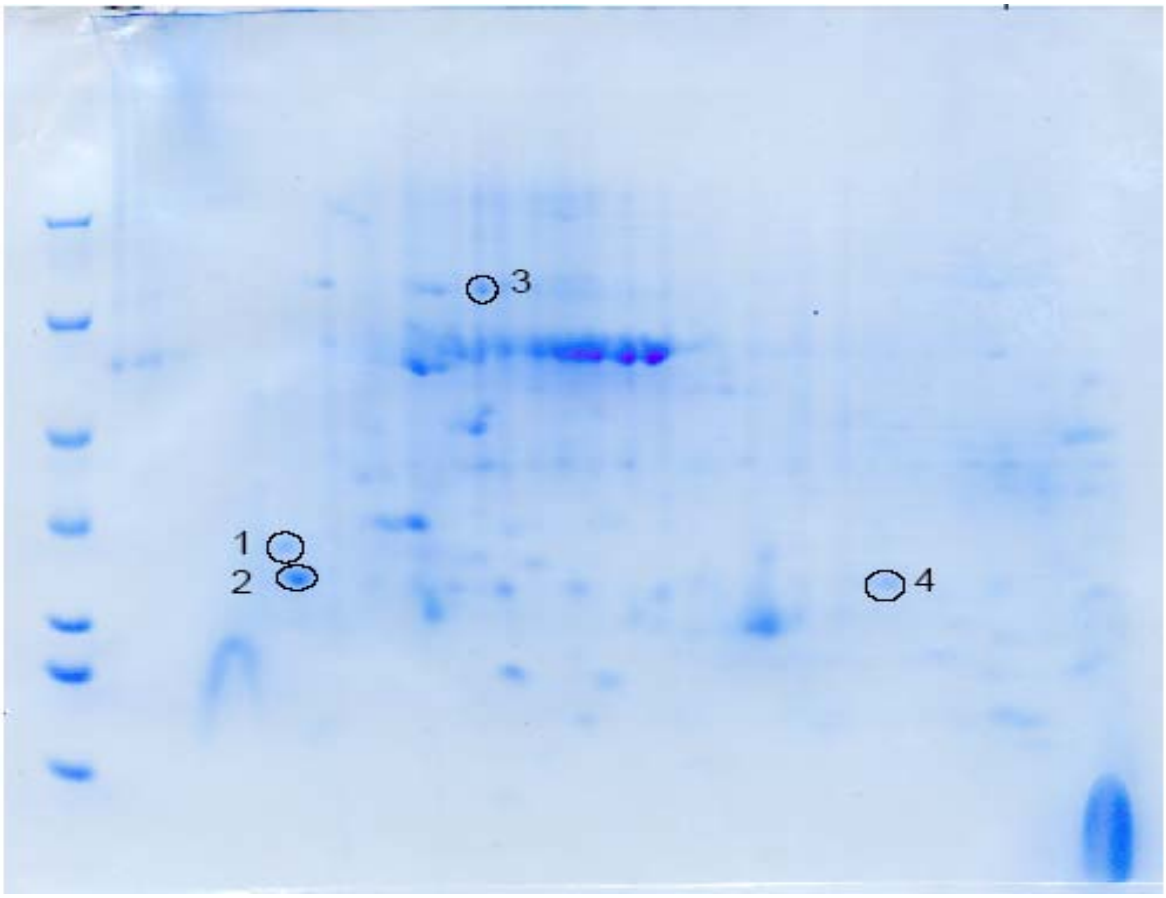
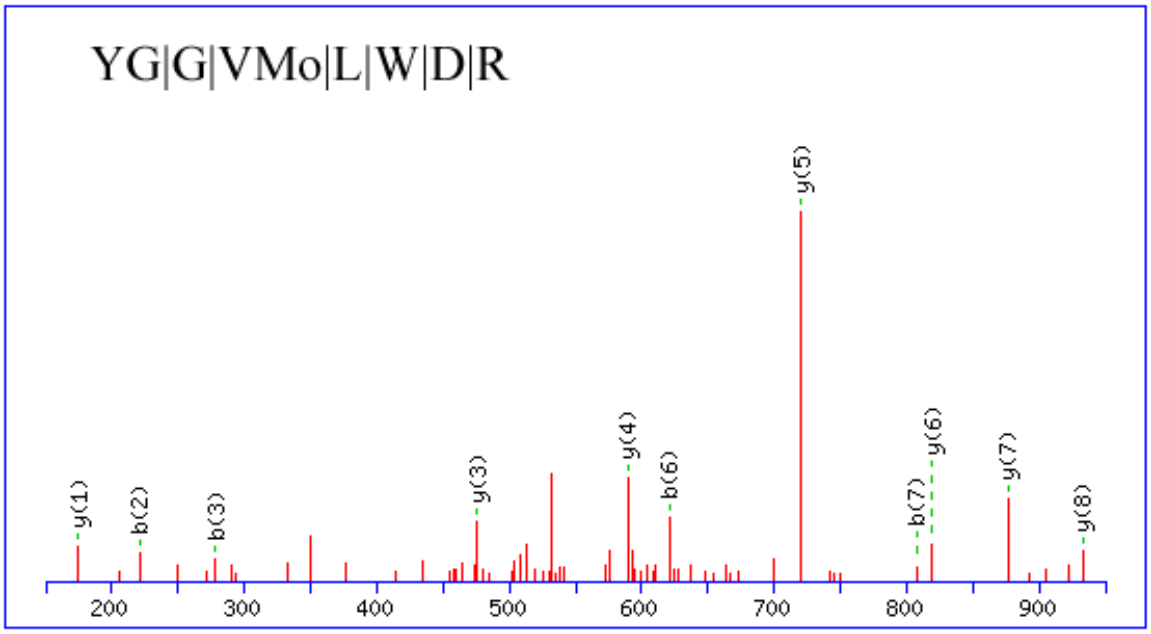


Figure 1-4. 2-D Gel Separation Teddy Bear Leaf Hydroponic Cd, Cr, Ni, Pb

(in submission and under review: Atkinson, I., Hasintha, C., vanKeulen, H., Cutright, T., Wei, R. Proteomic analysis of soluble proteins induced by toxic metals in *Helianthus annuus*. (Journal of Proteomics))

Table 1-II. 2-D Teddy Bear Leaf Hydroponic Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	pH	Accession no.	sp	Protein	Score	Peptide (Coverage)	Peptides
25-35	32	4-5	gi|126723930	*H. annuus*	Chitinase	62	1(2%)	YGGVMIWDR
25-35	32	4-5	gi|21633213	*H. vulgare*	jacalin - like protein LEM2	24	1 (3%)	SDVTLKALGVYVK
25-35	23	4-5	gi|53830843	*Solanum tuberosum*	Putative thaumatin-like protein	28	1(4%)	TQGGCNNPCTVYK
25-35	35	4-5	gi|30844168	*E. europaeus*	hevein-like antimicrobial peptide	19	1 (2%)	SRNPQVEDR
80-90	70	5-7	gi|18663	*Glycine max*	Heat Shock 70 kDa protein	41	4 (9%)	TTPSYVAFTDTER, NQVAMNPQNTVFDAK, SQVHEVVLVGGSTR, FELTGIPPAPR
80-90	71	5-7	gi|15219109	*Arabidopsis*	HSP 70B; ATP binding	41	4 (9%)	TTPSYVAFTDTER, IINEPTAAAIAYGLDK, FEELNMDLFR, NTIKDEK
20-35	32	8-10	gi|42573009	*Arabidopsis*	metal ion binding	32	1 (3%)	CCKGCQTKAKR



(A)

MEFTHPALLLLLFITVFSFLKPSTAAGIATYWGQQSDDTEGTLEAACATGNYQFV

NIAFLSTFGNNQQPVLNLAAHCDPASTCSRYSSQIKACQAQNVKVFLSIGGQRGS

YSLSSPQDAQQVADFLWNTYLGGQPATRPLGDAVLDGIDFDIEQGTDQFWSDLA

KAPAAYSSQKKVYLSAAPQCPFPSGDVRNQLLPAIREGLFDYVWVQFYNNEQCQ

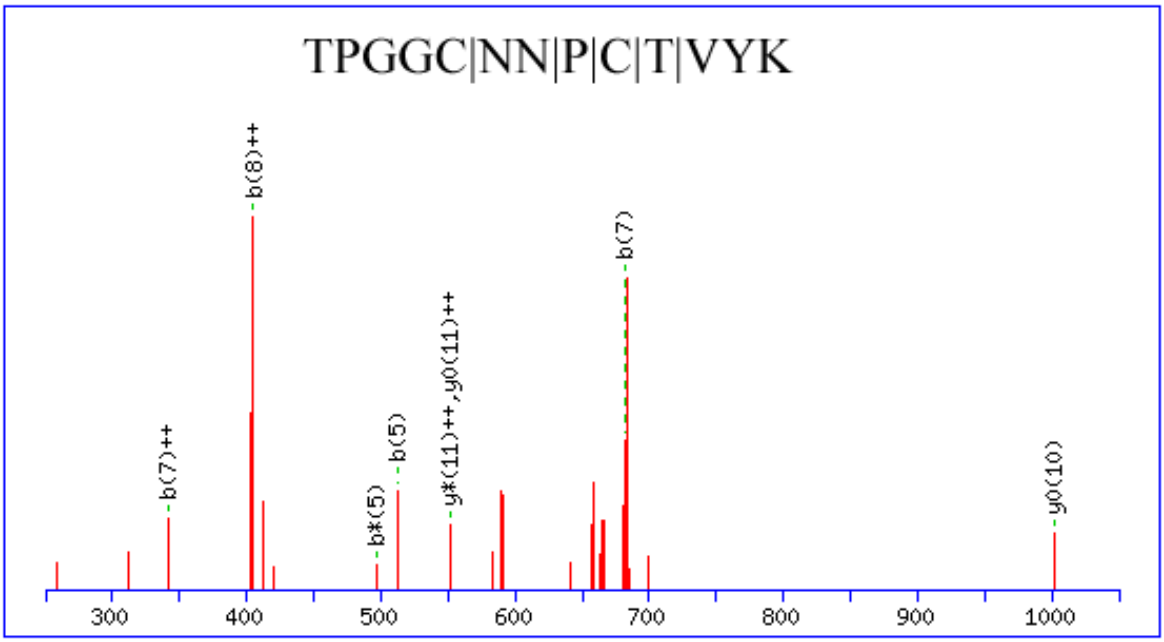
YGANADALLARWNEWTQVTTNTIFLGLPAAASGAAPSGGYIPPDILTSQILPSIKS

SPK**YGGVMLWDR**FYDKQSGYSDAIKGSIN

(B)

Figure 1-5. Tandem mass spectra of polypeptide (A) from Chitinase (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).

(548.75, 2+) to Chitinase *H. annuus* a 32 kDa protein (Fig. 1-5). MS/MS fragmentation yielded eleven fragment ions of the sixteen most intense peaks matching the expected masses of b_2, b_3, b_6, b_7, y_1, y_3, y_4, y_5, y_6, y_7, y_8. The second peptide identified in the first spot, SDVTLKALGVYVK, mass 1391.16 Da (696.59, 2+) matched to jacalin-like protein LEM2 *H. vulgare*, a 32 kDa protein (Fig. A-35). MS/MS fragmentation yielded ten fragment ions of the twenty-six most intense peaks matching the expected masses b_3^{0}, y_4, y_4^{++}, y_6, y_6^{++}, y_8^{++}, y_9^{++}, y_{10}^{*++}, y_{11}^{++}, y_{12}^{++}. The first peptide identified in the second spot, TQGGCNNPCTVYK, mass 1499.01 Da (750.54, 2+) matched to thaumatin-like protein *S. tuberosum*, a 25 kDa protein (Fig. 1-4). MS/MS fragmentation yielded eight fragment ions of the twenty-two most intense peaks matching the expected masses b_5, b_5^{*}, b_7, b_7^{**}, b_8^{**}, y_{10}^{0}, y_{11}^{*++}, y_{11}^{0++}. The second peptide in the second spot, SRNPQVEDR, 1099.38 Da (550.70, 2+) matched to hevein-like antimicrobial peptide, *E. europaeus*, a 32 kDa protein (Fig. A-36). MS/MS fragmentation yielded three fragment ions of the seven most intense peaks matching the expected masses y_3, y_4, y_6. The first set of peptides in the third spot, TTPSYVAFTDTER, 1486.44 Da (744.23, 2+), NQVAMNPQNTVFDAK, 1676.68 Da (839.35, 2+), IINEPTAAAIAYGLDK, 1658.56 Da (830.29, 2+), FEELNMDLFR, 1332.34 Da (667.18, 2+), SQVHEVVLVGGSTR, 1465.76 Da (733.89, 2+), FELTGIPPAPR, 1196.48 Da (599.25, 2+), NTIKDEK, 848.14 Da (425.08, 2+) matched to heat shock protein 70, *Arabidopsis*, a 70 kDa protein (Fig. A-37). MS/MS fragmentation of TTPSYVAFTDTER yielded eighteen fragment ions of the thirty-three most intense peaks matching the expected masses b_4, b_9, y_3^{*}, y_3^{0}, y_4, y_5, y_5^{*}, y_6, y_7, y_7^{0}, y_8, y_9, y_9^{++}, y_{10}, y_{11}, y_{11}^{++}, y_{11}^{*++}, y_{11}^{0++}. MS/MS fragmentation of NQVAMNPQNTVFDAK



(A)

MHFLKFFPLFVFLYFGQYYLYVTHAATFDITNRCTYPVWAAASPGGGRRLDSGQ

TWNLNVNPGTIQARIWGRTNCNFDGSGRGKCETGDCNGLLECQGYGSPPNTLAE

FALNQPNNLDFVDISLVDGFNIPMEFSPINGGCRNLLCNAPINDQCPNELR**TPGG**

CNNPCTVFKTNEFCCTNGPGSCGPTDLSRFFKQRCPDAYSYPQDDPTSLFTCPAG

TNYKVVFCP

(B)

Figure 1-6. Tandem mass spectra of polypeptide (A) from thaumatin-like protein (*S. tuberosum*) as indicated (bold) in the primary amino acid sequence (B).

yielded twenty fragment ions of the thirty-seven most intense peaks matching the expected masses b_4, b_5^*, b_6, b_{11}, b_{13}, y_4, y_6^*, y_7, y_8, y_9, y_9^{++}, y_{10}, y_{11}, y_{11}^*, y_{11}^{*++}, y_{11}^0, y_{11}^{0++}, y_{12}, y_{12}^{*++}, y_{13}. MS/MS fragmentation of IINEPTAAAIAYGLDK yielded eleven fragment ions of the twenty-six most intense peaks matching the expected masses b_3, b_5^{++}, b_8, b_9^{++}, b_{11}^{*++}, b_{11}^{0++}, b_{14}^{*++}, b_{14}^{0++}, b_{15}^{++}, b_{15}^{*++}, y_{11}^{*++}. MS/MS fragmentation of FEELNMDLFR yielded four fragment ions of the eleven most intense peaks matching the expected masses b_2, b_5, b_5^0, b_8. MS/MS fragmentation of SQVHEVVLVGGSTR yielded nine fragment ions of the thirty-five most intense peaks matching the expected masses b_8^{++}, y_5, y_5^*, y_6, y_7, $y_{8,}$ y_{10}, $y_{11,}$ b_{13}^{++}. MS/MS fragmentation of FELTGIPPAPR yielded ten fragment ions of the thirty most intense peaks matching the expected masses b_2, b_4, b_6, b_6^0, b_9^{0++}, b_{10}, y_5, $y_{7,}$ y_8, y_9. MS/MS fragmentation of NTIKDEK yielded six fragment ions of the eight most intense peaks matching the expected masses b_4^{++}, b_5^{++}, y_2^{++}, y_4^{*++}, y_4^{0++}, $y5^{++}$. MS/MS fragmentation of FELTGIPPAPR yielded ten fragment ions of the thirty most intense peaks matching the expected masses b_2, b_4, b_6, b_6^0, b_9^{0++}, b_{10}, y_5, $y_{7,}$ y_8, y_9. The second set of peptides in the third spot matched to RuBisCo, a protein whose strong band is evident in every gel, and seen as a co-contributing protein in many analyses. The peptide identified in the fourth spot, CCKGCQTKAKR, 1223.28 Da (612.65, 2+), matched to a metal ion binding protein, *Arabidopsis*, a 30 kDa protein (Fig. A-38). MS/MS fragmentation yielded four fragments of the eight most intense peaks matching the expected masses b_3, b_4, b_5, b_7.

1.4. Discussion

An abundance of defense and stress induced proteins have been identified within the samples analyzed. Protein families have been highlighted, and in many cases these proteins have a primary role associated with their function, and a seeming secondary or surveillance role involving stress tolerance. In our study we have seen proteins associated with biotic stress including mechanical wounding, bacterial, viral and fungal infection; and also with abiotic stress, such as drought, light, ABA, temperature, osmotic pressure and toxic chemicals. These defense/stress response proteins are responsible for maintaining proper cell function and homeostasis, implying the need for roles in maintaining cell structure and integrity, resource allocation and usage, material transport and storage, and repair mechanisms. To see such a wide array of stress response proteins exhibited in response to metal stress implies a more universal set of stress related functions of the proteins expressed, as opposed to many more specific responses observed from some proteins to individual stresses.

Chitinase, identified in samples from both Sundance and Teddy Bear cultivars, in roots and leaf tissues, exposed to both As and Pb in conjunction with Cd, Cr and Ni, enzymatically catalyzes the degradation of chitin, a linear ß1-4-linked polymer of N-acetyl-D-glucoseamine (GlcNAc). Chitinases are expressed by stress such as fungal and bacterial infection, and are also documented to be expressed under other stress conditions such as heavy metal exposure. Chitinases were positively determined to be up-regulated on the genetic level by As exposure using RT-PCR expression (Fig 1-7B, lane 6 & 7) and western blot analysis in *H. annuus* (Fig 1-8, Lane 3,4 & 8) [11]. This study described the isolation and determination of chitinase in plants exposed to As alone and in combination

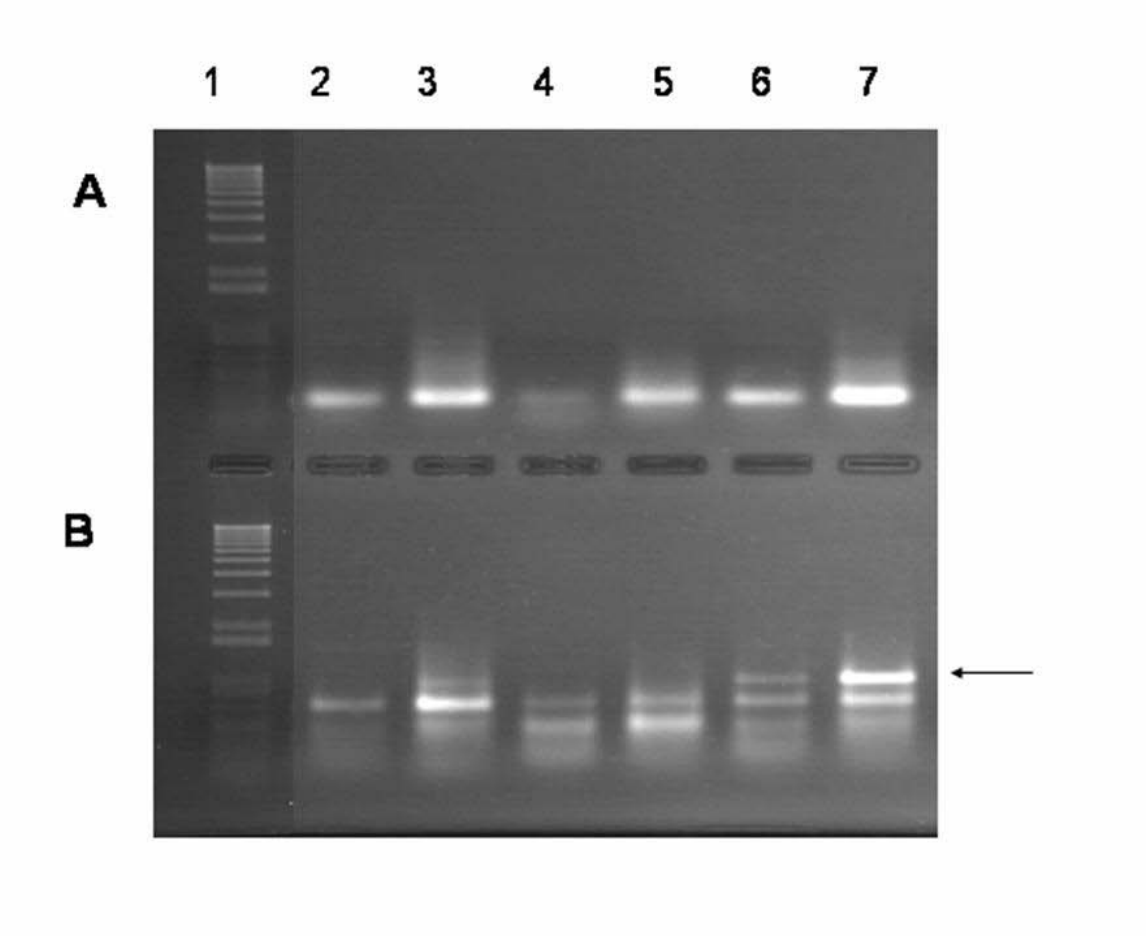


Figure 1-7. Semi-quantitative RT-PCR of total RNA samples showing an up-regulation of Chitinase in samples containing As [11].

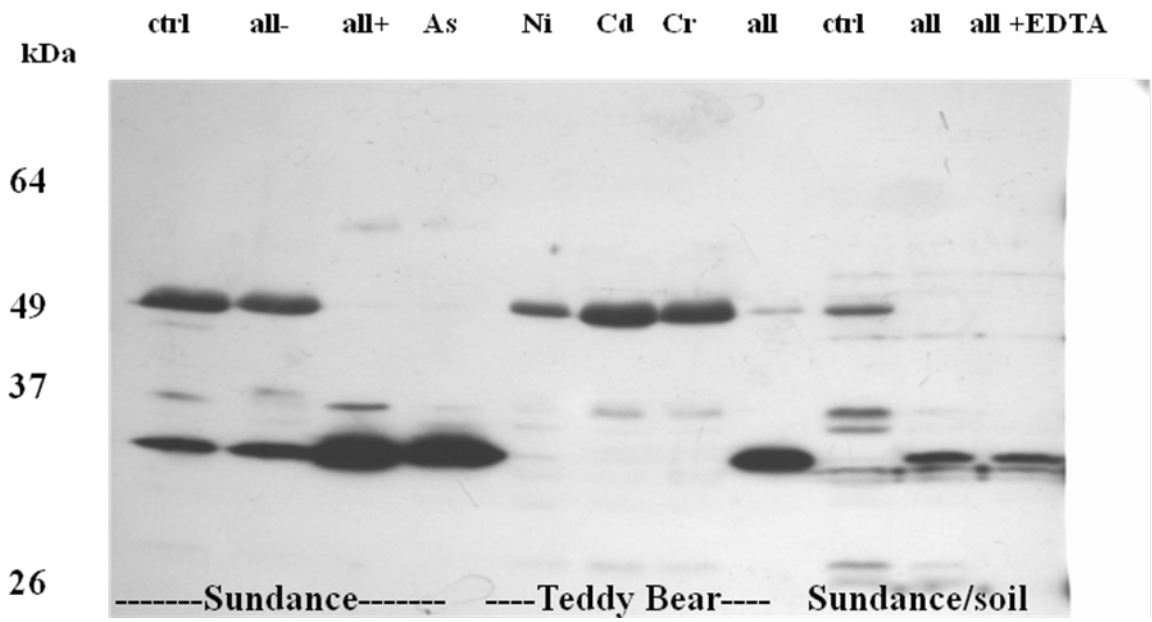


Figure 1-8. Western Blot analysis of Sundance and Teddy Bear samples showing an up-regulation of Chitinase in samples exposed to As [11].

with Cd, Cr, and Ni. In addition chitinase has now also observed to be up-regulated in the presence of Pb alone and in combination with Cd, Cr, and Ni.

Oxygen-evolving enhancer protein 1 (OEE1) (match to *Arabidopsis*), thylakoid membrane protein, was identified from Sundance leaves, upon exposure to Pb, Cd, Cr, and Ni. No OEE proteins has been described in the NCBI database for any species of *Helianthus.* This protein shares similarity with the proteins attributed to stabilize the manganese cluster which is the primary site of water splitting in the choloroplast, and is a subunit of the photosystem II complex. This protein has been documented to be increased under salt stress conditions in *Bruguiera gymnorrhiza* [19] to compensate for the depression of photosynthesis. This protein has also been found to exhibit thioredoxin activity in green algae [20], a possible mechanism to deal with environmental stresses on the photosystem II complex. The OEE1 protein has been attributed with regulating light induced oxidative damage to the photosystem II complex and maintaining the efficient turnover of the damaged subunit [21]. OEE2, an oxygen evolving complex protein, has also been documented to have increased expression in *Vigna unguiculata* under Mn stress [20].

An amino terminal coiled-coil, nucleotide-binding site containing, leucine-rich repeat resistance gene analog (CC-NBS-LRR RGA) protein was identified from Sundance roots exposed to As, Cd, Cr and Ni. These proteins are typically expressed in plants as a triggered defense response to a broad range of disease conditions. In addition to pathogen related expression a CC-NBS-LRR protein has been documented to be upregulated under water stress, to confer significant drought tolerance in *Arabidopsis* [22].

Lectin (*H. tuberosus*) was identified from Sundance roots exposed to As, Cd, Cr, Ni. No lectins exist in the database for *H. annuus*. Lectins are a group of proteins classified by their recognition and binding specificity to carbohydrate ligands that are widely distributed in higher plants and are believed to play a role in plant defense [23, 24, 25]. The family of jacalin-related plant lectins is mannose binding, specific for Man-α-1-3Man and Man-α-1-2Man, which are abundant in the glycoconjugates exposed on the surface of viruses, bacteria, and fungi and on the epithelial cells along the gastrointestinal tract of lower animals [26]. This specificity may allow these lectins to contribute to plant stress and defense response to pathogens. The jacalin-related lectins are composed of 1-6 copies of the jacalin-like lectin domain, a small 15 kDa domain that can explain why the protein was identified from a band in the 60 kDa range. Lectin is also observed to be up-regulated in the presence of Pb in combination with Cd, Cr, and Ni.

Actin was identified in Sundance leaves exposed to As alone, not in combination with any other metals. Actin is a highly conserved eukaryotic protein and one of the most abundant proteins in a eukaryotic cell. Actin is associated with cell structure and various cell maintenance functions thru interaction with a large number of Actin-binding proteins. Actin has been reported as up-regulated in *Cannabis sativa* under copper stress [16]

Glyceraldehyde 3-phosphate dehydrogenase B subunit (GAPDH) (*Arabidopsis*) was identified in Sundance leaves exposed to As alone, not in combination with any other metals. The peptides observed and the protein identified share homology with *H. annuus* GAPDH. GAPDH catalyzes the sixth step of glycolysis, playing a primary role in plant

metabolism. This protein has also been identified as an anti-stress protein due to its interaction with cytosolic thioredoxins [27].

Glycolate oxidase (GOX) (*Arabidopsis*) was identified in Sundance leaves exposed to As alone, not in combination with any other metals. There are no glycolate oxidase proteins in the database for *Helianthus.* GOX is a key photorespiratory enzyme, carrying out the oxidation of glycolate to glyoxylate. Its activity was documented to increase due to osmotic stress [28], salt stress [29] and drought stress [30].

Photosystem II protein D1 was identified in Sundance and Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. Photosystem II is the site of water oxidation in photosynthesis and is particularly susceptible to damage by light stress [21]. This proteins upregulation is observed in response to stress damage do to turnover and is closely associated with OEE1. [31]

Triosephosphate isomerase (TIM, *Arabidopsis*) was identified in Sundance and Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. TIM reversible interconverts dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate in glycolysis and gluconeogensis [1]. This is a general housekeeping enzyme found in all cells; no association of this enzyme with metal stress has been previously reported.

Hevein-like antimicrobial protein (*Euonymus europaeus*) was identified in Sundance and Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. There are no hevein-like proteins in the database for *Helianthus.* Hevein-like proteins are small proteins with chitin binding activity associated with the lectin super family. These proteins have antifungal activity, even against non-chitinous fungus, and are expressed in response to microbial

infection in plants [32, 33]. This activity suggests additional pathogen resistance properties associated with the protein in addition to its known chitin-binding function.

ATP synthase CF1 alpha subunit was identified in Sundance leaves exposed to Cd, Cr, Ni, and Pb. ATP synthase hydrolyzes ATP within the chloroplast and is driven by the electron-transport process. ATP synthases are up-regulated under osmotic and stress conditions [23,34]. ATP synthase activity is reduced under Fe^{2+} and oxidative stress conditions due to oxidative modification of amino acid residues required for proper function [35]

Drought-induced protein SDi-6 (designated sdi for sunflower drought induced) was identified in Sundance leaves exposed to Cd, Cr, Ni, and Pb. SDi-6 is a water stress induced protein whose presence is associated with an increase in plant tolerance to drought conditions [36]. SDi-6 has no matching plant family domain in the NCBI database although drought induced proteins are also associated with ABA stress as abscisic acid levels rise in low water conditions [36].

Glucosyltransferase-like protein (*A. thaliana*) was identified in Sundance leaves exposed to Cd, Cr, Ni, and Pb. There are no protein matches or homologies to glucosyltransferase proteins in the database for *Helianthus*. This metabolic protein type is documented to be upregulated in response to ABA, drought, salt and cold stress [37].

Quinone oxido-reductase-like was identified in Sundance stems exposed to Cd, Cr, Ni, and Pb. Quinone oxido-reductase (QOR) is regulated by the aryl hydrogen receptor complex as a detoxification mechanism for many toxic hydrocarbon compounds. QOR has been documented to be up regulated in response to Ar, Cd, and Cr exposure in mouse hepatoma Hepa 1c1c7 cells studied [38].

Phenylcoumaran benzylic ether reductase (PCBER) homolog TH3 (*T. heterophylla*) was identified in Sundance stems exposed to Cd, Cr, Ni, and Pb. There are no homologous proteins in the *Helianthus* database. PCBER is part of a class of related enzymes involved in the production of plant defense compounds in response to fungal stress. These related enzymes and their products are found across the entire plant kingdom and suggest the possibility of an underlying universal defense response [39].

Isoflavon reductase-homologue (IFR) (*N. tabacum*) was identified in Teddy Bear stems and leaves exposed to Cd, Cr, Ni, and Pb. There are no homologous proteins in the database for *Helianthus*. Phylogenetically related to phenylcoumaran benzylic ether reductase. IFR also plays a role in plant defense to fungal infection thru the production of antimicrobial phytoalexins [39].

Malate dehydrogenase (MDH) (*M. sativa*) was identified in Sundance and Teddy Bear stems and leaves exposed to Cd, Cr, Ni, and Pb. There are no MDH proteins in the database for *Helianthus*. MDH has an antioxidant role in cell defense through its conversion of malate to oxaloacetate, an ROS scavenging specie [40]. MDH has also been identified as an anti-stress protein due to its interaction with cytosolic thioredoxins [27].

Chloroplastic drought-induced stress protein (CDSP-34) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. CDSP-34 is a member of the plastid associated protein (PAP) fibrillin family. There exists no match in the NCBI protein database for this protein or family (PAP fibrillin super family) to any *Helianthus* species. CDSP-34 is documented to be expressed in response to many stresses; drought, high light exposure, oxidative, bacterial infection and mechanical wounding [41-44].

Heat-shock protein 70 (HSP-70, *Arabidopsis*) was expressed in Teddy Bear leaves exposed to As, Cd, Cr, and Ni and Pb, Cd, Cr, and Ni. HSP's are responsible for protein folding, assembly, translocation and degradation in many normal cellular processes. HSP's stabilize proteins and membranes, and can assist in protein refolding under stress conditions [45]. Hsp70 protein acts as a molecular chaperone binding to hydrophobic residues exposed by stress conditions, preventing proteins from agglomerating and helping to refold these denatured proteins. Hsp70 also assists in protein transport across membranes by its hydrophobic interaction, assisting to unfold, stabilize and refold proteins during transport. [45,46]. Members of the HSP 70 family are upregulated as a result of thermal and oxidative stress including heavy metal exposure [47]. HSP-70 was also identified as an anti-stress protein due to its interaction with cytosolic thioredoxins [27].

Thaumatin-like protein, a class 5 pathogensis-related proteins, were expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. Thaumatins are expressed in response to fungal and viral stresses as well as abiotic stress factors such as increased salinity and drought [48- 50]. Many thaumatin-like proteins also bind to ß1,3-glucans and also display endo- ß 1,3-glucanase activity, as associated with antimicrobial resistance [51]. A Thaumatin-like protein has been positively documented to be up-regulated in response to Pb exposure using RT-PCR expression in *H. annuus* (performed by Dr vanKeulen, results not shown).

Chloroplast ferredoxin-NADP+ oxidoreductase precursor (FNR) (*C. annuum*) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. There are no homologous

proteins in the database for *Helianthus*. FNR plays an antioxidant stress response role in cellular defense against oxidative damage [52, 53].

Villin 3 (*Arabidopsis*) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. There are no homologous proteins in the database for *Helianthus*. Villin is an actin binding, bundling, severing, nucleating and capping protein whose function is regulated by calcium binding and is expressed ubiquitously in *Arabidopsis* [54].

Cytokinin-O-glucosyltransferase (*M. trunculata*) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. There are no homologous proteins in the database for *Helianthus*. Cytokinin-O-glucosyltransferases glycosylate cytokinins, plant hormones essential for growth and development and that are also documented to provide protection against degradative enzymes [55]. Cytokinin-O-glucosyltransferase has also been documented to be upregulated in leaves of plants under drought stress [56].

Abscisic acid, water deficit (ABA/WDS) stress induced (*Orytza sativa*) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. No homologous proteins exist in the database for *Helianthus.* This is a family of proteins is induced by abscisic acid stress and drought/water deficit stress (WDS) and also has a noted upregulation under ripening [57]. The proteins role and function is not well understood.

NAC Domain protein (*H. annuus*) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. NAC Domain proteins are a family of transcription factors that have been observed to be upregulated in response to biotic and abiotic stressors, specifically cold, salt, water and wounding [58] as well as pathogen recognition and defense [59].

1-aminocyclopropane-1-carboxylate oxidase (ACO) (*Arabidopsis*) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. ACO converts 1-aminocyclopropane-

1-carboxylate to ethylene, a hormone in plants that plays a role in development, and response to stress and pathogen defense [60]. The ethylene plant stress-response pathway is induced by mechanical wounding, chemical, metal, drought, temperature stresses [61].

Jacalin - like protein LEM2 (*H. vulgare*) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. No homologous protein exists in the database for *Helianthus*, although this protein does share some family homology to the lectin identified in Sundance leaves exposed to Pb, Cd, Cr, Ni. Jacalin-like proteins are a class of lectins that bind carbohydrates and are expressed in response to abiotic and biotic stresses [62, 63].

Heavy-metal-associated domain-containing protein (HMA) (Arabidopsis thaliana) was expressed in Teddy Bear leaves exposed to Cd, Cr, Ni, and Pb. No homologous protein exists in the database for *Helianthus*. HMA domain containing proteins are known as transport and detoxification proteins. HMA proteins in addition to metal stress have been shown to be upregulated under drought, light, temperature and ABA stressors [64].

Cytosolic phosphoglycerate kinase (CytPGK) (*H. annuus*) was identified in Teddy Bear leaves exposed to Cr. CytPGK is a glycolysis enzyme that transfers a phosphate group from 1,3-bisphosphoglycerate to ADP, forming ATP. CytPGK has been shown to be upregulated under cold stress conditions [65].

Cytochrome f (*H. annuus*) was identified in Teddy Bear leaves exposed to Cr. Cytochrome *b*6*f* is the electron/proton transfer protein complex of oxygenic photosynthesis that functions between photosystems II and I. Cytochrome *f*, the largest subunit of the *b*6*f* complex, consists of an N-terminal 251-residue redox-active domain in the lumen-side aqueous phase, and a 35-residue C-terminal transmembrane anchor in the

photosynthetic membrane of *Chlamydomonas reinhardtii* [66]. There are no citable articles implying cytochrome f upregulation due to stress.

1.5. Conclusions

The expression of many proteins are regulated by biotic and abiotic stresses, suggesting the occurrence of complex mechanisms that control gene expression in response to environmental stresses. Similarly, evidence exists for crossed response by different stimuli [69]. It is well established that many environmental stimuli such as salinity, drought, heat or cold, elicit common stress responses in up-regulated proteins and that these proteins share a general role of protecting the plant from damage and repairing damage already incurred [70]. The proteins observed in *H. annuus* as a result of heavy metal exposure follow this response, as stress related proteins involved in defense and repair mechanisms, although not individually specific to heavy metal binding. This shared functionality of proteins under various stress condition is favored by evolution theory, which encourages the re-use and adaptation of existing proteins for multiple functions.

Proteins identified in the analysis encompass a wide role, including those classified with metabolic function, photosynthesis, general cellular maintenance, antimicrobial response, drought and water stress response, general stress and antioxidant roles. These potentially expressed proteins may play a role in rendering plants tolerant of these toxic metals, thus enabling the plants to hyperaccumulate them. Understanding how these proteins function in response to metal stress and determining how their roles confer stress tolerance will allow a greater application and effective implementation of phytoremediation.

The LC-MS identification of potentially expressed proteins still requires a need to set the bar for positive identification of a proteins gene upregulation. Coupling a matched peptide from the MASCOT database to a gene sequence for a protein allowed for the design of a primer and the expression of the gene for RT-PCR analysis. Further generation of an anti-serum to the determined protein sequence allowed Western Blot analysis. Using these techniques a positive upregulation of Chitinase in root and leaf tissues exposed to As and Pb alone and in combination with Cd, Cr, and Ni in *H. annuus* has been documented. Additionally a Thaumatin like proteins up-regulation in leaf tissues exposed to Pb in combination with Cd, Cr, and Ni in *H. annuus* has been documented.

Future work will be directed to addressing several issues this study has brought to light. First is the doubt of positive identification via LC-MS and bioinformatics alone. A protein identification based on a single peptide will always leave a question as to the confidence. Additional sequence coverage, even just a second tryptic peptide match, greatly increases confidence in the likelihood of correlation to the database resolved protein. To increase sequence coverage analysis methodology can be refined, shifting to a nano-flow analysis alone will greatly increases mass spectrometric results. Cross sequence coverage can also be performed using additional digest parameters with alternative enzymes which also allows for a very strong confidence even in the case of very limited sequence coverage. The second problem is the bioinformatics database itself. Protein identifications made are highly determined by the skill of the user in interpreting the results of the mass spectrum analyzed and properly filtering the database results. With the sheer majority of organisms' protein libraries being far from complete

the likelihood of fit to other species informatics need to be considered. Oftentimes an understanding of protein function and structure is needed, as some initial search results may not seem relevant or match the gel data determined (mass and P.I.) and be prematurely removed from possible matches. Finally there is high conservation of many peptide motifs across a wide range of proteins and protein families which can force it to the user's interpretation to determine a "goodness of fit" for a correct match. Further development of the databases utility and an increase in the repository of protein information will help resolve this issue.

In summary, the versatility of methodology, wide range of application and near limitless combinations of chromatography and mass spectrometry techniques 2D LC-Ion Trap MS/MS coupled with MASCOT bioinformatics is a powerful tool available to the biochemical researcher. The ability to tackle large mixtures of proteins and molecules, with basic preparation is an amazing feat of modern scientific application. Further "de novo" sequencing to determine identity of proteins unknown to the database can be performed by varying additional digest methods and conditions to increase the total peptide identification, elucidate secondary structure, and improve confidence. Separation gradients and column specifications can be refined with repetitive analysis to tailor fit each samples unique make-up. Additional steps can be incorporated to further identify or isolate particular peptides or proteins; such as the incorporation of affinity columns or size exclusion steps in series with or in place of the CapTrap. Peptide mass mapping with the sensitivity of the ion trap coupled with collision-induced dissociation spectra allows for an incredibly detailed picture of a protein to be created, including identification of posttranslational modifications and single amino-acid variations. Even

whole protein complexes are able to be studied with ESI-ion trap methodology, expanding the technique beyond protein identification to exploring protein function and interaction. In all, this technique stands on the forefront of a rapidly expanding field of Proteomics wherein it seems the only limitation is the imagination of the BioMolecular Chemist.

1.6. References

1. Voet, Voet & Pratt, Fundamentals of Biochemistry, John Wiley & Sons, Inc. 1999. Pg 100-103, 391-393.
2. Tang, L., Kebarle, P. (1993) Dependence of ion intensity in electrospray mass spectrometry on the concentration of the analytes in the electrosprayed solution. *Analytical Chem*, 65(22):3654-3668.
3. Baker, A.J.M., McGrath, S.P., Reeves, R.D., Smith, J.A.C. (2000). Metal hyperaccumulator plants: a review of the ecology and physiology of a biological resource for phytoremediation of metal-polluted soils. In: Terry, N., Banuelos, G. (Eds.). Phytoremediation of Contaminated Soil and Water. Lewis Publishers, Boca Raton, FL., pp. 85-107.
4. Cafer, T . Pepe, K. Cutright, T. (2005) The effect of EDTA on *Helianthus annuus* uptake, selectivity, and translocation of heavy metals when grown in Ohio, New Mexico and Colombia soils. *Chemosphere*, 5:1087-1095.
5. Gallego, SM, Benavides, MP, Tomaro, ML (1999). Effects of cadmium ions on antioxidative defense system in sunflower cotyledins. *Biologia Plantarum* 42: 49-55
6. Hamer, D.H. (1986). Metallothioneins. *Ann. Rev. Biochem*. 55: 913-951.
7. Steffens, J.C. (1990). The heavy metal-binding peptides of plants. *Ann. Rev. Plant Physiol.* 41: 553-575.
8. Rauser, W.A. (1995). Phytochelatins and related peptides. *Plant Physiol*. 109: 1141-1149.
9. Ow, D.W. (1996) Heavy metal tolerance genes: prospective tools for bioremediatuion. *Res. Conserv. Recycling* 18: 135-149.

10. Lasat, M. (2002) Phytoextraction of toxic metals: A review of Biological mechanisms. *J. Environ. Qual*. 31: 109-120.
11. vanKeulen, H., Cutright, T., Wei, R. (2008). Arsenate-induced expression of a class III chitinase in the dward sunflower *Helianthus annuus. Environmental and Experimental Botany*. 63: 281-288.
12. Rios-Arana, J.V., Gardea-Torresdey, J.L., Webb, R., Walsh, E.J. (2005). Heat shock protein 60 (HSP60) response of *Plationus pattulus* to combined exposures of arsenic and heavy metals. *Hydrobiologia* 546: 577-585.
13. Garcia, J. S., Gratao, P. L., Azevedo, R. A., Arruda, M. A. Z. (2006) Metal Contamination Effects on Sunflower (*Helianthus annuus* L.) Growth and Protein Expressoin in Leaves During Development. *Journal of Agricultural and Food Chemistry*, 54:8623-8630
14. Kieffer, P., Dommes, J., Hoffmann, L., Hausman, J-F, Renaut, J. (2008) Quantitative changes in protein expression of cadmium-exposed poplar plants. *Proteomics*, 8:2514-2530
15. Roth, U., Roepenack-Lahaye, E., Clemens, S. (2006) Proteome changes in *Arabidopsis thaliana* roots upon exposure to Cd2+. *J. of Experimental Biology*, 57 (15) 4003-4013
16. Bona, E., Marsano, F., Cavaletto, M., Berta, G. (2007) Proteomic characterizations of copper stress response in the *Cannabis sativa* roots. *Proteomics*, 7, 1121-1130
17. Fuhrs, H., Hartwig, M., Molina, L. E. B., Heintz, D., Dorsselaer, A. V., Braun, H. P, Horst, W. J. (2008) Early manganese-toxicity response in *Vigna unguiculata* L. – a proteomic and transcriptomic study. *Proteomics* 8: 149-159

18. Zhen, Y., et. al. (2007) Comparative proteome analysis of differentially expressed proteins induced by Al toxicity in soybean. *Physiologia Plantarum*, 131:542-554

19. Sugihara, K., Hanagata, N., Dubinsky, Z., Baba, S., Karube, I. (2000) Molecular Characterization of cDNA Encoding Oxygen Evolving Enhancer Protein 1 Increased by Salt Treatment in the Mangrove *Bruguiera gymnorrhiza*. *Plant and Cell Physiology* 41:1279-1285

20. Heinrich Heide1, Henryk M. Kalisz2, a, Hartmut F., (2004) The oxygen evolvinig enhancer protein 1 (OEE) of photosystem II in green algae exhibits thioredoxin activity. *Journal of Plant Physiology* 161: 139-149

21. Yamamoto, Y. (2001) Quality control of photosystem II. *Plant and Cell Physiology* 42(2): 121-128

22. Chini, A., Grant, J. J., Seki, M., Shinozaki, K., Loake, G. J. (2004) Drought tolerance established by enhanced expression of the CC-NBS-LRR gene, ADR1, requires salicylic acid, EDS1 and ABI1. *The Plant Journal* 38: 810-822

23. Zhang, W., et.al. (2000) Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants. *Planta* 210: 970-978

24. Braco, A.T., et.al. (2004) Expression and purificaton of the recombinant SALT lectin from the rice (*Oryza sativa* L.) *Protein Expression and Purification* 33: 34-38

25. Antonyuk, L.P., Evseeva, N.V. (2006) Wheat Lectin as a Factor in Plant-Microbial Communication and a Stress Response Protein. *Microbiology* 75(4): 470-475.

26. Suseelan, K.N., Mitra, R., Pandey, R., Sainis, K.B., Krishna, T.G. (2002) Purification and characterization of a lectin from wild sunflower (*Helianthus tuberosus* L.) tubers. *Archives of Biochemistry and Biophysics* 407: 241-247

27. Yamazaki, D., Motohashi, K., Kasama, T., Hara, Y., Hisabori, T. (2004) Target Proteins of the Cytosolic Thioredoxins in *Arabidopsis thaliana*. *Plant Cell Physiol.*45(1): 18-27

28. Kocheva, K. V., Georgiev, G. I., Vunkova-Radeva, R. V. (2007) Contribution of mineral nutrition to the reponse of barley seedlings to polyethyleve glycol-induced mild water stress. J. Plant Nutr. Soil Sci. 170:392-397

29. Chakrabarti, N., Mukherji, S. (2003) Alleviation of NaCl Stress by Pretreatment with Phytohormones in *Vigna radiate*. *Biolgia Plantarum*. 46(4):589-594.

30. Gong, H., Zhu, X., Chen, K., Wang, S.,Zhang, C. (2005) Silicon alleviates oxidative damage of wheat plants in pots under drought. *Plant Science*. 169:313-321

31. Geiken, B., Masojidek, J., Rizzuto, M., Pompili, M. L., Giardi, M. T. (1998) Incorporation of [^{35}S] methionine in higher plants reveals that stimulation of the D1 centre II protein turnover accompanies tolerane to heavy metals. *Plant, Cell and Environment* 21:1265-1273

32. Peumans, W.J., Van Damme, E.J., Barre, A., Rouge, P. Classification of plant lectins in families of structurally and evolutionary related proteins, *Adv. Exp. Med. Biol*. 491 (2001) 27– 54.

33. Van Damme, E.J., Peumans, W.J., Barre, A., Rouge, P. (1998) Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. *Crit. Rev. Plant Sci.* 17:575–692.

34. Xu, X., Abo, M., Okubo, A., Yamazaki, S. (2001) Salt-Stress-Responsive Membrane Proteins in *Rhodobacter sphaeroides* f. sp. *denitrificans* IL106. *Journal of Bioscience and Bioengineering* 91(2):228-230

35. Ribeiro, A. S., Souza, M. O., Scofano, H.M., Creczynski-Pasa, T. B., Mignaco, J. A. (2007) Inhibition of spinach chloroplast F_0F_1 by an Fe^{2+}/ascorbate/H_2O_2 system. *Plant Physiology and Biochemistry* 45:750-756

36. Ouvrard, O., Cellier, F., Ferrare, K., Tousch, D., Lamaze, T., Dupuis, J., Casse-Delbart, F. (1996) Identification and expression of water stress and abscisic acid-regulated genes in a drought-tolerant sunflower genotype. *Plant Molecular Biology*. 31 : 819-829

37. Fujita, M., Fujita, Y., Maruyama, K., Seki, M., Hiratsu, K., Ohme-Takagi. (2004) A dehydration-induced NAC protein, RD_26, is involved in a novel ABA-dependent stress-signaling pathway. *The Plant Journal*. 39(6): 863-876.

38. Elbekai, R. H., El-Kadi, A.O.S. (2004) Modulation of aryl hydrocarbon receptor-regulated gene expression by arsenite, cadmium, and chromium. *Toxicology* 2002:249-269

39. Gang, D.R., Kasanhara, H., Xia, Z., Mijnsbrugge, K.V., Bauw, G., Boejan, W., Montagu, M.V., Davin, L.B., Lewis, N.G. (1999) Evolution of Plant Defense Mechanisms.

40. Oh, T. J., Kim, I.G., Park, S.Y., Kim, K.C., Shim, H.W. (2002) NAD-dependant malate dehydrogenase protects against oxidative damage in *Escherichia coli* K-12 through the action of oxaloacetate. *Environmental Toxicology and Pharmacology* 11:9-14

41. Chen, H. C., Klein, A., Xiang, M. H., Backhaus, R. A., Kuntz, M.,(1998) Drought and wound induced expression in leaves of a gene encoding a chromoplast carotenoid associated protein *Plant J.*, *14*, 317–326

42. Rey, P., Gillet, B., Romer, S., Eymery, F. *et al.*(2000) Over-expression of a pepper plastid lipid-associated protein in tobacco leads to changes in plastid ultrastructure and plant development upon stress, *Plant J.*, *21*, 483–494.

43. Langenkamper, G., Manac'h, N., Broin, M., Cuine, S. et al.(2001), Plants and the Environment. Accumulation of plastid lipid-associated proteins (fibrillin/CDSP34) upon oxidative stress, ageing and biotic stress in Solanaceae and in response to drought in other species. J.Exp. Bot. 2001, 52, 1545–1554.

44. Jones, Alexandra M. E., Bennett, Mark H., Mansfield, John W., Grant, Murray (2006) Analysis of the defence phosphoproteome of *Arabidopsis thaliana* using differential mass tagging. *Proteomics* 6(14), 4155-4165.

45. Wang, W., Vinocur, B., Shoseyov, O., Altman, A. (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *TRENDS in Plant Science* 9 (5): 244-252

46. Wegele, H., Muller, L., Buchner, J (2004). Hsp70 and Hsp 90-arelay team for protein folding. *Rev Physiol Biochem Pharmaco*l 151:1–44

47. Damelin, L.H., Vokes, S., Whitcutt, J.M., Damelin, S.B., Alexander, J.J. (2000) Hormesis: a stress response in cells exposed to low levels of heavy metals. *Human & Experimental Toxicology* 19 (7): 420-43

48. Hu, X. and Reddy, A.S. (1997) Cloning and expression of a PR5-like protein from *Arabidopsis*: inhibition of fungal growth by bacterially expressed protein. *Plant Mol. Biol*. 34, 949–959.

49. Selitrennikoff, C.P. (2001) Antifungal proteins, *Appl. Environ. Microbiol*. 67: 2883–2894.

50. Brandazza, A., Angeli, S., Tegoni, M., Cambillau, C., Pelosi, P. (2004) Plant stress proteins of the thaumatin-like family discovered in animals. *FEBS letters* 572: 3-7.

51. Menu-Bouaouiche L., Vriet C., Peumans W.J., Barre A., Van Damme E.J., Rouge P., (2003) Plant a-amylase inhibitors and their interaction with insect a-amylases. *Biochimie*, 85:. 123-131

52. Krapp, A.R., Tognetti, V.B., Carrillo, N., Acevedo, A. (1997) The role of ferredoxin-$NADP^+$ reductase in the concerted cell defense against oxidative damage. Eur. J. Biochem 249:555-563

53. Palatnik, J.F., Valle, E.M., Carrillo, N. (1997) Oxidative stress causes ferredoxin-NADP+ reductase solubilization from the thylakoid membranes in methyl viologen treated plants. *Plant Physiology*. 115(4):1721-1727

54. Klahre, U., Friedrich, E., Benedikt, K., Louvard, D, Chua, N. (2000) Villin-Like Actin Binding Proteins are Expressed Ubiquitiously in *Arabidopsis*. *Plant Physiol* 122:35-48

55. Meek, L., Martin, R.C., Shan, X., Karplus, P.A. (2008) Isolation of Legume Glycosyltransferases and Active Site Mapping of the *Phaseolus lunatus* Zeatin O-glucosyltransferace ZOG1. *J. Plant Growth Regul.* 27:192-201

56. Havlova, M., Dobrev, P.I., Moyka, V., Storchova, H., Libus, J., Dobra, J., Malbeck, J., Guadinova, A., Vankova, R. (2008) The role of cytokinins in responses to water deficit in tobacco plants over-expressing trans-zeatin O-glucosyltransferase gene under 35S or SAG12 promoters. *Plant, Cell and Environment* 31:341-353

57. Wang, J.-T., Gould, J.H., Padmanaghan, V., Newton, R.J. (2003) Analysis and Localization of the Water-Deficit Stress-Induced Gene (lp3)

58. Nogueira, F.T.S., Schlögl, P. S., Camargo, S. R., Fernandez, J. H., De Rosa, V. E., Pompermayer, P., Arruda, P. (2005) SsNAC23, a member of the NAC domain protein family, is associated with cold, herbivory and water stress in sugarcane. *Plant Science* 169(1):93-106

59. Oh, S.-K., Lee, S., Yu, S.H., Choi, D. (2005) Expression of a novel NAC domain-containing transcription factor (CaNAC1) is preferentially associated with incompatible interactions between chili pepper and pathogens. *Planta*. 222(5):876-887

60. Gang, P., (2008) Isolation of an 1-aminocycloporpane-1-carboxlyate oxidase gene from mulberry (*Morus alba* L.) and anaylsis of the function of this gene in plant development and stresses response. *Journal of Plant Physiology*. 165:1204-1213

61. Wang K.L., Li H., Ecker J.R. (2002) Ethylene biosynthesis and signaling networks. *Plant Cell*. 14:S131–51.

62. Grunwald, I., Heinig, I., Thole, H.H., Neumann, D., Kahmann, U., Kloppstech, K., Gau, A.E. (2007) Purification and characterization of a jacalin-related, coleoptiles specific lectin from *Hordeum vulgare*. *Planta*. 226:225-234

63. Barth, O., Zschiesche, W., Siersleben, S., Humbeck, K. (2004) Isolation of a novel barley cDNA encoding a nuclear protein involved in stress response and leaf senescence. *Physiologia Plantarum*.31(6): 473-482

64. Wang, J., Zhang, H., Huang, R. (2007) Expression Analysis of Low Temperature Responsive Genes in *Eupoatorium Adenophorum Spreng* Using cDNA-AFLP. *Plant Mol Biol Rep.* 25:37-44

65. Sainz, G., Carrell, C.J., Ponamarev, M.V., Soriano, G.M., Cramer, W.A., Smith, J.L. (2000) Interruption of the Internal Water Chain of Cytochrome f Impairs Photosynthetic Function. *Biochemistry*. 39 :9164-9173

66. Zhu, Jian-Kang (2001) Plant salt tolerance. *Trends in Plant Science*. 6(2): 66-71

CHAPTER II

STUDY ON ADDUCTION REACTIONS IN CHYMOTRYPSIN'S ACTIVE SITE

2.1. Background and Significance

PAHs, composed of hydrogen and carbon, are hydrophobic pollutants made up of two or more fused benzene rings. In addition, some may contain alkyl and other substituent groups. They are formed as the byproducts of the pyrolysis and incomplete combustion of organic matter [1]. In industrialized countries, the major sources of PAHs are petroleum products such as oil, coal, tar, creosote, and vehicle emissions [2], though they are also formed as byproducts of energy usage, heating, and cooking [3]. Human exposure is predominantly through breathing, ingestion, dermal contact, and cigarette smoke [4]. Though inherently found in the environment as a consequence of natural phenomena (e.g. volcanic eruptions, forest fires) their presence has greatly increased since the advent of industry, the mainstream use of fossil fuels as power and heat sources, the development of the combustion engine, and the grand-scale deforestation of natural habitats. The thermal stability of PAHs coupled with the wide array of potential sources has lead to the ubiquitous nature of the pollutants' dispersal; the range of concentration of BaP throughout the environment is shown in Fig. 2-1.

Air	1.3–500 ng/m^3
Soil	0.8 ng/kg – 100 mg/kg
Tap water	2.5–9 ng/l
Surface water	130–500 ng/l
Plants	Up to 150 μg/kg
Food	0.1–20 μg/kg

Figure 2-1. Conc. of Benzo[a]pyrene detected in air, soil, water, and plants and food [3].

Research has been undertaken to *qualify* and *quantify* the carcinogenic/mutagenic nature of PAHs. In an attempt to utilize the protein-PAH adduct as a biomarker, blood proteins from industrial workers as well as urban and rural residents of many industrialized countries have been evaluated, with correlating studies being conducted on control patients exposed to similar contaminant levels [5-14]. Attempts to correlate the level of PAH exposure to cancer rates and tumor growth in lab animals have also been conducted, yielding a positive correlation between PAH exposure and tumor development [15]. Of these studies the focus on protein-PAH adduct formation has primarily been on hemoglobin and albumin; due to the vast amount of literature present, the ease of isolation and their large quantity per gram of blood.

The qualification of PAH toxicity has been attributed to the PAHs' metabolic cycle in the body. Examination of the biochemical degradation steps of the PAHs' metabolism in humans have found that CPY1A1 is the principle enzyme involved in their initial monooxygenation [3]. This enzyme is responsible for the PAHs' "activation" to an electrophilic species capable of reacting with nucleophilic sites on DNA or Protein. The various metabolites of the CPY1A1 reaction with BaP can be seen in Fig. 2-2.

Though research into the effects that the PAH – DNA adduct has on helix structure and its subsequent isolation could undeniably pin point genetic mutation/mistranslation events there are two distinct drawbacks to studying the DNA adduct. Firstly, DNA inherently has a specific repair mechanism not present in proteins that may allow for adduct removal and proper helix restoration. Secondly the vast amount of intron and satellite DNA in exposed cells makes risk assessment of toxins very difficult to quantify. The seemingly rare possibility of an adduct formation event occurring within an exon as

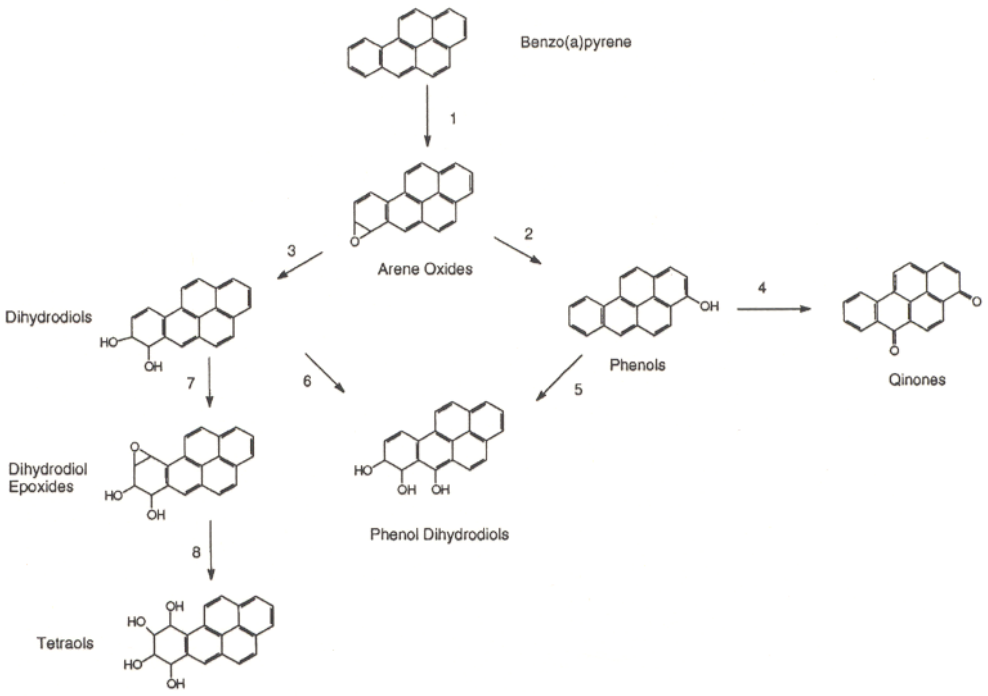


Figure 2-2. Metabolism of Benzo [a] Pyrene [1].

opposed other “non-transcribed/translated” genetic material makes the *immediate* mutagenicity of the DNA-PAH adduct difficult to correlate with abnormal cellular function.

The protein adducts that are formed upon contact with the CPY-1A1 altered PAH in the cellular or extra-cellular environment lack a specific repair mechanism and allow for a decisive determination of, if any, altered proteolytic function. These altered proteins can still be involved in biochemical reactions, only now in a non-homologous state with their unaltered forms for the remainder of their “lifetime”. The effects that these alterations may have on the proteins kinetics and structural conformation are the objectives that this research hopes to address.

This study will follow suit with past work done in the preparation and purification of protein-PAH adducts and proposes a next step into understanding protein alterations. Past research has focused on two separate methods for determining adduct formation. The initial research focused on determining adduct placement on the primary amino acid sequencing of the protein by lysing the denatured protein and piecing together the observed fragments [16-17]. Further methods have isolated the specific residues involved in adduct formation by hydrazinolysis [18-20]. The end goal of these studies was to quantify the level of adduct formation with the subjects exposure, again attempting to utilize the protein adduct as a “bio-marker”.

To additionally qualify an adducts effect on protein function an analysis of protein kinetics will be performed. For this analysis a good protein model needs to be chosen. Contributing to this are the findings of prior experimentation into protein-PAH adduct formation that revealed the high affinity of BaP diol epoxides for the imidazole nitrogen

of histidine residues [16–22], and the potential affinity that electrophiles have in nucleophilic substitution reactions within receptor binding sites [23]. This alkylation affinity of the metabolically activated PAH leads to the choice of chymotrypsin (CHY) for the study. CHY, a member of the serine peptidase class of enzymes, has a catalytic core whose functionality is based in part on a centrally located histidine residue Fig 2-3. The high potential for disruptions of the natural function of these proteins by PAH adduct formation with the catalytic core's histidine residue makes this an excellent enzyme for analysis.

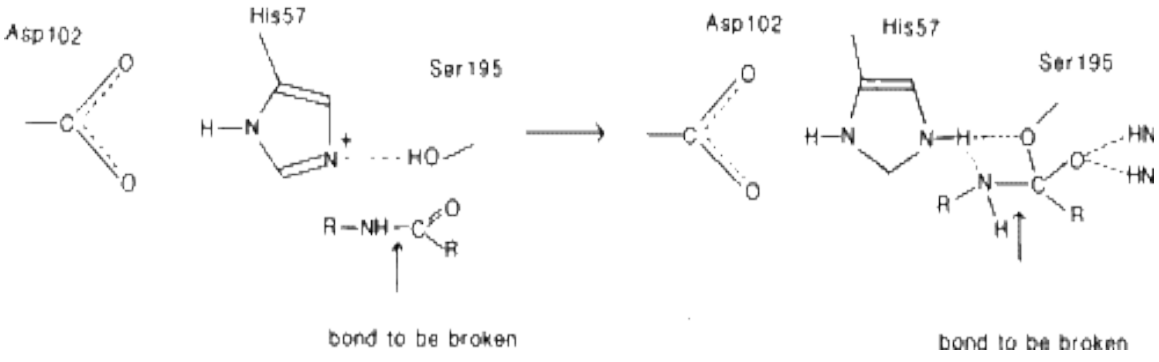


Figure 2-3. Catalytic core of Chymotrypsin, a model Serine Peptidase.

To determine and optimize conditions for CHY-PAH adduct analysis a known CHY-adduct will be assessed, the irreversible inhibition of chymotrypsin by exposure to tosyl phenylalanyl chloromethylketone (TPCK). TPCK irreversibly inhibits CHY by binding to histidine 57 located in the center of CHYs active sites catalytic triad (Fig. 2-4). This reaction is specifically mediated by CHY's hydrophobic pocket located adjacent to its active site. This pocket facilitates the placement and orientation of the aromatic side

groups on TPCK, much as in CHY's normal enzymatic function. When oriented in the hydrophobic pocket TPCK's chlorinated methyl group reacts with the active site histidine. The pocket's hydrophobic character and facilitation of substrate orientation for nucleophilic substitution reaction is potentially a viable course of reaction after exposure of the serine proteases to PAH metabolites.

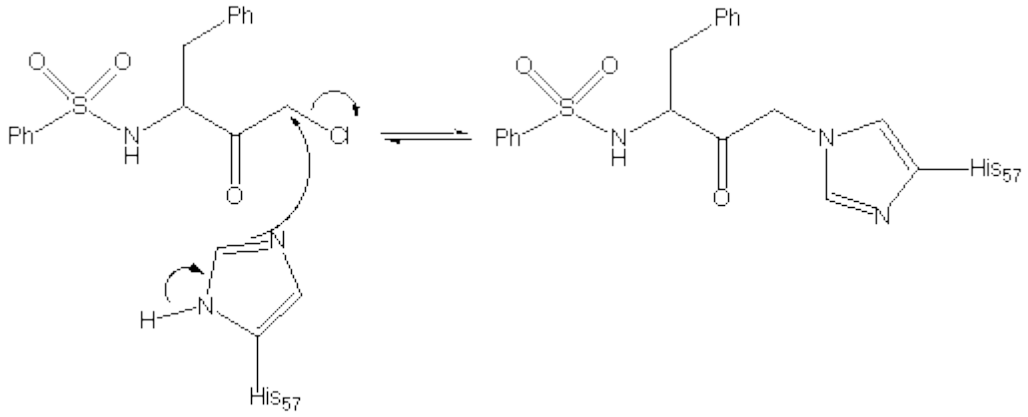


Figure 2-4. TPCK reaction/inhibition of CHY.

This hydrophobic pocket on CHY has also been studied using a theoretical software program for "goodness of fit" for a large group of hydrophobic compounds as reversible inhibitors [28]. This work is of interest in applying the effects of PAH and PAH metabolites on CHY kinetics in a real time experiment. Preliminary trials of CHY + PAH solutions had some interesting effects on absorbance data, particularly the 2 and 3 ring species, and their reduction of CHY Kinetics.

Additional theoretical support is provided by a series of calculations that describe a phenomenon wherein highly reactive metabolites, such as BPDE, follow a very specific

site-mediated reaction pathway instead of simply reacting with any available nucleophile [23]. Various experiments have been cited as evidence involving electrophiles, PAH species in particular, that involve a non-covalent interaction prior to a substrate-ligand binding type reaction. Site-specific reactions such as that proposed for CHY-BPDE are typically explained by chemical reactivity, steric accessibility or active site proximity to nucleophilic residues, but no individual explanation has been broadly successful in describing why only certain residues are targeted, and why only those found in specific positions as opposed to that residue type within the entire protein.

A kinetic scheme for the reaction of small electrophiles with proteins can also be established. Assigning k_a and k_{-a} as the rate constants for reversible, noncovalent protein-electrophile binding, the ratio of k_a to k_{-a} then defines the equilibrium association constant. k_1 is representative of the rate constant for reaction of an electrophile E with a surface-accessible nucleophile N_1. It is assumed that k_1 is not substantially different from the rate constant that characterizes the same reaction with a free amino acid or any other small molecule that has N_1 as part of its structure. k_2 is representative of the rate constant for covalent bond formation between E and nucleophile N_2, which is located within the binding site for E. Some values for k_a, k_{-a}, and k_1 have been measured or estimated, but values for k_2 are largely unknown (Fig. 2-5) [23].

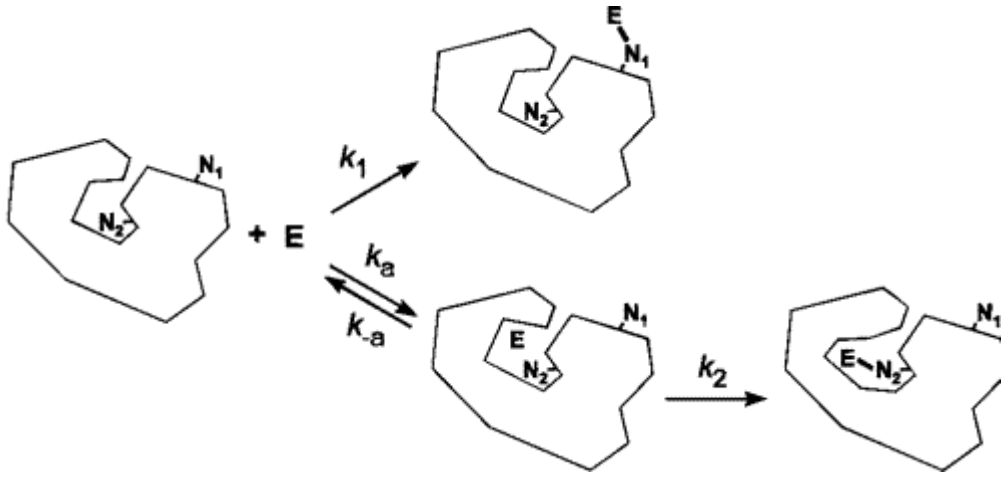


Figure 2-5. Representative diagram of potential equilibrium reactions between an electrophile E and neuleophilic residues N_1 and N_2 [23].

Equilibrium association constants are driven by association kinetics that are extremely fast as compared to nucleophilic substitution reactions [23]. Implying a potentially high kinetic favoring of the 2 step, non-covalent - covalent binding, over the one step, unassisted reaction. To give further evidence of this case the work describes the intercalation of Benzo[a]Pyrene-diol-epoxide (BPDE) with DNA occurring on a millisecond timescale, with a calculated $K_{eq} = 10^5$ $M^{-}s^{-}$ for the BDPE-DNA association. This suggests that even low affinity binding will probably exhibit rapid association kinetics, regardless of the strength of K_{eq}.

Applying this theory of a "docking" based mechanism for this reaction type to a albumin-BPDE adduct wherein a known reaction occurred in the 1B sub domain, a low affinity-binding domain, gives further evidence of its plausibility. The adduct products obtained were a His^{146} adduct with the (+) enantiomer and an Asp^{187} or Glu^{188} adduct with the (-) enantiomer. Upon determination of Albumins structure via x-ray

crystallography it was shown that these residues are in close proximity to each other and the selectivity is understandable given that Albumin only exists as a single enantiomer itself. This example elucidates the intercalation event that coordinates the PAH, leaving only the front or back side reaction pending the specific BPDE enantiomer in question [18].

Applying these theoretical calculations to CHYs histidine residue located in the center of its catalytic triad associated with the presence of a hydrophobic pocket within the receptor binding site, one specifically "designed" to accommodate the large hydrophobic residues; tyrosine, tryptophan and phenylalanine, allows for a possibility of the CHY-BPDE adduct formation. Additionally CHY only has 11 surficial nucleophiles, 7 aspartic acid and 4 glutamic acid residues, giving the enzyme a low nucleophilic character, which also implies a likelihood for receptor site mediated binding. Incorporating CHY's small number of external nucleophiles and its hydrophobic pocket adjacent to the active site this theoretical work bolsters the hypothesis that the active site of Chymotrypsin makes a very "enticing" nucleophile for the BPDE specie.

2.2. Methodology

2.2.1. CHY adduct formation

2.2.1.1. CHY-TPCK adduct reaction

10 mg of protein was dissolved in 1 mL of 10 mM sodium/potassium phosphate buffered normal saline, pH 7.4. To this is added 100 μL of an approximately 1 mM solution of TPCK in 2:1 methanol (MeOH)/H_2O. The reaction is allowed to proceed for 1 h at room temperature.

2.2.1.2. Proposed CHY-PAH adduct reaction

The BPDE-protein reaction is a modification of the methods described for the reaction of ± (anti) r-7, t-8- dihydroxy- t-9, 10-epoxy-7, 8, 9, 10 tetrahydro Benzo[a]pyrene with human hemoglobin, human serum albumin and bovine serum albumin [18].

100 mg of protein will be dissolved in 4mL of 10 mM sodium/potassium phosphate buffered normal saline, pH 7.4. To this is added 1 mg of the BaP diol epoxide in 100 μL tetrahydrofuran (THF). The reaction is allowed to proceed for 2 h at 37 °C. The reaction is terminated by extraction with ethyl acetate and the remaining solution is then washed 3 times with 4 mL aliquots of n-butanol. 10 mL diethyl ether is then added to the organic phase extract and it is dried under a stream of N_2 [18].

2.2.2. Whole protein adduct analyses

2.2.2.1. LC separation and QQQ analysis:

Separation was performed by high pressure liquid chromatography (HPLC) using a reversed phase Vydac protein C-4 100 x 0.3 mm, 5 μm, 300A column. Solvent A consists of H_2O with 1 % acetonitrile (ACN) and 0.1 % formic acid. Solvent B was an ACN solution containing 0.1 % formic acid. 50 mg of the dried protein extract was dissolved in 2 mL 50:50 H_2O/ACN for evaluation. The column is pre-equilibrated with 10 % solvent B. Initial linear gradient is from 10 – 50 % in 30 min, followed by a 5 min isocratic elution, followed by a second gradient segment from 50 – 85 % acetonitrile solution in 10 min followed by a 15 min isocratic elution.

Average molecular masses were measured by triple quadrupole mass spectrometry to determine adduct formation, and number of adducts/molecule if multiple adducts are observed. Protein mass is determined using Mass Lynx software package. Settings for whole protein were used, CHY samples were prepared and both infused directly in the MS by controlled flow needle injection system or introduced by HPLC. Samples introduced via HPLC injection were analyzed using the described gradient ramp. A special apparatus was constructed to control back pressure and optimize LC conditions to maintain post column flow of 45μL (Appendix B, Fig. B-1). The pump was set at approximately 0.60 mL / min to maintain 55 psi on column. From these samples ESI polarity, Cone Voltage, Capillary Ion Flow, HV lens, scan time and data range were optimized on analysis for sample detection.

To verify LC QQQ MS retention time reproducibility control samples were analyzed at the start and finish of assigned MS research days. Column and LC-MS conditions were evaluated using a tryptic myoglobin digest sample. Tryptic Digest Conditions: Scan Range 280-1500, 45 cone voltage, 2.5 step, LC gradient 0-80 min 2-70% B, 80-90 min 70-90%B 110 min end of run.

2.2.2.2. LC separation and Ion Trap analysis:

HPLC ESI-Ion Trap MS was carried out by injecting the sample offline into a secondary LC loop using a Loading Phase of 2% ACN in H_2O with 0.2% HCOOH and collecting the molecules in a C-4 Cap Trap (MicromResources). The sample was washed for 10 min in the Cap Trap and then main column switch brings the CapTrap online into the primary Agilent LC mobile phase ramp; the column switch is diagramed in appendix B, Fig. B-2.

Post isocratic cap-trap enrichment, samples were analyzed under a gradient ramp, as described above. LC flow conditions of 5µL / min to maintain a graded 75 – 30 psi on column corresponding with net mobile phase polarity. From these samples ESI polarity, cone voltage, capillary ion flow, HV lens, nebulizer psi, flow rate of dry gas, temperature, ESI voltage, scan mode, range, speed, ICC Smart Target, maximum accumulation time and scan range were optimized on analysis for sample detection.

As a control for LC ESI-Ion Trap the 66 kDa albumin marker from SDS-PAGE gel samples was chosen. This control was used to optimize protein parameters and used for all Ion Trap experiments.

2.2.3. Hydrazinolysis analysis

2.2.3.1. Hydrazinolysis

Ten mg of the dried protein extract was dissolved in 0.5 mL hydrazine and kept at 100 °C in a reflux condenser. After 15 h the solution was chilled in an ice bath. The solution was then applied on water-flushed C-18-SPE column, prewashed with MeOH. The column was then sequentially washed with 2 mL water and 2 mL MeOH/water (3:7). The BPDE products were then eluted with 2 mL MeOH/water (8:2) and the solvent was evaporated to dryness under a stream of N_2 [23]. The sample was reconstituted in 100 µL 50 % ACN / 50 % H_2O for injection.

2.2.3.2. LC separation and QQQ analysis

Separation of hydrazinolysis products was performed by HPLC using a reversed phase Vydac C-18 column, 100 x 0.3 mm, 5 µm, 300A. Solvent A consists of H_2O with 1% ACN and 0.1 % formic acid. Solvent B was an ACN solution containing 0.1 % formic acid. 10 µL of the peptide solution was injected for evaluation. The column was pre-equilibrated with 10 % acetonitrile solution. Initial linear gradient is from 10 – 40 % acetonitrile solution in 30 min, followed by a 5 min isocratic elution, followed by a second gradient segment from 40 – 60 % acetonitrile solution in 10 min and by a third gradient to 90 % acetonitrile solution in 10 min, with 5 min of isocratic elution.

HPLC ESI-QQQ MS was carried out by splitting the outlet of the HPLC column in two directions: one to the UV detector and the other to mass spectrometer with electrospray ion source. A second apparatus was constructed to control back pressure on column for sample injection with outlet split to maintain 45 µL / min post column flow

rate (shown in appendix B, Fig. B-3). UV detection was set at 345nm, specific to BPDE products. Average molecular masses measured by mass spectrometry was be used to determine adduct formation and initial amino acid identity of the amino acid residue-PAH adducts.

For optimization of experimental conditions and QQQ settings for small molecules, histidine, B[a]P, a 7, 8 diol Benzo[a]pyrene(BPD), and TPCK solutions were prepared and infused directly by controlled flow needle injection system. Simple QQQ mass spectra for these samples can be seen in Fig. 2-6. From these samples ESI polarity, Cone Voltage, Capillary Ion Flow, HV lens, scan time and data range were optimized for sample infusion. Parameters for collision induced dissociation (CID) and formation of daughter fragments were also optimized as seen in the infusion experiment of TPCK (Fig. 2-7).

Optimization of parameters for small molecules by sample introduction via LC injection was also performed for method development. Samples were injected under the gradient described at 0.625 mL /min, 48 psi to maintain post C-18 column flow of 45 µL. For tandem MS analysis fractions were collected at 1 min intervals for subsequent infusion and CID daughter analysis.

2.2.3.3. LC Separation and Ion Trap analysis

Separation of hydrazinolysis product was performed by HPLC using a reversed phase Vydac C-18 column, 100 x 0.3 mm, 5 µm, 300A. Solvent A consists of H_2O with 1% ACN and 0.1 % formic acid. Solvent B was an ACN solution containing 0.1 % formic acid. 10 µL of the peptide solution is injected for evaluation. The column was pre-

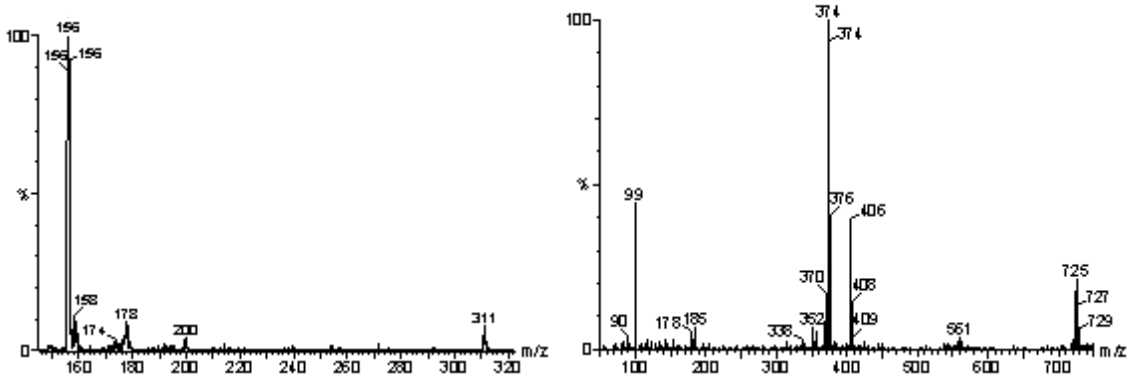


Figure 2-6. QQQ Mass spectra of histidine and TPCK, optimization of small molecule analysis.

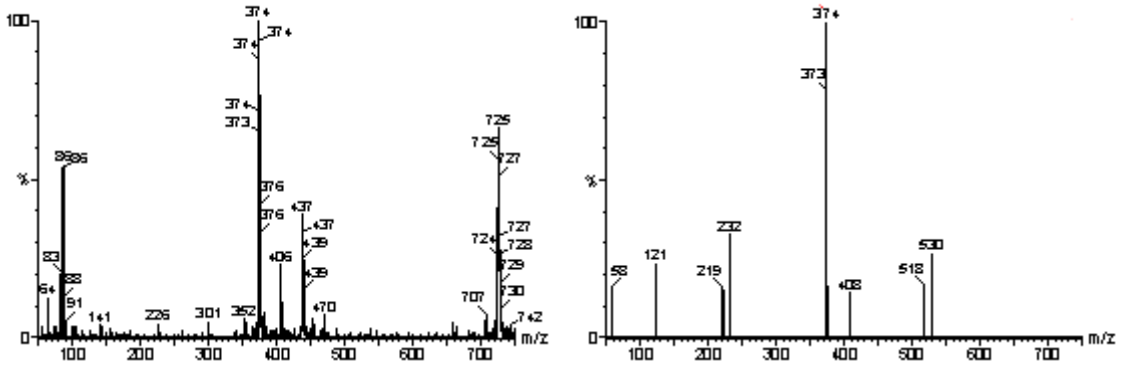


Figure 2-7. QQQ Mass spectra of TPCK and daughter analysis of parent molecule, optimization of small molecule analysis CID.

equilibrated with 10 % acetonitrile solution. Initial linear gradient is from 10 – 40 % acetonitrile solution in 30 min, followed by a 5 min isocratic elution, followed by a second gradient segment from 40 – 60 % acetonitrile solution in 10 min and by a third gradient to 90 % acetonitrile solution in 10 min, with 5 min of isocratic elution.

HPLC ESI-Ion Trap MS was carried out by injecting the sample "offline" into a secondary LC loop using a Loading Phase of 2% ACN in H_2O with 0.2 % HCOOH and collecting the molecules in a C-18 Cap Trap (MicromResources). The sample was washed for 10 minutes in the Cap Trap and then main column switch brings the CapTrap "online" into the primary Agilent LC mobile phase ramp (appendix B, Fig. B-2).

For optimization of experimental conditions and Ion-Trap settings for small molecules histidine, B[a]P, a 7, 8 diol Benzo[a]pyrene(BPD), and TPCK solutions were prepared and introduced by HPLC. Direct infusion was not utilized for these samples to allow for additional sample enrichment and to remove any buffer prior to column and trap exposure. Post isocratic cap-trap enrichment, samples were analyzed under the described gradient ramp at a LC flow rate of 5µL / min to maintain < 75 psi on column. From these samples ESI polarity, cone voltage, capillary ion flow, HV lens, nebulizer psi, flow rate of dry gas, temperature, ESI voltage, scan mode, range, speed, ICC Smart Target, maximum accumulation time and scan range were optimized for sample detection.

2.2.4. Enzymatic digest analysis

Trypsin was obtained from Promega or Pierce. Protein digestion was performed on 1-D SDS-PAGE gel. Gel bands of interest were excised, cut into approximately $1mm^3$ pieces and stored at -20°C until digestion. For digestion gel pieces were

washed/destained with 200 µL wash reagent (50% Et-OH / 5% acetic Acid / 45% H_2O), for 1 hr. Wash was discarded and wash step repeated for an additional hour. Gels were then dehydrated by adding 200 µL acetonitrile for 10 min. Acetonitrile was discarded and samples were rehydrated in ammonium bicarbonate for 5 min. Ammonium bicarbonate was discarded and samples are again dehydrated with acetonitrile for 10 min. Acetonitrile was discarded and the samples were dried in SpeedVac for approximately 3-5 min. Samples are rehydrated in 50 µL trypsin, prepared to approximately 20 ng/µL and incubated at room temperature (23-25°C) overnight.

After digestion step samples were microfuged and the supernatant digest mixture collected. To the remaining gel pieces 30 µL extraction reagent was added (50 % ACN / 5 % formic acid / 45 % Water), sample was briefly vortexed and left for 10 min. The extraction mixture was removed and added to the digest mixture, the extraction step is repeated once. Combined supernatants were placed in SpeedVac to reduce volume to approximately 10 µL, about 45 min – 60 min. Sample volumes were reconstituted to approximately 25 µL total volume with 1 % acetic acid for LC-MS analysis.

HPLC ESI-Ion Trap MS was carried out by injecting the sample offline into a secondary LC loop using a loading phase of 2% ACN in H_2O with 0.2 % HCOOH, the molecules are enriched on a C-18 CapTrap (MicromResources). The sample was washed for 10 minutes in the CapTrap and then main column switch brings the CapTrap online into the primary Agilent LC mobile phase ramp as described in 2.2.2.3. Ion Trap conditions were set to those determined for small molecule analysis and optimized on analysis as needed.

2.2.5. Functional analysis of protein

The activity of chymotrypsin was based on the following reaction [24]:

N-Benzoyl-L-Tyrosine Ethyl Ester + H_2O → N-Benzoyl-L-Tyrosine + Ethanol

The reaction velocity was determined by measuring an increase in absorbance at 256 nm resulting from the hydrolysis of benzoyl-L-tyrosine ethyl ester (BTEE). One unit hydrolyzes one micromole of BTEE per min at pH 7.8 and 25ºC under the specified conditions. The reaction was continuously observed in a spectrophotometer at 256 nm, 25º.

Ten mg of protein extract was dissolved in 10 mL 0.001 N HCl. The protein solution was then diluted with 0.001 N HCl to give 10-30 μg protein/mL. Into a cuvette the following reagents were added; 500 μL 0.08 M Tris HCl buffer, pH 7.8 with 0.1 M $CaCl_2$, 467 μL 0.00107 M BTEE in 50 % w/w methanol (63 mL absolute methanol added to 50 mL Mili Q water). The reaction mixture was then incubated in the spectrophotometer at 25ºC for 5 minutes to achieve temperature equilibrium. 33 μL of the protein solution was then added and the absorbance at 256 nm is recorded for 20 min.

2.3. Results & Discussion

2.3.1. QQQ MS characterization of adduct in chymotrypsin

Preliminary model analysis settings for QQQ MS operation were established as described. All analysis was conducted under positive polarity. Cone voltage settings were modified and optimized for whole protein, peptide, or small molecule analysis as a function of optimal scan range. Cone voltage was adjusted accordingly: average cone

voltage: 35 v = 200 m/z – 400 m/z, 40 v = 500m/z, 70 v = 600 m/z – 700 m/z, 90 v = 2000 m/z. A setting of 35 v for hydrazine digest, 70 v for tryptic digest and 75 v for whole protein analysis was selected. HV lens, scan time and data range were selected as a function of sample type, a shorter scan time and data range chosen for the whole protein analysis and a longer scan time with a wider data range for the peptide/small molecule analysis.

Whole protein samples of CHY and CHY-TPCK were analyzed by direct infusion and the LC C-4 methods described (Fig. 2-8). Hydrazine protein digests were then analyzed using the described gradient with LC C-18 separation to establish retention times and develop fractionation procedure. Comparative data of CHY and the CHY-TPCK digest chromatograms display high peak elution conservation with an interesting peak seen only in the adduct sample (Fig. 2-9) Of some concern in observing our adduct, was the potential cleavage of the TPCK-Histidine bond due to a Gabriel synthesis type reaction of the Hydrazine with the unknown reactivity of the imidazole bound nitrogen. This potential reaction could limit detection of the model adduct, and our proposed PAH adduct; based on current literature no documentation of TPCK-Histidine adduct has been characterized. However our successful observation of the TPCK-Histidine product has verified that any additional reaction by hydrazinolysis, if it occurs, is not complete under these conditions.

Theoretical amino-acid hydrazide molecular weights were calculated (Table 2-I**)** and compared to QQQ-MS results. The peak observed at 42.6 min in the CHY-TPCK digest not seen in the CHY digest was collected by fractionation and further analyzed by infusion QQQ MS (Fig 2-10) fitting a potential match to the His-TPCK hydrazide.

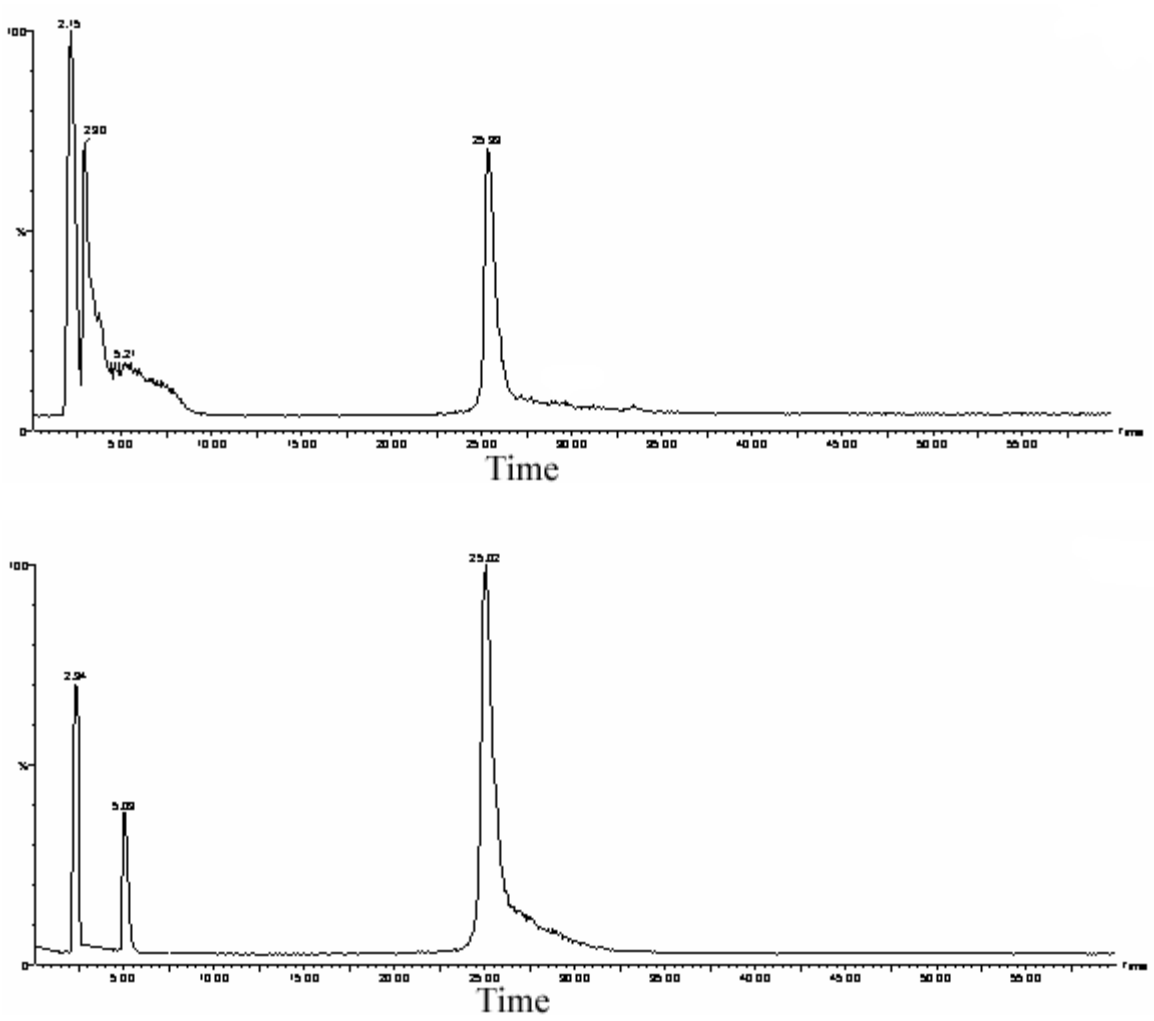


Figure 2-8. LC-MS QQQ Whole CHY-TPCK (top) and CHY (bottom).

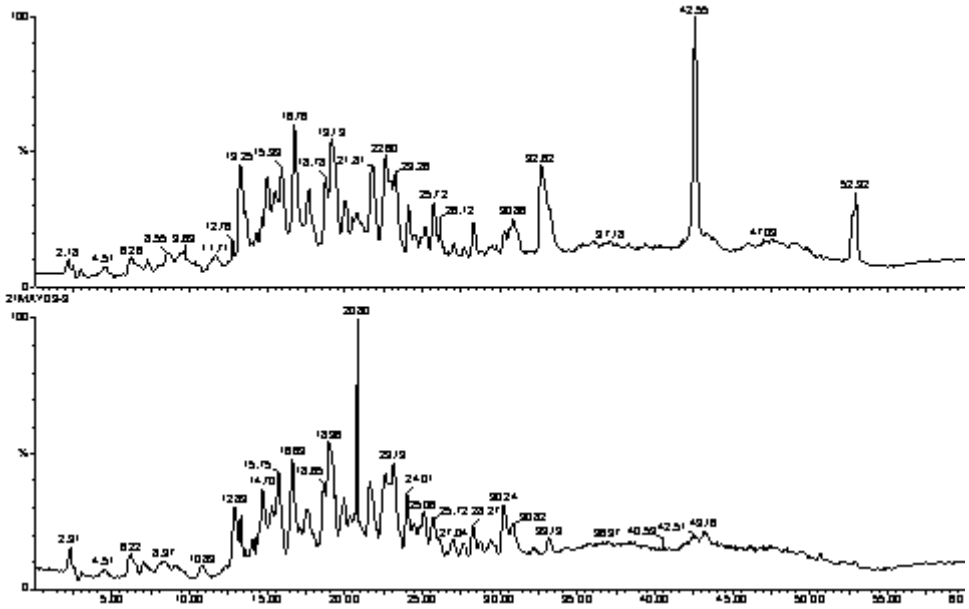


Figure 2-9. LC MS QQQ Separation of hydrazine digests of CHY-TPCK (top) and CHY (bottom).

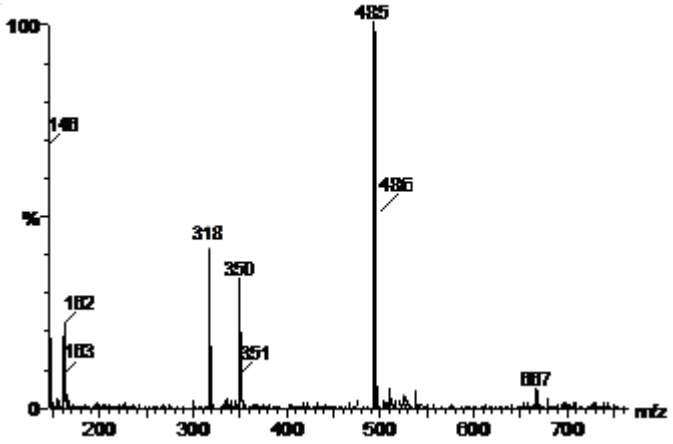


Figure 2-10. LC MS QQQ hydrazinolysis peak spectra at 42.6 min in CHY-TPCK digest corresponding to Histidine – Adduct.

CGVPAIQPV……IVNGEEAVPGSWPWQVSLQDKTGFHFCGGSLINENWVVTAAH

CGVTTSDVVVAGEFDQGSSSEKIQKLKIAKVFKNSKYNSLTINNDITLLKLSTAAS

FSQTVSAVCLPSASDDFAAGTTCVTTGWGLTRY..ANTPDRLQQASLPLLSNTNC

KKYWGTKIKDAMICAGASGVSSCMGDSGGPLVCKKNGAWTLVGIVSWGSSTCS

TSTPGVYA*RVTALVNWVQQTLAAN*

-Underlines refer to Sheet segments and Italics to Helix segment

Residue		# in CHY	Monoisotopic Mass	AA-Hydrazide Mass
Glycine	(G)	22	57.02146	89
Alanine	(A)	22	71.03711	104
Valine	(V)	23	99.06841	131
Leucine	(L)	17	113.08406	145
Isoleucine	(I)	10	113.08406	145
Proline	(P)	9	97.05276	130
Phenylalanine	(F)	6	147.06841	179
Tyrosine	(Y)	4	163.06333	195
Tryptophan	(W)	8	186.07931	219
Cysteine	(C)	10	103.00919	135
Methionine	(M)	2	131.04049	163
Serine	(S)	26	87.03203	119
Threonine	(T)	22	101.04768	133
Lysine	(K)	14	128.09496	$160/161^{+}$
Arginine	(R)	3	156.10111	$188/189^{+}$
Histidine	(H)	2	137.05891	$169/170^{+}$
Asparagine	(N)	13	114.04293	146
Glutamine	(Q)	10	128.05858	160
Aspartic Acid	(D)	9	115.02694	147
Glutamic Acid	(E)	5	129.04259	161

Table 2-I. Secondary structure of chymotrypsin, individual amino acids molecular weight and the molecular weight of the expected amino acid hydrazides.

LC parameters and MS-QQQ settings were continuously monitored to optimize data collection. Key focus was placed on peak resolution and t_R reproducibility for later fraction collection. These continued modifications of digest protocol and refinement of LC-MS parameters allowed for a large amount of accurate data on the digest system to be acquired.. The drawback to this system is that initial data is only MS^1 as the QQQ-MS lacks the specificity to accurately verify compounds without fractionation and subsequent infusion. Arguably samples could be analyzed in repetition and a precisely programmed protocol to alter MS^* settings for each peak of interest could be established. This additional procedure would then be required for each system being observed, presenting the immediate drawback of large sample volume and long MS analysis times. At this point the project was moved to the departments newly acquired the Ion-trap MS and the work on the QQQ was re-focused for the development of ESI-Ion Trap experimentation that would allow for MS* analysis of compounds directly in-line with the initial identification step.

2.3.2. LC ESI-MS Ion characterization of adduct in chymotrypsin

Whole protein samples of CHY and CHY-TPCK were analyzed by the LC C-4 methods described. Hydrazine protein digests of CHY and CHY-TPCK were then analyzed using the described gradient with offline captrap washing and enrichment followed by C-18 separation and MS. Comparative data of CHY and the CHY-TPCK are shown in (Fig. 2-11). MS^2 spectra were acquired for peaks and a potential Histidine-TPCK hydrazide peak was observed at a similar retention time, 46 min (Fig. 2-12).

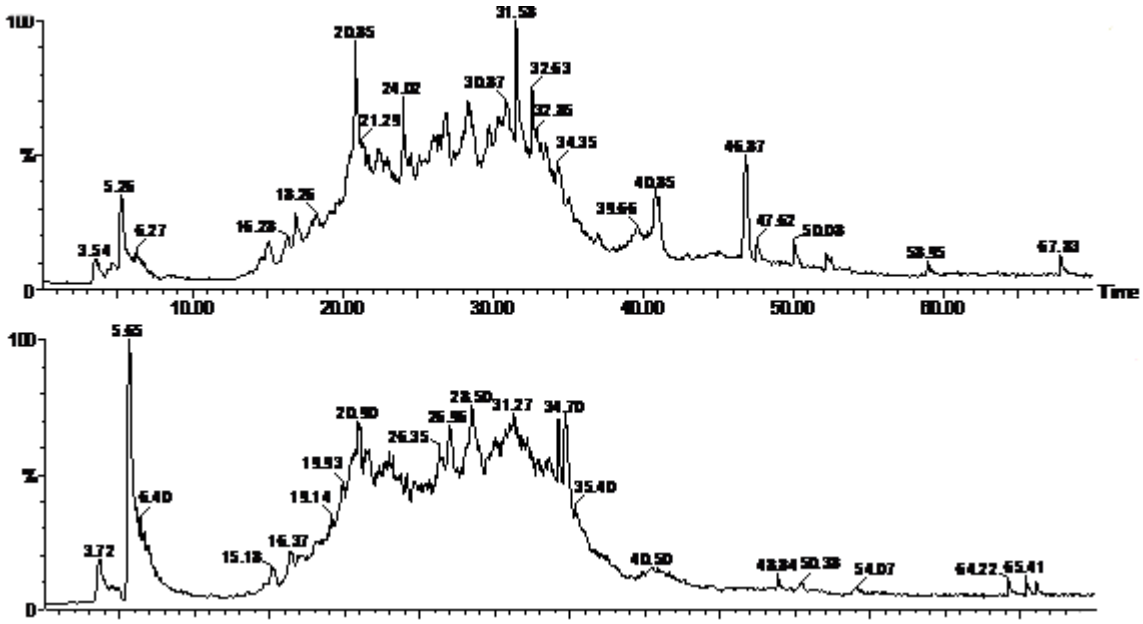


Figure 2-11. LC MS Ion Trap Separation of hydrazine digests of CHY-TPCK (top) and CHY (bottom).

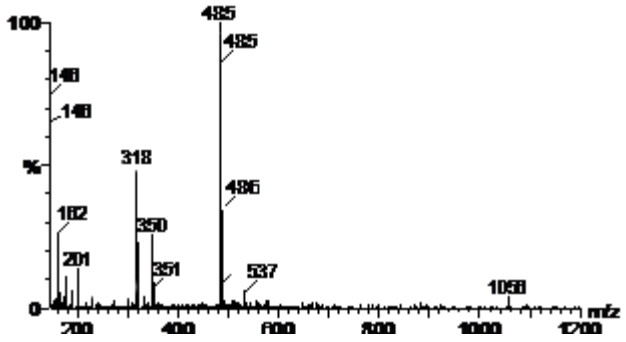


Figure 2-12. LC MS Ion Trap hydrazinolysis peak spectra at 46.4 min in CHY-TPCK digest corresponding to Histidine – Adduct.

2.3.3. UV Spectroscopic Analysis

For optimal activity CHY was prepared in a Ph 7.8 Tris buffer system. Reaction kinetics were evaluated using UV-Spec of various concentration TPCK – CHY solutions and allowed varied exposure times prior to analysis for determination of optimal inhibition. Enzyme activity was reduced substantially after 30 min allowing an effective and fast means of determining loss of proteolytic activity in CHY (Appendix B, Fig. B-4).

2.3.4. Tryptic Digest

As further development of mass spectrometric methods of analysis and sample preparation for study was employed an additional means of enzyme digestion was explored. With a wealth of bioinformatics data available a tryptic digest would prove very useful in protein – adduct experiments and was primarily pursued here as a model method for potential future studies.

Trypsin, a serine protease, is a digestive enzyme that specifically cleaves polypeptides on the carboxylic side of lysine and arginine residues. Native trypsin is subject to autolysis, a self cleaving process that can produce pseudo-enzymes that exhibit a wider range of cleavage sites, including chymotrypsin like activity. Trypsin Gold is manufactured porcine trypsin in which the lysine residues have been modified by reductive methylation, resulting in a highly stable enzyme. Additionally the prepared trypsin is treated with TPCK to inhibit any chymotrypsin activity and then purified further by affinity chromatography and lyophized.

CHY was digested both in solution and via SDS-PAGE Gel isolation with Promega mass spectrometry grade trypsin gold. In this experiment the tryptic digest of CHY did not yield conclusive peptide-adduct data, most likely as a result of CHY and trypsin's inherently bound functional roles.

2.4. Conclusions

Developing an efficient and effective means of evaluating PAH-protein adducts is an important step in understanding the full effect these pollutants may have on cellular function and biochemical pathways. The ubiquitous nature of these pollutants alone should interest all sectors of society in the determination of the immediate effect of toxin exposure on proteins in terms of both structural modification and proteolytic function.

This study positively identified an adduct event on the whole protein by loss of enzymatic activity and by expected mass shift, also by chemical degradation of the protein to its constituent amino acids and detection of an adduct-hydrazide. The reaction with hydrazine allows the determination of adduction to a specific residue type, but offers no additional information as to where on the proteins structure the adduct may be.

Utilizing an enzymatic digestion of a protein could offer additional insight into the adduct formation, and possibly identify the specific sight of adduction. Upon tryptic digestion of the protein the resulting peptides are well matched to ESI ion trap analysis by nature of their N and C termini. CID analysis of the peptides within the trap would allow for further sequence information of the peptides secondary structure. Taking this a step further one can incorporate knowledge of a proposed adduct reaction type and write a modification specific search protocol for the available bioinformatics database. The

data then obtained from the MS can be searched against the expected tryptic peptides for the protein while allowing for the potential mass shift of the adduct. In whole this would offer a much greater confirmation than the chemical degradation alone.

Future steps in this project call for an exposure of CHY to BPDE. CHY function can be reliable studied by enzymatic assay and whole protein and amino acid modification determined by LC ESI ion trap MS. Additional research direction would involve the incorporation of tryptic digestion of proteins of interest upon exposure to PAHs for more accurate determination of adduction sites.

2.5. References

1. *Environmental Health Criteria 202, IPCS*, World Health Organization: Geneva, 1998; 47- 96.
2. Cerniglia C.E. (1992) Biodegradation of polycyclic aromatic hydrocarbons. *Biodegradation* 3:351- 368.
3. Pavanello, S., Clonfero, E. (2000) Biological indicators of genotoxic risk and metabolic polymorphisms. *Mutation Research 463*:285 – 308.
4. Angerer, J., Mannschreck, C., Gundel, J. (1997) Biological monitoring and biochemical effect monitoring of exposure to polycyclic aromatic hydrocarbons. *Int Arch Occup Environ Health* 70:365-377.
5. Hemminki K., et al. (1990) DNA adducts in humans related to occupational and environmental exposure to aromatic compounds. In: Vainio H, Sorsa M, McMichael AJ (eds) Complex mixtures and cancer risk. IARC, Lyon pp181-192.
6. Ovrebo S, et al. Biological monitoring of exposure to polycyclic aromatic hydrocarbon in an electrode paste plant. *J Occup Med* 36:303 – 310.
7. Phillips D H, et al. (1998) *Mutat Res* 204:5331-541.
8. Reddy M V, Hemminiki K, Randerath K (1991) Post labeling analysis of polycyclic aromatic hydrocarbon-DNA adducts in white blood cells of foundry workers. *J Toxicol Environ Health* 34:177-185.

9. Petera F P, et al. (1994) Carcinogen-DNA adducts and gene mutation in foundry workers with low-level exposure to polycyclic aromatic hydrocarbons. *Carcinogenesis* 15:2905-2910.

10. Hemminiki K, et al (1990) Aromatic DNA adducts in white blood cells of coke workers. *Int Arch Occup Environ Health* 62:467 – 470.

11. Ovrebo S, et al (1995) Studies of biomarkers in aluminum workers occupationally exposed to polycyclic aromatic hydrocarbons. *Cancer Detect Prev* 19:258-267.

12. Herbert R, et al. (1990) detection of adducts of DNA in white blood cells of roofers by ^{32}P post labeling. *Scand J Work Environ Health* 16:135-143.

13. Santella R M, et al. (1995) Polycyclic aromatic hydrocarbon-DNA and protein adducts in coal tar treated patients and controls and their relationship to glutathione S-transferase genotype. *Mutat Res* 334: 117-124.

14. Kure E H, Andreassen A, Ovrebo S, et al. (1997) BaP-hSA adducts in humans exposed to PAHs in an industrialized area of Poland. *Nor – Occup Environ Med* 54/9:662- 666.

15. Willett K L, et al. (1998) Inhibition of CYP1A1- Dependant Activity by the Polynuclear Aromatic Hydrocarbon Flouranthene. *Biochem Pharm* 55: 831-839.

16. Brunmark P, et al (1998) *Acta Chem Scand* 52:797.

17. Ferranti P, et al. (1996) *Carcinogenesis* 17:2661.

18. Helleberg H, and Tornqvist M. A new approach for measuring protein adducts from BaP diolepoxide by HPLC/MS, *Rapid Communication in Mass Spectrometry*. 14:1644-1653

19. Mraz J, et al *6th European Meeting on Mass Spectrometry in Occupational and Environmental Health, September 1-3 1999, Stockholm, Scientific Program and Book of Abstracts*, Ostman C (ed). Acro Media, Stockholm University: Stockholm, 1999; 70-71

20. Sabbioni G., et al. (2001) *Chem Res Toxicol.* 13:82.

21. Brunmark P., Harriman S., Skipper P.L., Wishnok J.S., Amin S., Tannenbaum S.R. (1997) *Chem. Res. Toxicol.* 10:880.

22. Day B.W., Skipper, P.L., Zaia J., and Tannenbaum S.R. (1991) Benzo[a]pyrene *anti*-diol epoxides covalently modifies human serum albumin carboxylate side chains and imidiazole side chain of histidine (146). *J. Am. Chem. Soc.* 113:8505–8509.

23. Skipper, Paul L. (1996) Influence of Tertiary Structure on Nucleophilic Substitution Reactions of Proteins. *Chem. Res. Toxicol.* 9: 918–923.

24. Hummel, B. (1959) A Modified Spectrophotometric Determination of Chymotrypsin, Trypsin and Thrombin. *Can. J. Biochem. Physiol.* 37:1393

CHAPTER III

MS DETERMINATION OF POLYCYCLIC AROMATIC HYDROCARBONS (PAHS) IN SEDIMENT AND WATER SAMPLES FROM BOLGODA AND BEIRA LAKES

3.1. Abstract

Eleven polycyclic aromatic hydrocarbons (PAHs) have been quantified in surface sediment and aqueous samples from the Bolgoda and Beira lakes, Sri Lanka. This study represents an initial attempt to study the potential impacts of PAHs on the lakes. Water and sediment samples were collected from four sites in Beira Lake which is situated in downtown Colombo, capital city of Sri Lanka, and from six sites in Bolgoda Lake which is located on an outskirt of Colombo but is surrounded by a number of busy townships dotted with a variety of small industries. The lake also receives massive amounts of industrial effluents from an industrial center located upstream. Total concentrations (sum of 11 PAHs) of PAHs among the four sediment samples in the Beira Lake showed a wide site to site variability ranging from 151 to 568 μg/kg, while the concentration range in the surface water was 46 – 84 ng/L. The PAH concentrations in the sediments in Bolgoda Lake ranged from 115 to 934 μg/kg, and their level in water samples was between 40 and 127 ng/L. In general the PAH composition in the sediments varied from that in the water column. The total PAH concentrations in the sediments appeared to reflect the proximity of the sampling site to anthropogenic sources. The PAHs of Beira Lake sediments and waters were dominated by those with 4-5 rings: 66 – 75% in water and 71 – 92% in sediments. Similar concentration pattern was observed for the Bolgoda Lake: 57 – 61% for water and 71-89% for sediments, a strong indication of pyrolytic inputs. The applications of anthracene/phenanthrene and fluoranthene/pyrene ratios also indicated dominant pyrolytic inputs. In addition, the

presence of high 1-nitropyrene levels at several sites indicates a strong contribution from mobile sources; especially from diesel-powered vehicles. Short-term solid-phase Microtox assays showed toxicity of the sediments at all sites; however there was little correlation between the results of the assay and total PAH concentrations.

3.2. Introduction

The pollution of polycyclic aromatic hydrocarbons (PAHs) has been widely used to assess the potential impact of anthropogenic activities on aquatic environments because their occurrence in water is closely tied to urban activities. Many PAHs possess mutagenic and carcinogenic properties [1,2]. PAH distribution and toxic potentials therefore have been the focus of numerous studies in waterways including the Great Lakes [3], Yanisei Bay [4], and the Fraser River basin [5]. Sri Lanka, a small island nation with a dense population of about 20 million people, faces a multitude of environmental stresses ranging from deforestation to traffic congestion and the deterioration of water quality. This study was undertaken to understand the occurrence, sources, and potential impacts of PAH in the waterways in Sri Lanka. Two lakes, Beira and Bolgoda lakes, were selected for the study due to their economic value and high level of pollution. Beira Lake, situated in downtown Colombo, the capital city of Sri Lanka, is highly polluted. Sources of pollution are multivarious. For instance, clusters of communities have sprung up along the edges of the lake in recent times and many shacks have been built. These communities are generally not connected to municipal sewer systems and substantial quantities of domestic sewage and untreated wastewaters are discharged directly into the lake. Small industries have also grown rapidly around the lakes, most of which are not believed to have adequate facilities to treat industrial wastewater, especially organic wastes. In addition, Sri Lanka has experienced an upsurge of motor vehicles, including millions of three-wheelers and minivans that are powered by leaded-gasoline and diesel fuels. Traffic congestion and severe air pollution due to vehicle emissions now is a common daily occurrences and is considered major potential

sources of PAH in the lakes. Although Bolgoda Lake is situated some distance from Colombo, it is heavily polluted due to the growing number of towns with attendant increase in small businesses and various industries along its shores. These new developments have undoubtedly impacted the lake through the discharge of PAHs and other anthropogenic chemicals present in industrial wastewater and from street runoffs. The lake, additionally, receives a large quantity of pollutants from the industrial zone in the north.

The pollution caused by PAH has led to various studies on the distribution and origin of PAHs in environment [6,7]. Based on the proportions of different PAH, most studies aim to distinguish PAHs of petrogenic sources from those of pyrolytic origins. The PAHs of petrogenic origin, prevalent in coals and fossil fuels, are formed from diagenesis of sedimentary organic material under low to moderate temperature and tend to consist of low molecular weight PAHs with two to three aromatic rings [8]. The pyrolytic PAHs, on the other hand, are formed at much higher temperatures, for example, greater than 500^{o}C, and consist mainly of four and higher number of rings [9]. Thus, an increase in the proportion of higher molecular weight PAHs is interpreted to indicate contamination of mainly of pyrolytic origin. The prevalence of high molecular weight PAHs in the urban dusts [10] and in atmospheric particles [11] illustrates the chemistry of their formation at high temperature.

For this study 11 known potential toxinogenic and mutagenic PAHs [12,13] were chosen: naphthalene, acenaphthene, phenanthrene, anthracene, benzo(*k*)fluoranthene, fluoranthene, pyrene, chrysene, 1-nitropyrene, benzo(*a*)pyrene, and dibenz(*a*,*h*)anthracene. Of these PAHs, five have 4-5 aromatic rings

(benzo[*k*]fluoranthene, pyrene, chrysene, benzo[*a*]pyrene, and dibenz[*a,h*]anthracene) and the remaining four have 2-3 rings (naphthalene, acenaphthene, phenanthrene, and anthracene).

3.3. Methodology

3.3.1. Sample Site Selection & Collection

Fig. 3-1 shows a map of Beira and Bolgoda Lakes, indicating the various sampling sites. In comparison to most lakes around the world, both lakes are characterized by extreme shallowness with an average depth of 2–3 m, and being in a tropical region near the equator, they experience little seasonal variation of water temperature (26-28^{o}C) and receive frequent rainstorms, ranging from local thunderstorms that can occur almost daily to the southwest monsoons from May to September. In the catchment area where the lakes are located, rainfall ranges from about 4 inches in January to nearly 14 inches in October.

Beira Lake is situated well within the city limit of Colombo and opens to the Indian Ocean through a short and narrow canal in the west, and extends eastward by branching into West and East Lakes. The West Lake narrows to a strip of water and meanders into another small lake, S.W. Lake. The exchange of water between these lakes is very slow. Fresh water flows into the eastern and northern parts of the lake through a number of small tributaries. Three sampling sites BA-1 and BA-3, and B-4 in Beira Lake, are in the vicinity where reclamation work was being performed. These areas are surrounded by a high density of shops and busy streets, and a large number of

people have built shacks along the edges of the lake. Site BA-2 is in a region where the shops are not as numerous and the traffic is less congested.

Bolgoda Lake is also located in the western Sri Lanka and opens to the Indian Ocean via the two narrow, winding canals. Fresh water flows into the lake mainly by the Weras Ganga from the Dehiwala area and the eastern drain of Lady Catherine industrial region. In addition, the lake is surrounded by a substantial number of small shops and industries, some of which are considered as potential sources of PAHs. Sampling sites WG-1 and WG-2 are located in the mouth of North Lake and BNL-1 is in the middle of North Lake. These sites are therefore susceptible to being impacted by the water pollutants and contaminants originating from the industrial zone in the north. BNL-2 is close to the canal where the exchange of water between the ocean and lake occurs. Site BG is away from the river Ganga, hence the water quality is expected to be more reflective of the immediate surroundings. Site BSL is along the canal area close to Bolgoda south and is distant from the other study sites.

Surface sediment samples (0-3 cm) were collected using a scoop in the shallow point of selected sites during 2003-2004. Water samples were also collected from the surface in 1 L amber glass bottles with screw caps. Samples, collected in glass jars were stored frozen at -20^{o}C until analysis. Both colloid and surface sediment samples have been collected in duplicate from selected sites at various times of the year.

3.3.2. Sample Preparation

Moist sediments were subjected to Soxhlet extraction. The methodology for the extraction and fractionation is similar to that described [14]. Briefly, 10 g sediments were extracted with 50 mL of dichloromethane at three cycles per hour for 24 hours. After cooling to ambient temperature, residual water was removed by adding 2-3 g of sodium sulfate. Then, the contents were evaporated at 35^{o}C to near dryness (1-2 mL) and to dryness under a gentle stream of nitrogen. The PAH fraction was enriched using sephadex and silica gel chromatography in the following manner: the fraction is mixed with a slurry of activated silica gel in dichloromethane and transferred to a small glass column (10.0 cm x 1.0 cm id). The gel was equilibrated with 30 mL of pentane, 2 mL of the extract was loaded, 25 mL of pentane was passed through, and the PAH fraction was eluted with 25 mL of dichloromethane:pentane (4:6, v/v) and was taken to dryness under a gentle stream of nitrogen. For analysis by high performance liquid chromatography (HPLC), the extract was exchanged into acetonitrile. PAH recoveries were determined by adding known quantities of PAH mixtures to the sediments (5-10 grams) and they were subjected to the same extraction procedures.

The water samples containing both dissolved and colloid-adsorbed PAHs were extracted using a separatory funnel. Into a 1 L separatory funnel, 500 mL of water sample and 50 mL of CH_2Cl_2 were added and extracted for three times. PAH recoveries were checked by adding known quantities of PAH mixtures to distilled water. These spiked waters were then extracted using the same procedures described in the previous paragraph.

3.3.3. HPLC Sample Analysis

PAHs were analyzed by high performance liquid chromatography (HPLC) and gas chromatography coupled to mass spectrometry (GC-MS). The HPLC procedure was based on the U.S. Environmental Protection Agency (U.S. EPA) Method 8310 using high performance liquid chromatograph (Agilent 1100) with a gradient pumping system and a 250 mm by 2.6 reverse phase HC-ODS Sil-X 5-micron particle size column. A UV detector with excitation wavelength 254 nm was coupled to the fluorescence detector. The fluorescence detector uses an excitation wavelength of 280 nm and emission greater than 289 nm cutoff. Quantitation was performed by the external standard method with individual PAH standards.

3.3.4. GCMS Sample Analysis

The GC/MS methodology was based on U.S. EPA 8100 Method. Samples were quantified by gas chromatography/mass spectrometry (GC/MS) on a Thermo Electron GCQ plus ion trap mass spectrometer. Spiked samples containing a mixture of 11 PAHs and two reference compounds were analyzed for every 10 samples. Perfluoro-*n*-butylamine was used as the calibration compound to ensure instrument performance. Samples were separated and analyzed on an Altech DB-5MS capillary column (0.32 mm I.D. x 0.1 mm film thickness x 30 m). A Thermo Electron HS 2000 autosampler was used for all sample injections. Helium was used as carrier gas at a constant flow rate of 1.4 mL/min. The gas chromatograph oven temperature was programmed as follows: the initial temperature was maintained at 100^{o}C and held for 1.0 min following injection. The temperature was then ramped at 20^{o} C/min to 200^{o} C and held there for 10 min. Finally

the temperature was ramped at a rate of 10^{o} C/min to 300^{o} C. The injection volume was 5 μL and the injector was maintained at 250^{o} C. A splitless injection was used with septum purge operative throughout the analysis. The mass spectrometer was operated in Electron Impact mode for all analysis with an electron energy of 70 eV, scanning at a mass range of 50 – 400 amu every 0.8 s. A five point calibration routine was used for all analysis of each PAH with at least one calibration point of higher or lower concentration than the concentration of the unknown compound under consideration. Peak areas were used to determine the response factors of each of the parent PAHs GC/MS retention times. The identity of PAHs was accomplished by comparison of spectral pattern of ions obtained to the libraries of mass spectral data. Perdeuterated aromatic compounds were used as internal standards for quantitative analyses in each aromatic series.

3.3.5. Evaluation of Sediment Toxicity

Vibrio fischeri NRRL B-11177 was obtained from Azur Environmental, Carlsbad, CA, USA). The organisms were received freeze-dried and stored below – 20^{o} C. Cells of *V. fischeri* were diluted to an absorbance of 0.4 at 650 nm in a simple medium that contained KH_2PO_4 (18.4 mM), NaCl (0.5 M), and $MgSO_4$ (4.1 mM), The medium was adjusted to pH 7.2 and autoclaved prior to use

The solid phase Microtox assay was conducted as recommended by the manufacturers. The procedure involved placing a 1-g dry sediment sample in 5 mL assay medium in 10 mL borosilicate tubes and was used to make a series of sediment dilutions. The sediment samples were serially diluted and 2 mL aliquots of evenly suspended samples were mixed with equal volume of test culture in polystyrene dishes (35 x 10

mm). After 15 min of gentle shaking on a platform linear shaker at room temperature, the contents were transferred to a polystyrene cuvette (12 x 75 mm) and luminescence was measured in a tube luminometer (Zylux FB15). Luminescence in the assay medium served as control. Phenol was used as reference toxin and EC50 values were maintained within 13-26 mg phenol/L as recommended by the manufacturer and the luminescence change of the test organism in the medium without the sediment served as positive control.

Calculation of EC50 values were obtained from a plot of cumulative response (0-100%) versus log sediment amount. Inhibition of luminescence intensity in the short-term test, for instance, was based on the following formula (McConkey et al 1997) = 100 [1-(LfCi/LiCf)], where Li is the initial luminescence prior to and Lf is the luminescence following a 15-min exposure to sediment, and Ci and Cf are the initial and final luminescence of the bacteria that were not exposed to the sediment. The data for % light inhibition = $100/1 + e^{-\beta(\ln x - \ln \mu)}$, where x is the amount of sediment, μ is the EC50, and β is the slope of the amount-response curve.

3.4. Results & Discussion

3.4.1. Beira Lake

As shown in the map (Fig. 3-1 A) sampling in Beira Lake was conducted at four sites over a period of 18 months. Composite sediment and overlying water samples were collected from the sites and were analyzed for eleven parent PAHs. The overall concentration range for the sediments was found to be between 152 and 569 µg/kg with average 329 µg/kg dry sediment (Table 3-I). When the individual sites were compared, the total PAH concentrations from the BA-2 site were substantially below those obtained from the other sites (BA-1, BA-3, and BA-4). These sites are adjacent to an area of high traffic density, with numerous shops and small industries of various kinds in close proximity, thus direct local inputs of PAHs from the street runoff into the lake is probably most significant. Parts of the region around these sampling sites had recently undergone sediment dredging (hatched area) and are speculated to be heavily contaminated by metals and other contaminants. The region around the BA-2 site by comparison is less populated and the traffic not as congested.

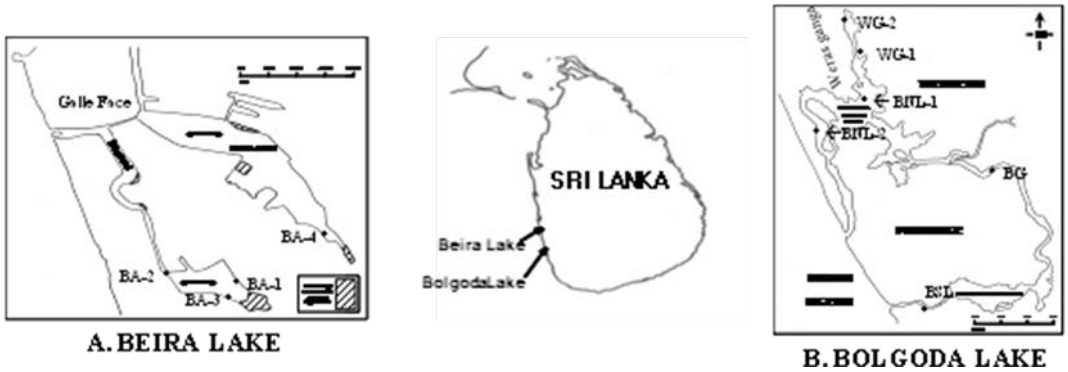


Figure 3-1. Sampling sites in Beira Lake (A) and Bolgoda Lake (B).

Table 3-I. Average PAH levels in Beira and Bolgoda Lakes [a,b].

A. Beira Lake	Sampling Sites		
BA-1	BA-2	BA-3	BA-4
Water (ng/L)			
78 ± 5.0	46 ± 3.0	112 ± 4.7	85 ± 5.6
Sediment (µg/Kg)			
569 ± 33.0	152 ± 8.1	465 ± 25.0	371 ± 21.0

B. Bolgoda Lake	Sampling Sites				
WG-1	WG-2	BNL-1	BNL-2	BSL	BG
Water (ng/L)					
40 ± 2.2	127 ± 6.3	70 ± 3.7	67 ± 2.5	71 ± 2.6	71 ± 3.1
Sediment (µg/Kg)					
406 ± 24.1	865 ± 50.1	115 ± 6.2	934 ±24.0	121 ± 6.8	249 ± 13.0

[a] Average of 11 PAH species: naphthalene, acenaphthene, phenanthrene, Anthracene, benzo[k]fluoranthene, fluoranthene, pyrene, chrysene, 1-Nitropyrene, benzo[o]pyrene, dibenz [a,h] anthracene

[b] The mean ±SEM (N = 3 - 5)

The results from the analysis of overlying water samples in Beira Lake (Table 3-I A) show that PAH concentrations in the water column at the BA-2 site were also lower (46 ± 3.0 ng/L) than those of other three sites, which ranged 78 – 85 ng/L. Thus, there was some correspondence between the aqueous and sediment samples in the PAH concentrations at this site. For the most part, however, there were low correlations in the concentration of PAHs between the phases. A comparison of individual PAHs also

indicated that PAH concentrations in the sediments did not reflect those of water column (Fig. 3-2), probably due to a rapid mixing and transport of the PAH-adsorbed colloidal particles in shallow water column and that the majority of PAHs in water column are expected to be associated with colloidal particles since PAHs have low aqueous solubility as indicated by their high log K_{ow} values (3.36 for naphthalene to 6.90 for coronene) [15]. The reduced site-to-site variability of the PAH concentrations in the water column could also be explained by the rapid equilibration of the colloidal particles. Various external factors such as wind velocity over the evaporating surface, frequency and intensity of rain storming events around the lakes, the effects of intense sunlight on the PAH degradation, and the relationship between year-round warm water temperature probably influence the loading of PAH in the lake.

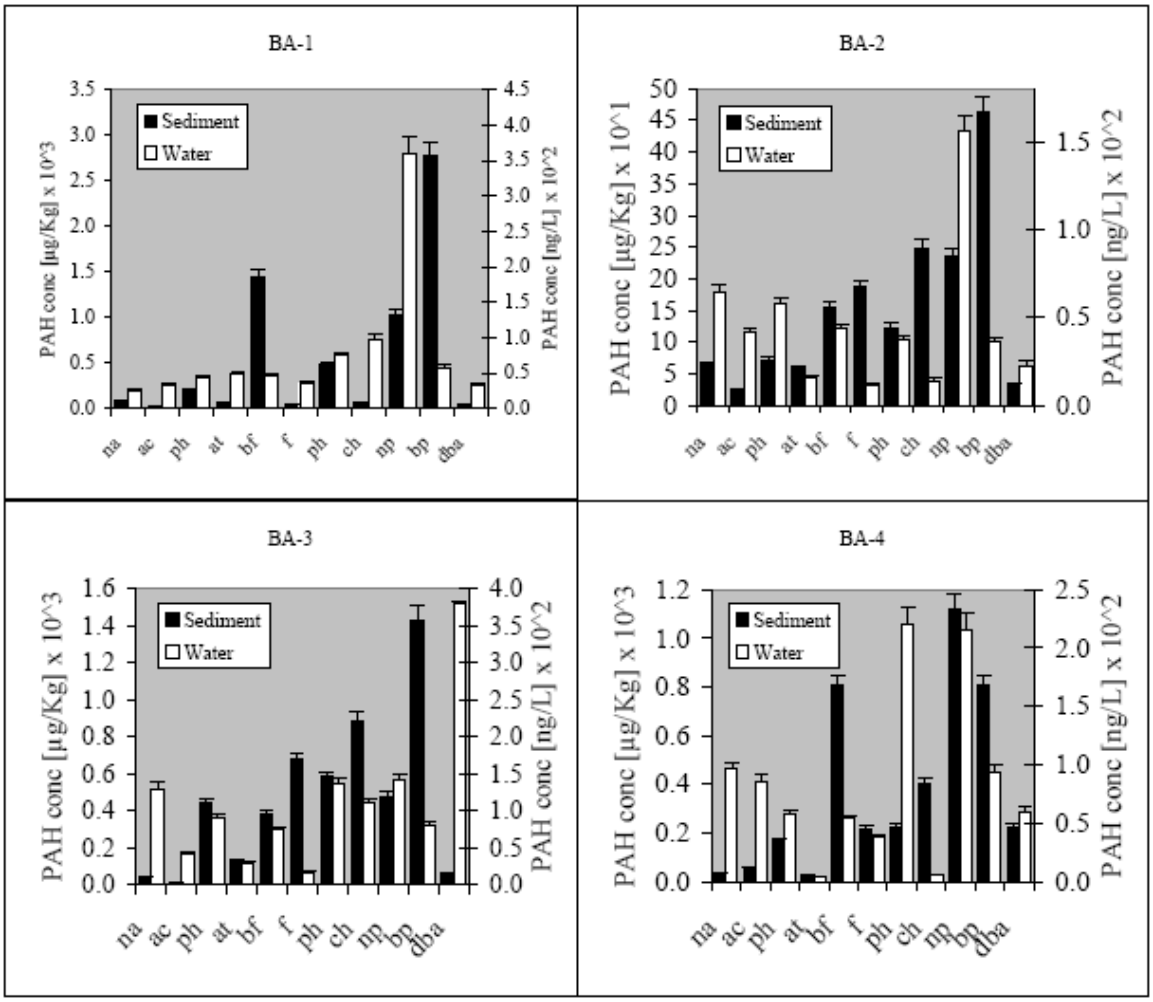


Figure 3-2. Relative distributions of PAHs for four sites in the water column and sediments in Beira Lake.

Values are mean ± standard error of mean. np=naphthalene, ac=acenaphthene, ph=phenanthracene, at=anthracene, bf=benzo(k)fluoranthene, f=fluoranthene, p=pyrene, ch=chrysene, np=1-Nitropyrene, bp=benzo(a)pyrene, dba=dibenzo(a,h)anthracene.

3.4.2. Bolgoda Lake

Bolgoda Lake (Fig. 3-1 B), like Beira Lake, is also located in the catchment area but some distance away from the city of Colombo. However, it is expected to experience a substantial PAH loading since the lake is surrounded by a fair number of small industries, such as incinerators and electricity power plants that are situated close to the lake, many of which produce gaseous emissions and generate wastewaters that are expected to contain high concentrations of PAHs. The PAH levels for the sediments ranged from 121 ± μg/kg at the BNL-1 to 934 ± μg/kg sediment at the BNL-2 site with the mean of 448 μg/kg sediment (Fig. 3-3). Overall PAH levels in sediment samples in the Bolgoda Lake thus were somewhat higher than those of Beira Lake. The concentrations at WG-2 site were also high at 865 μg/kg. The potential sources of PAHs around the site are 16 small industries including a number of gas stations and a chemical industry, most of which have the potential of discharging significant quantities of PAHs through wastewaters as well as gaseous emissions. In addition, Bolgoda Lake receives massive amounts of industrial effluents via the Ganga river, a major tributary, flowing out of the Lady Catherine industrial zone in the north (Dehiwala area), and that the region around site WG-2 was probably impacted by the effluents from the industrial zone more than the other sites because of its location and other geologic characteristics. On the other hand, heavy contamination of the sediments in BNL-2 site in the west was probably attributed to the Panandura estuary where the exchange of lake water with the polluted sea coast occurs and possibly reflects localized geographic variation of some unknown origin. The extent of contribution of the PAHs present in the industrial effluents from the industrial zone in Dehiwala area is unlikely to be important at this site because of

the low levels of PAHs observed around the site BNL-1, which is closer to the WG-2 site. Located in the general vicinity of BNL-1 site are number of small industries such as a garment factory, several gas stations, and a large farm. Also, the general area is not as congested and the population is somewhat less than the areas in the vicinity of other sites. These conditions perhaps explain the low levels of PAHs observed around the site. The PAH levels for the remaining sites (WG-1, BSL, and BG) were also low, ranging between 115 – 247 µg/Kg. BG and BSL sites have one to two gasoline stations each. None of the small industries in the area (for instance, hotels and shops) support activities expected to be directly associated with generation of PAHs. The total PAH concentrations for the six aqueous samples ranged from 40 to 127 ng/L with the average of 79 ng/L (Table 3-I B). The PAH distribution in the water column thus was similar to that of Beira Lake in that the PAH concentrations in the water column showed less variability for similar reasons.

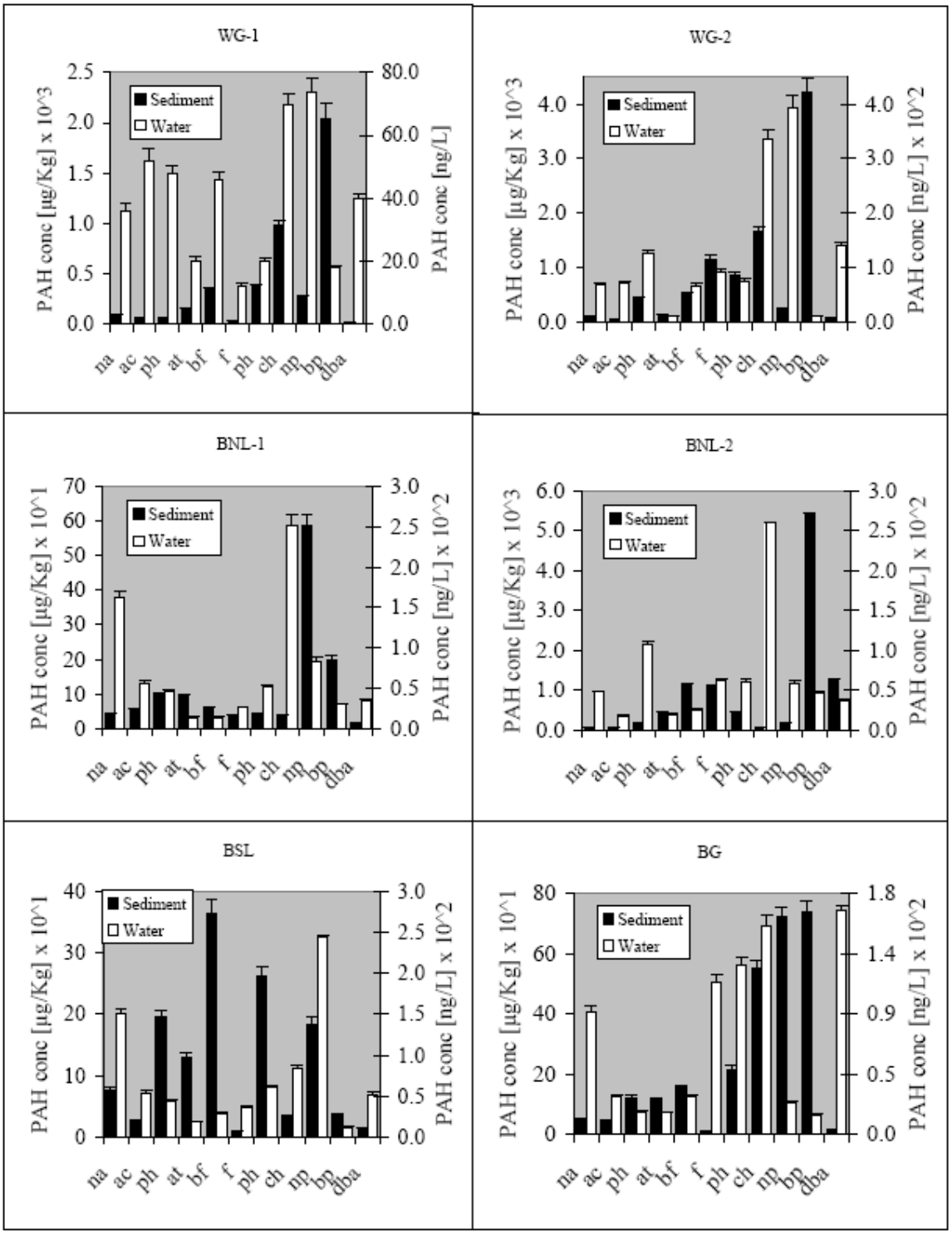


Figure 3-3. PAH Relative distributions of PAHs for six sites in the water column and sediments in Bolgoda Lake.

Values are mean ± standard error of mean. np=naphthalene, ac=acenaphthene, ph=phenanthrene, at=anthracene, bf=benzo(k)fluoranthene, f=fluoranthene, p= pyrene, ch=chrysene, np=1-Nitropyrene, bp=benzo(a)pyrene, dba=dibenzo(a,h)anthracene.

3.4.3. PAH Ratios

An objective of this study was to determine potential sources of PAHs at each site. Total concentrations of four- to five-ring PAHs were compared to those of two- to three ring PAHs. Of the eleven major parent PAHs analyzed, five had 4-5 aromatic rings (benzo[*k*]fluoranthene, pyrene, chrysene, benzo[*a*]pyrene, and dibenz[*a,h*]anthracene) and the remaining four PAHs had 2-3 rings (naphthalene, acenaphthene, phenanthrene, and anthracene). In Beira Lake the inputs of the PAH with 4-5 rings ranged from 58 to 66% of total concentrations for the water samples and 71 to 92% for the sediments (Table 3-II). A similar profile existed for Bolgoda Lake, where the PAHs with 4-5 rings comprised 57 – 62% for the water samples and 62-86% in the sediments. Thus, the distribution patterns of PAH in both lakes are indicative of dominant contributions from the pyrolytic sources at all sites, although significant site-to-site variation existed. Alternatively, concentration ratios of phenanthrene/anthracene (P/A) and fluoranthene/pyrene (F/P) have been applied to distinguish between PAHs of pyrogenic and petrogenic sources [16-18]. This approach is based on the rationale that for parent PAHs, combustion processes under the conditions of limited oxygen availability tend to produce proportionately more PAH isomers with greater thermodynamic stability [19]. For example, for the three ring PAHs, the proportion of more stable phenanthrene is expected to be produced in higher proportion than its isomer, anthracene, during the combustion processes. One report presented showed that PAHs of pyrolytic origin generally have P/A ratios of 1-10 [7]. The study further showed that fluoranthene/pyrene (F/P) ratios fall between 1-2 for the pyrolytic PAHs. The calculated values in Beira and Bolgoda lakes generally indicated dominance of pyrogenic sources

Table 3-II. Ratios[a] of two-three ring[b] to four-five ring[c] PAHs for Water and Sediment samples from Beira and Bolgoda Lakes[d]

A. Beira Lake	Sampling Sites			
	BA-1	BA-2	BA-3	BA-4
Water (ng/L)				
2-3 ring PAH	38 ± 2.0	38 ± 2.0	62 ± 4.0`	56 ± 3.4
4-5 ring PAH	112 ± 7.3	52 ± 3.5	154 ±5.2	108 ± 7.3
Ratio	75%	57%	71%	66%
Sediment (µg/Kg)				
2-3 ring PAH	84 ± 4.3	82 ± 4.8	262 ± 12.5	100 ± 5.5
4-5 ring PAH	973 ± 56.9	210 ± 10.9	635 ± 35.5	597 ± 33.8
Ratio	92%	72%	71%	86%

B. Bolgoda Lake	Sampling Sites				
WG-1	WG-2	BNL-1	BNL-2	BSL	BG
Water (ng/L)					
2-3 ring PAH					
34 ± 2.3	74 ± 3.5	61 ± 3.2	51 ± 2.3	60 ± 2.9	53 ± 2.6
4-5 ring PAH					
45 ± 2.1	171 ± 8.6	78 ± 4.2	81 ± 2.7	80 ± 2.4	86 ± 3.5
Ratio					
57%	70%	56%	61%	57%	62%
Sediment (µg/Kg)					
2-3 ring PAH					
83 ± 4.4	377 ± 22.8	65 ± 3.4	360 ± 18.5	88 ± 4.7	67 ± 3.8
4-5 ring PAH					
674 ± 40.6	1273 ± 72.9	156 ± 8.5	1413 ± 28.5	149 ± 8.5	400 ± 20.6
Ratio					
89%	77%	71%	80%	63%	86%

[a] Average 4-5 ring PAH / Average 2-3 ring PAH x 100
[b] Average of 5 PAH species: naphthalene, acenaphthene, phenanthrene, anthracene, fluoranthene
[c] Average of 6 PAH species: benzo[k]fluoranthene, pyrene, chrysene, 1-Nitropyrene, benzo[o]pyrene, dibenz [a,h] Anthracene
[d] The mean ±SEM (N = 3 – 5)

Table 3-III. Phenanthrene/anthracene (P/A) and fluoranthene/pyrene (F/P) ratios for water and sediment samples from Beira and Bolgoda Lakes.

A. Beira Lake	Sampling sites			
	BA-1	BA-2	BA-3	BA-4
Water				
P/A	0.9	3.6	3.0	14.5
F/P	0.5	0.3	0.1	0.2
Sediment				
P/A	3.1	1.2	3.3	5.9
F/P	0.1	1.5	1.2	1.0

B. Bolgoda Lake	Sampling sites				
WG-1	WG-2	BNL-1	BNL-2	BSL	BG
Water					
P/A 2.4	10.5	3.3	5.4	2.4	1.0
F/P 0.6	1.2	0.5	1.0	0.6	0.9
Sediment					
P/A 0.4	3.5	1.0	0.4	0.4	1.5
F/P 0.1	1.3	1.0	2.4	0.0	0.0

although there were exceptions (Table 3-III). For instance, the water samples collected from the BA-4 site had the P/A ratio of 14.5 and F/P ratio of 0.2, which suggested a strong input of the petrogenic PAHs. The calculated values from this application are therefore in general agreement with the method based on the ratios of 4-5 ring/2-3 ring PAH ratios.

In addition, concentration of 1-nitropyrene was examined since the PAHs are known to react with certain functional groups such as nitro groups ($-NO_2$) under high temperature conditions, forming nitrated PAHs [20]. For instance, emissions from

combustion of fossil fuels, diesel fuels in particular, produce 1-nitropyrene and other related products. 1-Nitropyrene is potentially mutagenic and carcinogenic in humans [13]. As shown in Table 3-I, its levels were twice the average PAHs both in the sediment and water samples from Beira Lake, particularly in samples from the BA-1, BA-3, and BA-4 sites where the traffic flows are generally heavy and the percentage of the vehicles powered by diesel fuels is high. Thus the results appear to reflect high local inputs of 1-nitropyrene generated by motor vehicles.

3.4.4. Microtox Assay

Various tests have been developed for evaluating the toxicity of PAHs in soils and sediments, including the Microtox assays. The results are shown in (Fig. 3-4). The EC50 values for the sediments indicated that all sediments were toxic to the test organisms but that the toxicity data did not correlate well to total PAH concentrations in the sediments. This discrepancy suggested that the toxicity was affected by the factors other than the PAH concentration or specific PAHs present, such as particle size and the sulfur content in the sediments. From this brief study in the evaluation of toxicity by the Microtox assays, the following conclusions may be reached: the results from the short-term solid-phase assays may not be accurate in evaluating sediment toxicity and that other approaches, long-term assays should be also tried.

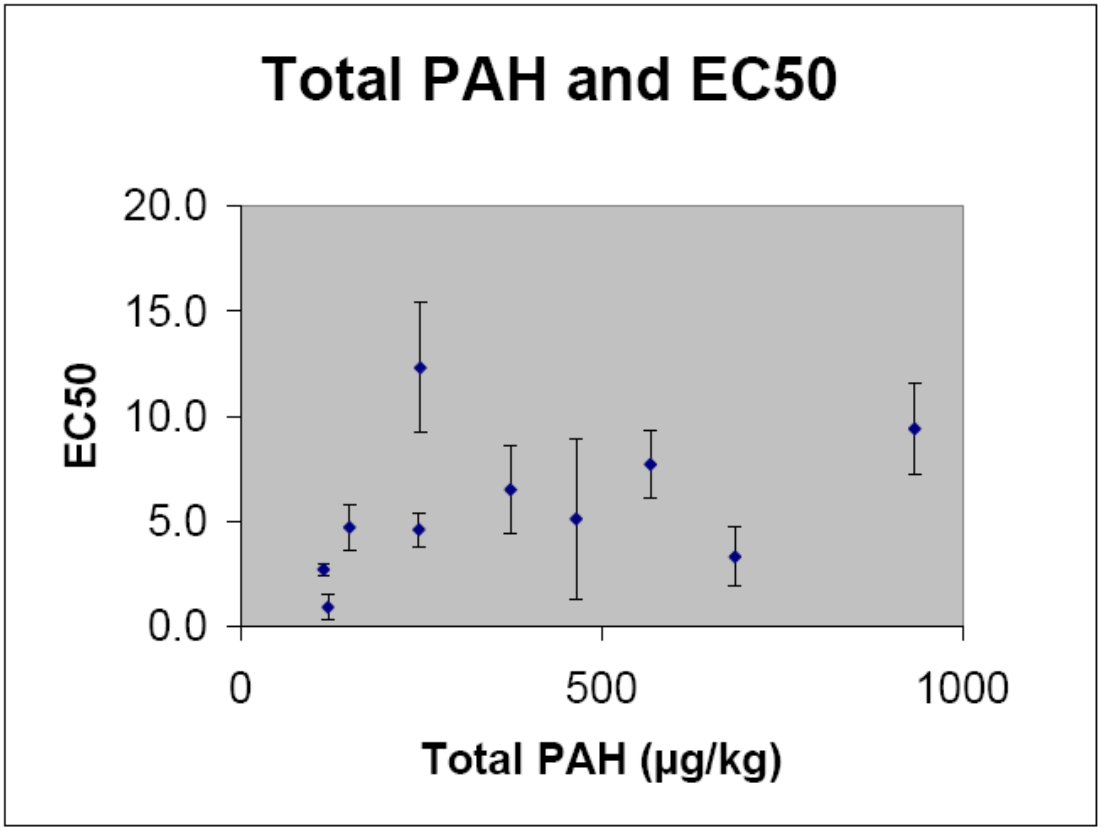


Figure 3-4. Correlation between EC50 Values obtained by Microtox assays and total PAHs for the sediments.

Acknowledgements- Thank the National Science Foundation of Sri Lanka (RG/2001/C/05) for financial support and Dr David Hehemann and the NASA Glen Research Center for assistance and expertise with the GC-MS.

3.5. References

1. Menzie CA, Potocki BB, Santodonato J (1992) Exposure to carcinogenic PAHs in the environment. Environ Sci Technol 26: 1278-1284.
2. Zedeck MS (1980) Polycyclic aromatic hydrocarbons: a review. J of Envir Pathol & Toxicol 3: 537-567.
3. USEPA Report: Assessment and Remediation of Contaminated Sediments (ARCS) Program, Final Summary Report, EPA-905-594-001(1994), 48 pp.
4. Dahle S, Savinov VM, Matishov GG, Evenset A, Naes K (2003) Polycyclic aromatic hydrocarbons (PAHs) in bottom sediments of the Kara Sea shelf, Gulf of Ob and Yanisei Bay. The Science of Total Environ 306:57-71.
5. Yunker MB, Macdonald RW, Vingarzan R, Mitchell RH, Goyette D, Sylvestre S (2002) PAHs in the Fraser River Basin: a critical appraisal of PAH ratios as indicators of PAH source and composition. Organic Geochem 33:489-515.
6. Yunker MB, Snowdon LR, Macdonald RW, Smith JN, Fowler MG, Skibo DN, McLaughlin FA, Danyushevskaya AJ, Petrova VI, Ivanov GI (1996) Polycyclic aromatic hydrocarbon composition and potential sources for sediment samples from the Beaufort and Parents Seas. Environ Sci Technol 30: 1310-1320.
7. Budzinski H, Jones I, Bellocq J. Pierard, Garrigues P (1997) Evaluation of sediment contamination by polycyclic aromatic hydrocarbons in the Gironde estuary. Marine Chem 58:85-97.
8. Potter TL, Simmons KE (1998) Composition of petroleum mixtures, vol 2. In: Total Petroleum Hydrocarbon Criteria Working Group Series, Amherst Scientific Publishers, Amherst, MA.

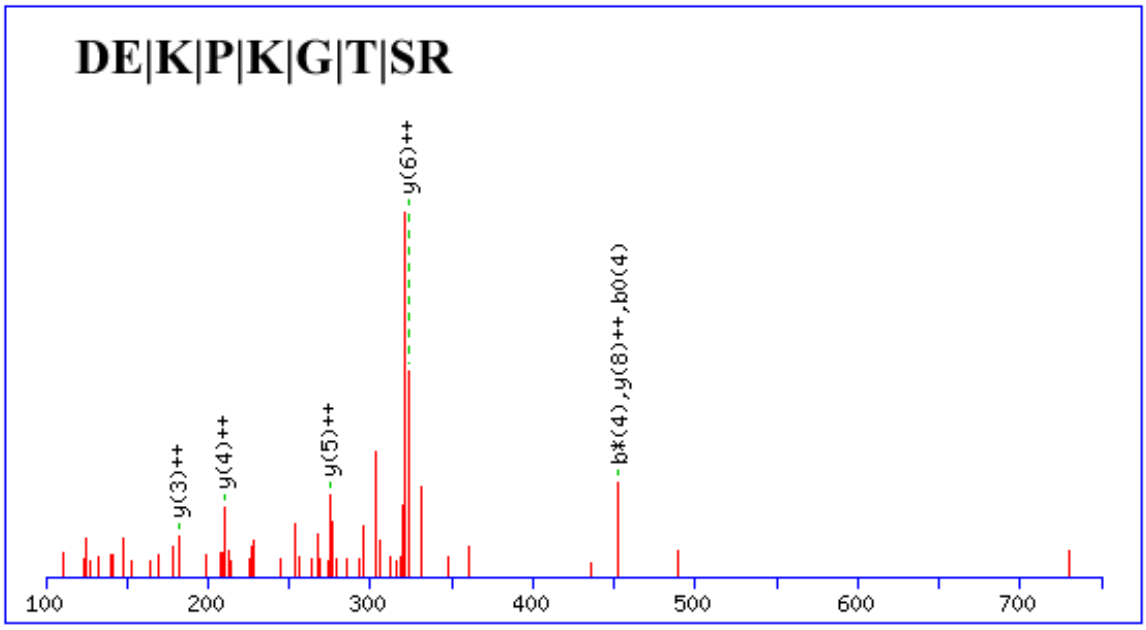
9. Commins BT (1969) Formation of polycyclic aromatic hydrocarbons during pyrolysis and combustion of hydrocarbons. Atmos Environ 3:565-572.
10. Wise SA, Hilpert LR, Rebbert RE, Sander LC, Schanz MM, Chesler SN, May WE (1988) Standard reference materials for the determination of polycyclic aromatic hydrocarbons. Fresenius A Anal Chem 332:576-582.
11. Sicre MA, Marty JC, Saliot A, Aparicio X, Grimalt J, Albaiges J (1987) Aliphatic and aromatic hydrocarbons in different sized aerosols over the Mediterranean Sea: Occurrence and origin. Atmos Environ 21:2247-2259.
12. IARC (International Agency for Research on Cancer (1989) Monographs and Evaluation of the Carcinogenic Risks to Humans: In Diesel and Gasoline Engine Exhausts and Some Nitroarene, vol. 46, IARC, Lyon, France.
13. IARC.1997. 1-nitropyrene. IARC Monograph on Evaluation of Cancer Risk of Chemical to Humans 46:321-358.
14. O'Malley VP, Abrajano TA, Hellou J (1994) Determination of 13C/12C ratios of individual PAHS source apportioned . Organic Geo Chem 21:809-822.
15. Neff JM. 1979. Polycyclic aromatic hydrocarbons in the aquatic environment: sources, fates and biological effects, Applied Science Publishers Lts, Essex, England.
16. Benner BA, Bryner NP, Wise SA, Mulholland GH, Lao RC, Fingas MF (1990) Polycyclic aromatic hydrocarbon emissions from combustion of crude oil on water. Environ Sci Technol 24: 1418-1427.
17. Lake JL, Norwood C, Dimock C, Bowen R (1979) Origins of polycyclic aromatic hydrocarbons in estuarine sediments. Geochim Cosmochim Acta 43:1847-1854.

18. Gschwend PM, Hites RA (1981) Fluxes of polycyclic aromatic hydrocarbons to marine and lacustrine sediments in the northeastern United States. Geochim Cosmochim Acta 45:2359-2367.
19. Albert RA, Reif AK (1988) Standard chemical thermodynamic properties of polycyclic aromatic hydrocarbons and their isomer groups. I. Benzene series. J Phys Chem Ref Data 17: 241-253.
20. Paputa-Peck MC, Mrano RS, Schuetzle D, Riley TC, Hampton CV, Prator TJ, Skewes LM, Jensen TE (1983) Determination of nitrated polynuclear aromatic hydrocarbon chromatography with nitrogen selective detection Anal Chem 55: 1946-1954.

APPENDICES

APPENDIX A

ADDITIONAL TABLES & FIGURES FOR CH I



(A)

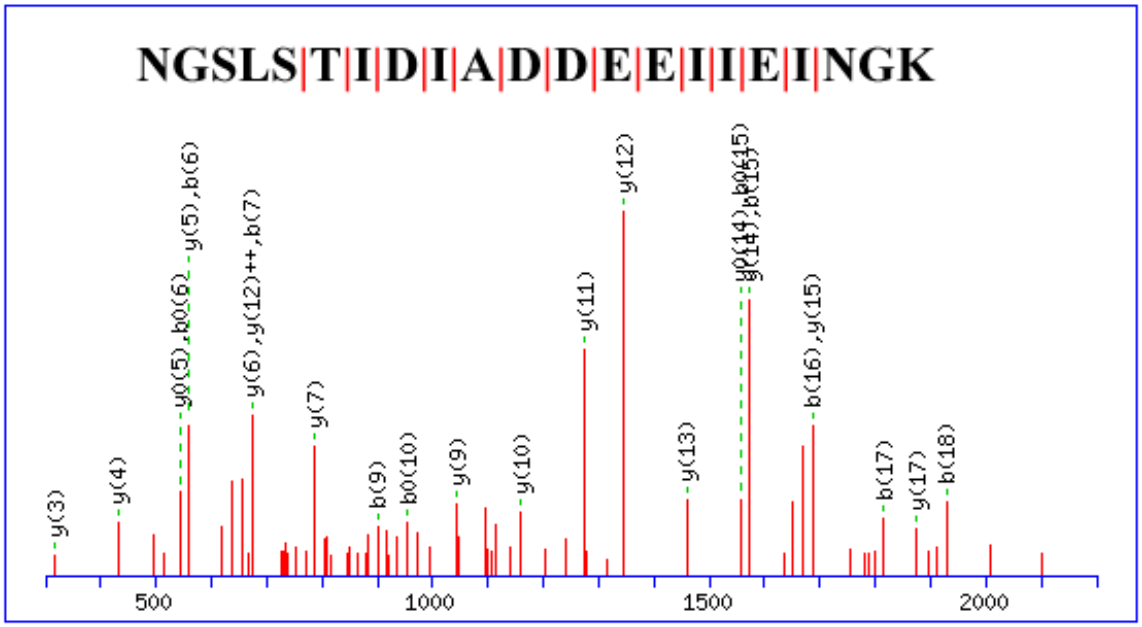
LNDASHKEITDESVRVWLNSLQHLAYDIDDVLDDVATEAMHRELTPESEASTSM

VRKLIPTCCTKFSLSHRLTPKLDSINTQLQQLEKQKSDLGLIVK**DEKPKGTSR**RN

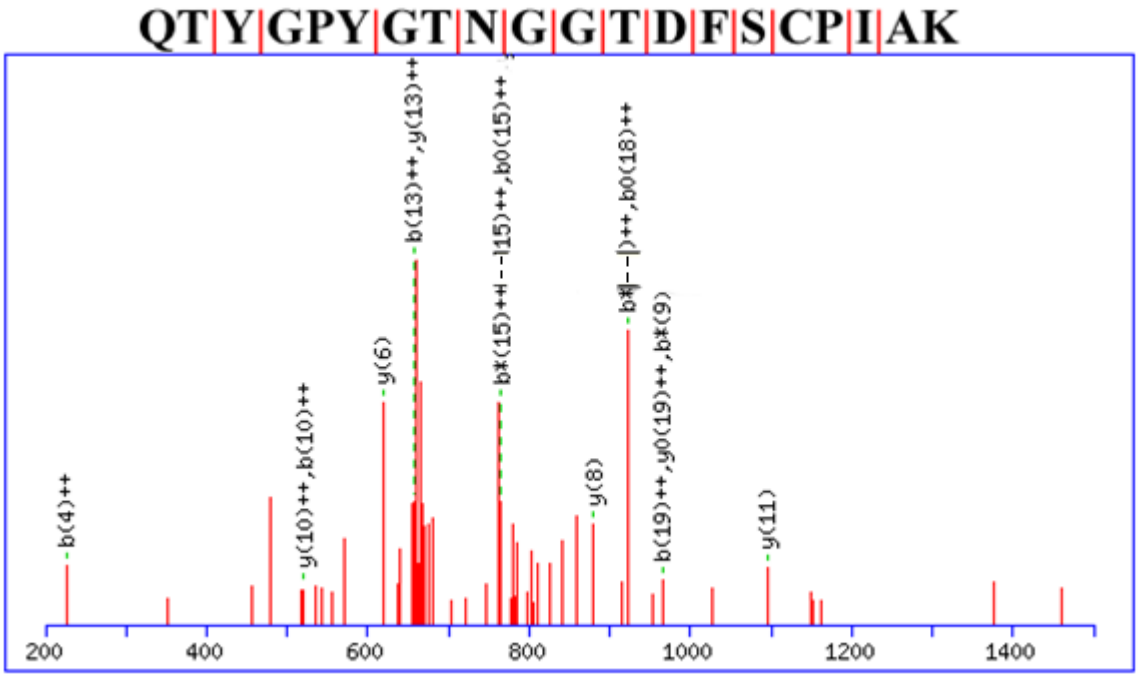
ETSLLESDVVGREGEKKR

(B)

Figure A-1. Tandem mass spectra of peptide (A) matched to a CC-NBS-LRR RGA type protein (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).

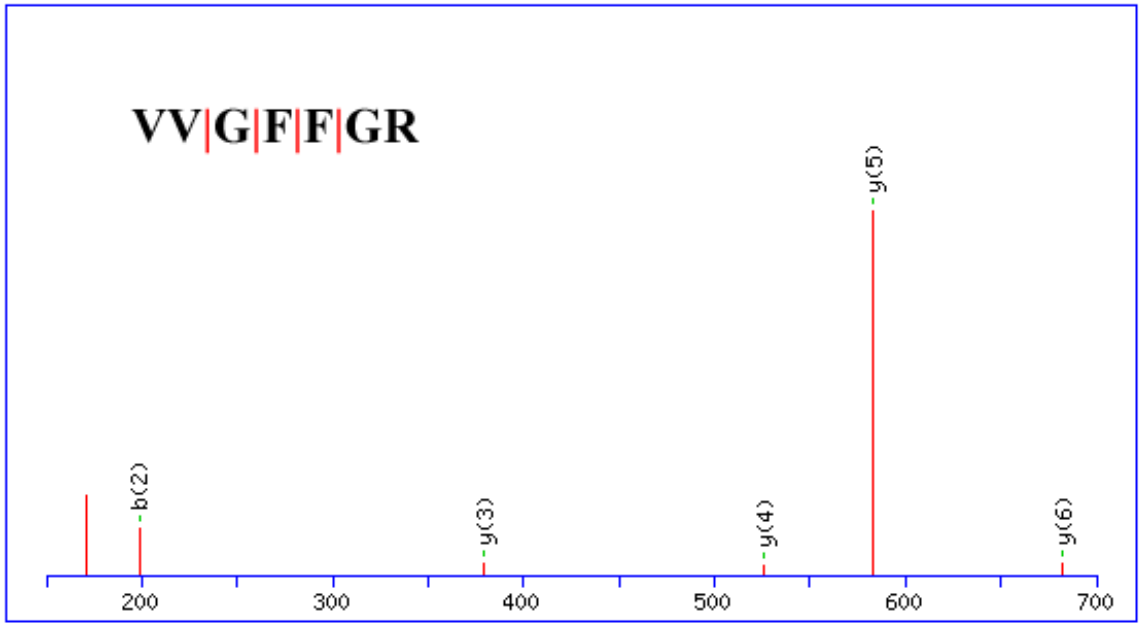


(A)



(B)

Figure A-2. Tandem mass spectra of polypeptides (A,B,C) matched to Lectin (*H. tuberosus*) as indicated (bold) in the primary amino acid sequence (D).



(C)

MANNYVEVGPWGGSGGANPWSIIPNDGRITRINVRSGAIVDAIYFGYTEGGISYE

TAIFGGR**NGSLSTIDIADDEEIIEINGK**VGTLENLNLVSQLTFVTNK**QTYGPYGT**

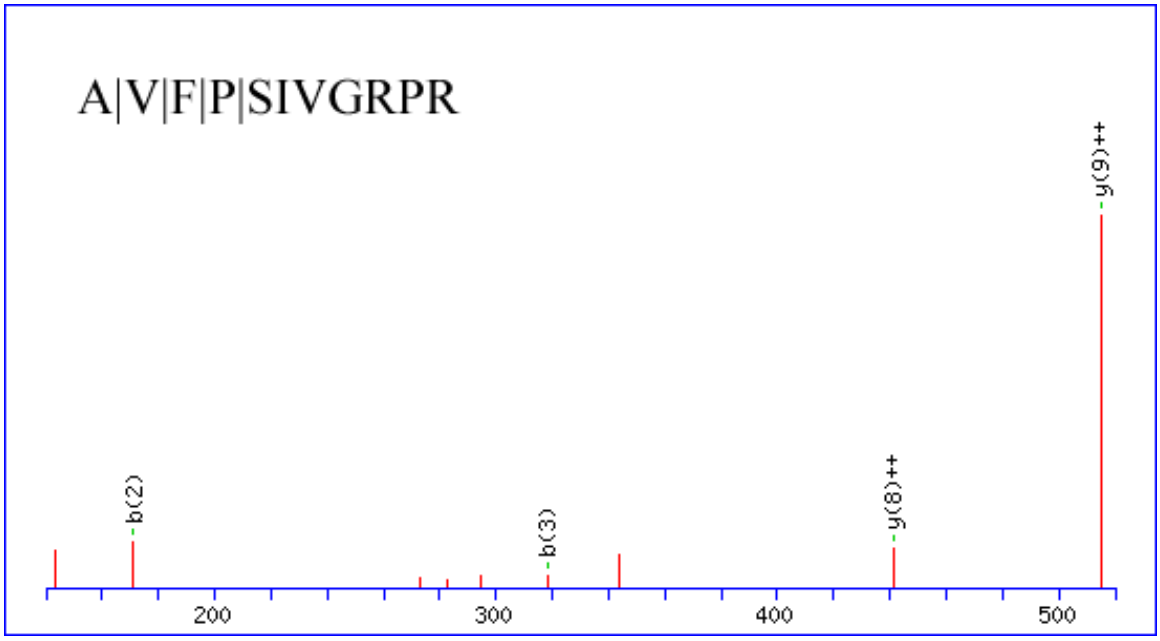
NGGTDFSCPIAKGK**VVGFFGR**YGCLPGRYWCRT

(D)

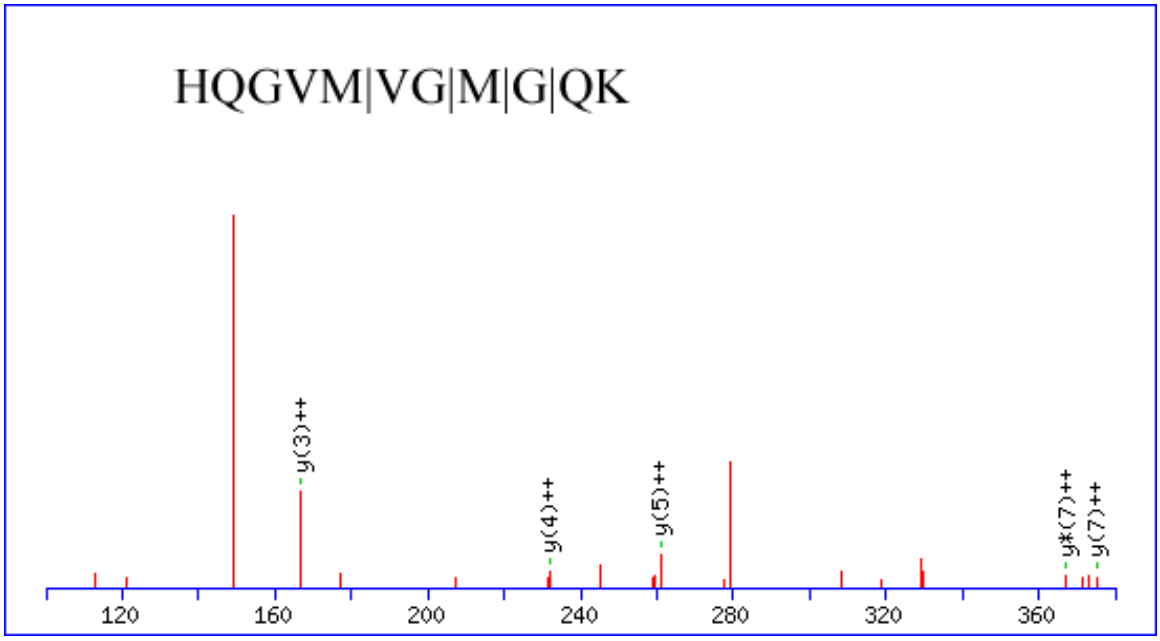
Figure A-2 (cont'd). Tandem mass spectra of polypeptides (A,B,C) matched to Lectin (*H. tuberosus*) as indicated (bold) in the primary amino acid sequence (D).

Table A-I. 1-D Hydroponic Sundance Leaf As only

Marker Region kDa	Theoretical Mass kDa	Accession no.	sp.	Protein Name	Score	Peptide (Coverage)	Peptides
45	42	gi|3127135	*M. viride*	Actin	79	4 (16%)	HQGVMVGMGQK, AVFPSIVGRPR, SYELPDGQVITIGNER, VAPEEHPILLTEAPLNPK
45	44	gi|336390	*Arabidopsis*	glyceraldehyde 3-phosphate dehydrogenase	50	2 (5%)	HIQAGASK, AVSLVLPQLK
45	40	gi|30683149	*H. annuus*	glycolate oxidase	24	1 (5%)	IAIQAGAAGIIVSNHGAR

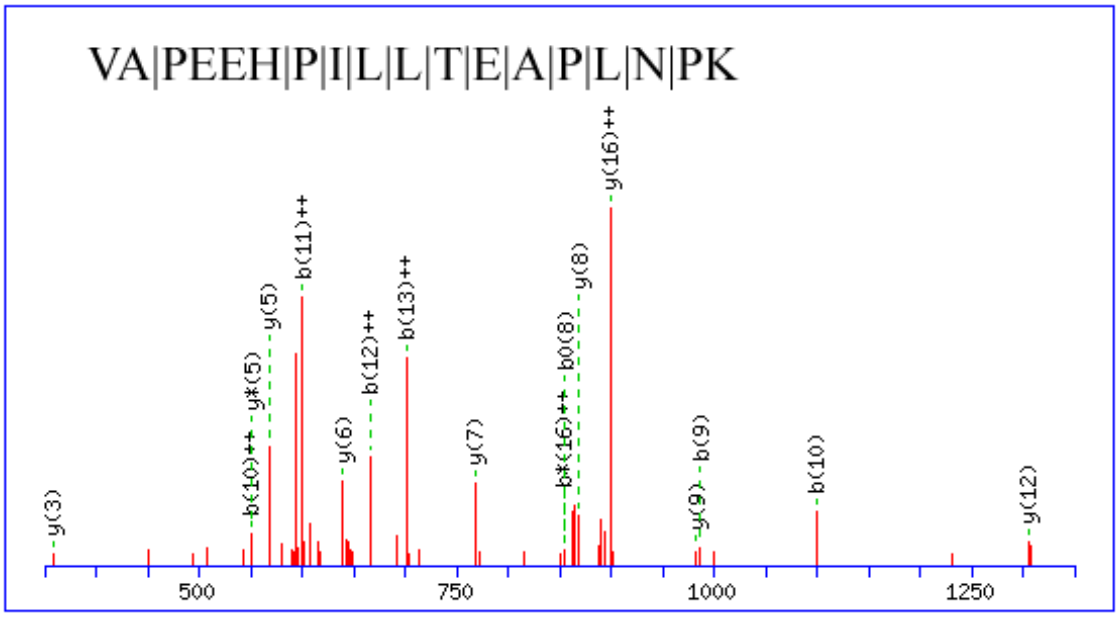


(A)

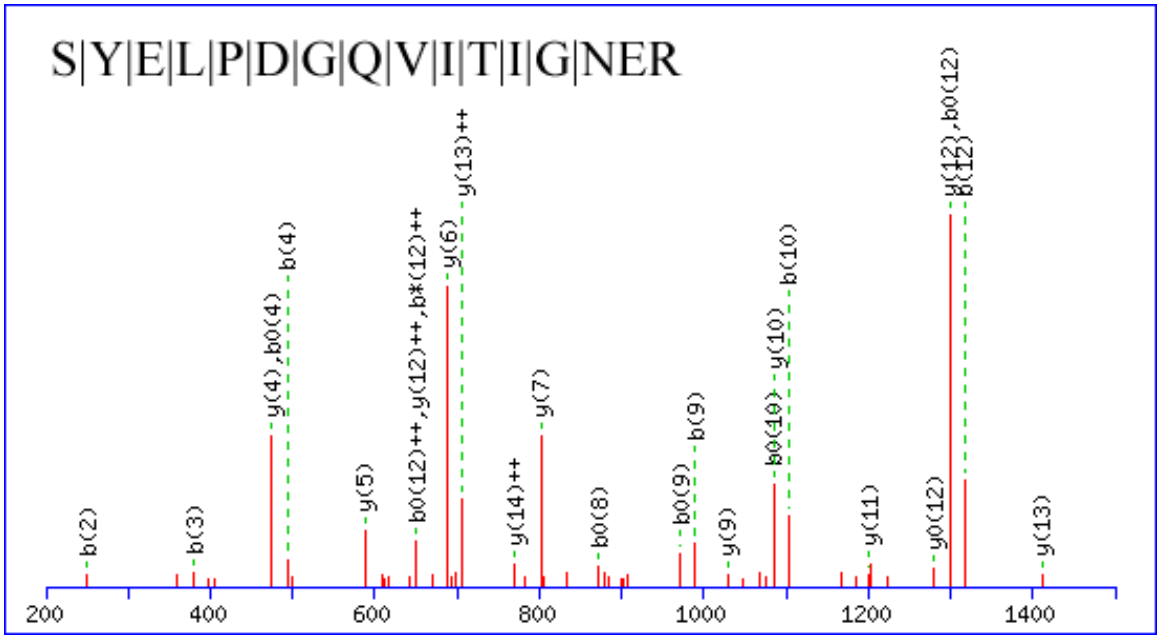


(B)

Figure A-3. Tandem mass spectra of peptides (A,B,C,D) matched to Actin (*M. viride*) as indicated (bold) in the primary amino acid sequence (E) in comparison to primary amino acid sequence of (*H. annuus*) (F).



(C)



(D)

Figure A-3 (cont'd). Tandem mass spectra of peptides (A,B,C,D) matched to Actin (*M. viride*) as indicated (bold) in the primary amino acid sequence (E) in comparison to primary amino acid sequence of (*H. annuus*) (F).

MADEGEVSALMCDNGSGMVKAGFAGDDAPR**AVFPSIVGRPRHQGVMVGMG**

QKDAYVGDEAQSKRGILTLKYPIEHGIVTNWDDMEKIWHHTFYNELR**VAPEEH**

PVLLTEAPLNPKANREKMTQIMFETFNVPAMYVAIQAVLSLYASGRTTGIVLDS

GDGVTHTVPIYEGYALPHAILRLDLAGRDLTDYLMKILTERGYSFTTTAEREIVR

DIKEKLAYVAIDYEAELATANTSSSIEK**SYELPDGQVITIGNER**FRCPEVLFNPGY

VGMESAGIHETTYNSIMRCDVDIRKDLYGNIVLSGGSTMXPGIADRMSKEITALA

PSSMKIKVVAPPERKYSVWIGGSILASLSTFQAMWIAKSEYDESGPSIVHRKCF

(E)

MAHEGEIQPLVCDNGTGMVKAGFAGDDAPR**AVFPSIVGRPRHTGVMVGMGQ**

KDAYVGDEAQSKRGILTLKYPIEHGIVSNWDDMEKIWHHTFYNELR**VAPEEHP**

VLLTEAPLNPKANREKMTQIMFETFNVPAMYVAIQAVLSLYASGRTTGIVLDSG

DGVSHTVPIYKGYALPHAILRLDLAGRDLTDSLMKILTERGYMFTTTAEREIVRD

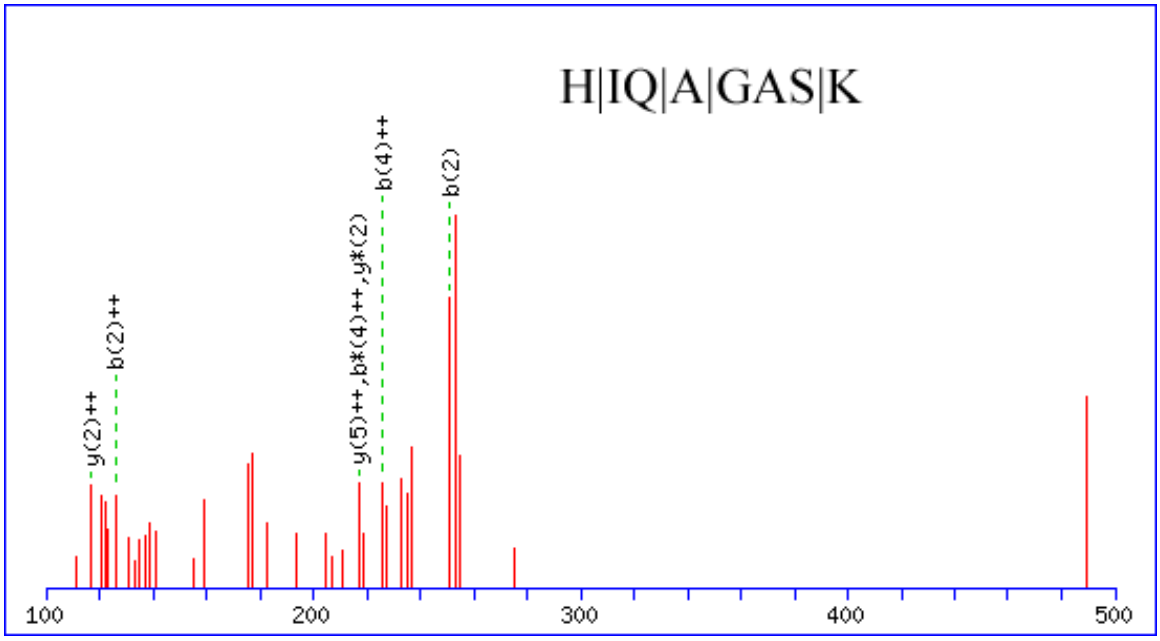
MKEKLAYVALDYEQELETAKSSSSVEK**NYELPDGQVITIGAER**FRCPEVLFQPS

LIGMEAAGIHETTYNSIMKCDVDIRKDLYGNIVLSGGSTMFPGIADRMSKEITALA

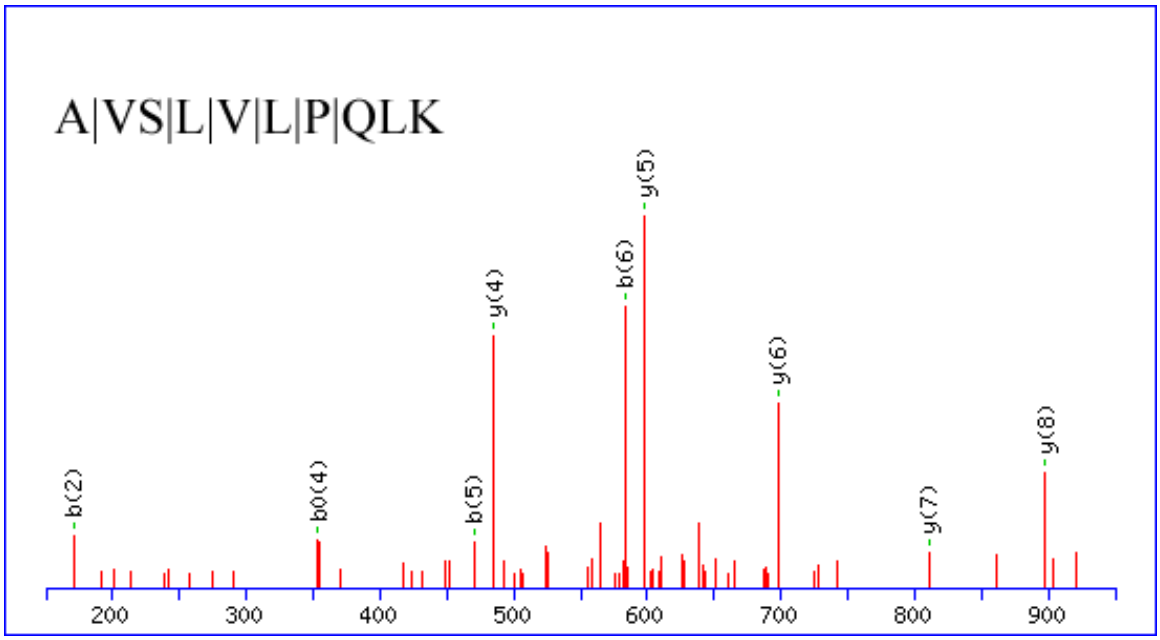
PSSMKIKVVAPPERKYSVWIGGSILASLSTFQQMWISKGEYDESGPSIVHRKCF

(F)

Figure A-3 (cont'd). Tandem mass spectra of peptides (A,B,C,D) matched to Actin (*M. viride*) as indicated (bold) in the primary amino acid sequence (E) in comparison to primary amino acid sequence of (*H. annuus*) (F).



(A)



(B)

Figure A-4. Tandem mass spectra of peptides (A,B) matched to Glyceraldehyde 3-phosphate dehydrogenase B subunit (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (C) in comparison to primary amino acid sequence of Glyceraldehyde 3-phosphate dehydrogenase (*H. annuus*) (D).

MSSIGGEASFFDAVAAQIIPKAVTTSTPVRGETVAKLKVAINGFGRIGRNFLRCW

HGRKDSPLEVVVLNDSGGVKNASHLLKYDSMLGTFKAEVKIVDNETISVDGKLI

KVVSNRDPLKLPWAELGIDIVIEGTGVFVDGPGAGK**HIQAGASK**VIITAPAKGAD

IPTYVMGVNEQDYGHDVANIISNASCTTNCLAPFAKVLDEEFGIVKGTMTTTHSY

TGDQRLLDASHRDLRRARAAALNIVPTSTGAAK**AVSLVLPQLK**GKLNGIALRVP

TPNVSVVDLVINVEKKGLTAEDVNEAFRKAANGPMKGILDVCDAPLVSVDFRCS

DVSTTIDSSLTMVMGDDMVKVVAWYDNEWGYSQRVVDLAHLVASKWPGAEA

VGSGDPLEDFCKTNPADEECKVYD

(C)

EAGAKKVLITAPGKGDIPTYVVGVNAELYNHSEAIISNASCTTNCLAPFVKVIDQ

KFGIIKGTMTTTHSYTGDQRLLDASHRDLRRARAAALNIVPTSTGAAK**AVALVL**

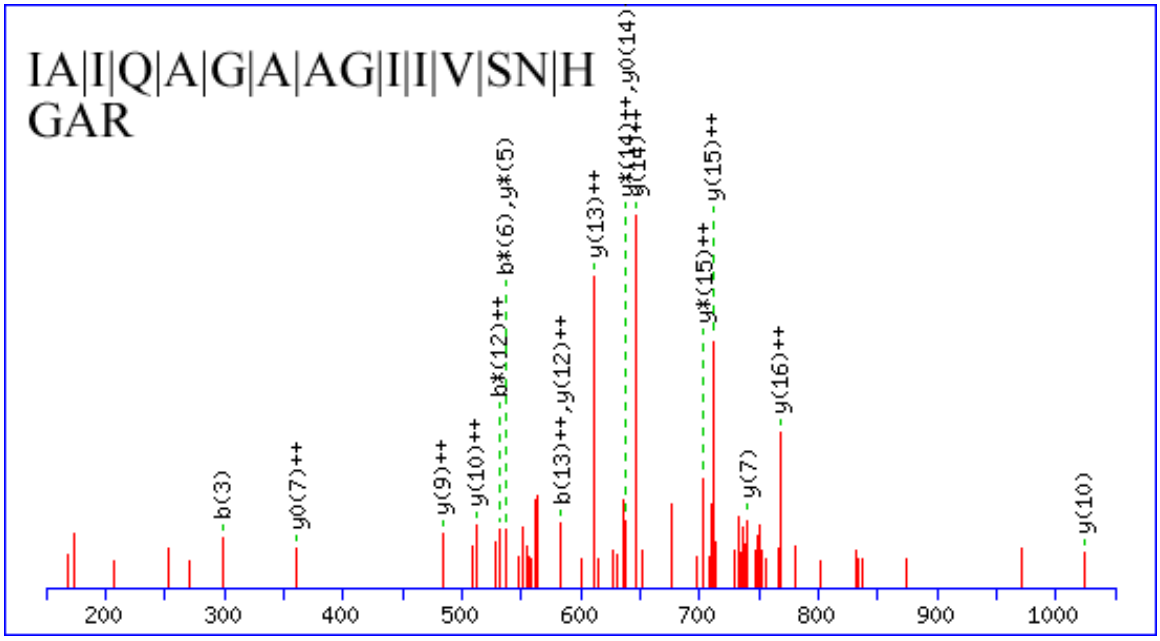
PSLKGKLNGIALRVPTPNVSVVDLVVQVSKKTFAEEVNAAFREAADNELNGILS

VCDEPLVSVDFRCSDVSSTVDSSLTMVMGDDMVKVIAWYDNEWGYSQRVVDL

ADIVANNWK

(D)

Figure A-4 (cont'd). Tandem mass spectra of peptides (A,B) matched to Glyceraldehyde 3-phosphate dehydrogenase B subunit (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (C) in comparison to primary amino acid sequence of Glyceraldehyde 3-phosphate dehydrogenase (*H. annuus*) (D).



(A)

MEITNVTEYDAIAKQKLPKMVYDYYASGAEDQWTLQENRNAFARILFRPRILID

VSKIDMTTTVLGFKISMPIMVAPTAMQKMAHPDGEYATARAASAAGTIMTLSS

WATSSVEEVASTGPGIRFFQLYVYKNRNVVEQLVRRAERAGFKAIALTVDTPRL

GRRESDIKNRFTLPPNLTLKNFEGLDLGKMDEANDSGLASYVAGQIDRTLSWKD

VQWLQTITKLPILVKGVLTGEDAR**IAIQAGAAGIIVSNHGAR**QLDYVPATISALE

EVVKATQGRIPVFLDGGVRRGTDVFKALALGASGIFIGRPVVFSLAAEGEAGVRK

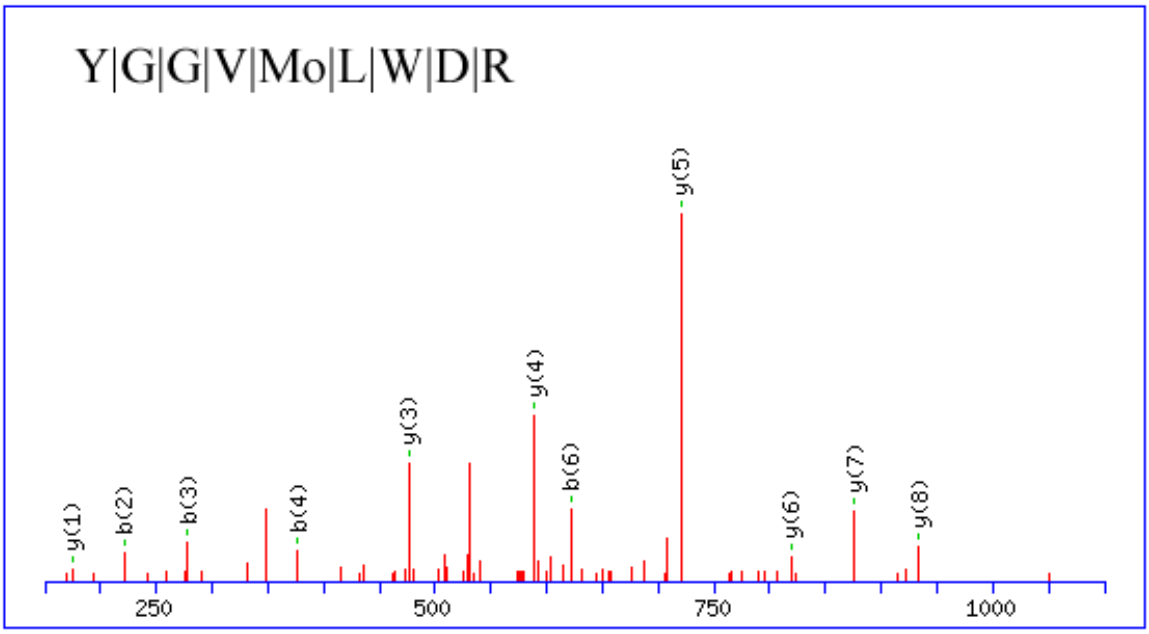
VLQMLRDEFELTMALSGCRSLKEISRNHITTEWDTPRPSARL

(B)

Figure A-5. Tandem mass spectra of peptide (A) matched to Glycolate Oxidase (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (B).

Table A-II. 1-D Sundance Leaf Hydroponic Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	Accession no.	sp.	Protein Name	Score	Peptide (Coverage)	Peptides
37	31	gi|126723930	*H. annuus*	Chitinase	66	1 (3%)	YGGVMoLWDR
37	33	gi|22571	*Arabidopsis*	33 kDa oxygen-evolving protein	51	1 (5%)	GGSTGYDNAVALPAGGR
< 37	38	gi|94502471	*H. annuus*	photosystem II protein D1	47	1 (3%)	VINTWADIINR
< 37	33	gi|15226479	*Arabidopsis*	TIM Triosephosphate Isomerase	45	3 (12%)	DEFIGKK, VASPQQAQEVHVAVR, IIYGGSVNGGNSAELAK
< 37	34	gi|30844170	*E. europaeus*	hevein-like antimicrobial peptide	68	1 (3%)	GPLQLTWNYNYAQAGR
< 37	55	gi|94502485	*H. annuus*	ATP synthase CF1 alpha subunit	23	2 (3%)	RPPGR, LIESPAPGIISR
< 37	18	gi|7489327	*H. annuus*	drought-induced protein SDi-6	31	1 (4%)	SGATGATPPAEAPLVTEEK
< 37	37	gi|21594027	*Arabidopsis*	glucosyltransferase-like protein	40	1 (2%)	FMTEVR



(A)

MEFTHPALLLLLFITVFSFLKPSTAAGIATYWGQQSDDTEGTLEAACATGNYQFV

NIAFLSTFGNNQQPVLNLAAHCDPASTCSRYSSQIKACQAQNVKVFLSIGGQRGS

YSLSSPQDAQQVADFLWNTYLGGQPATRPLGDAVLDGIDFDIEQGTDQFWSDLA

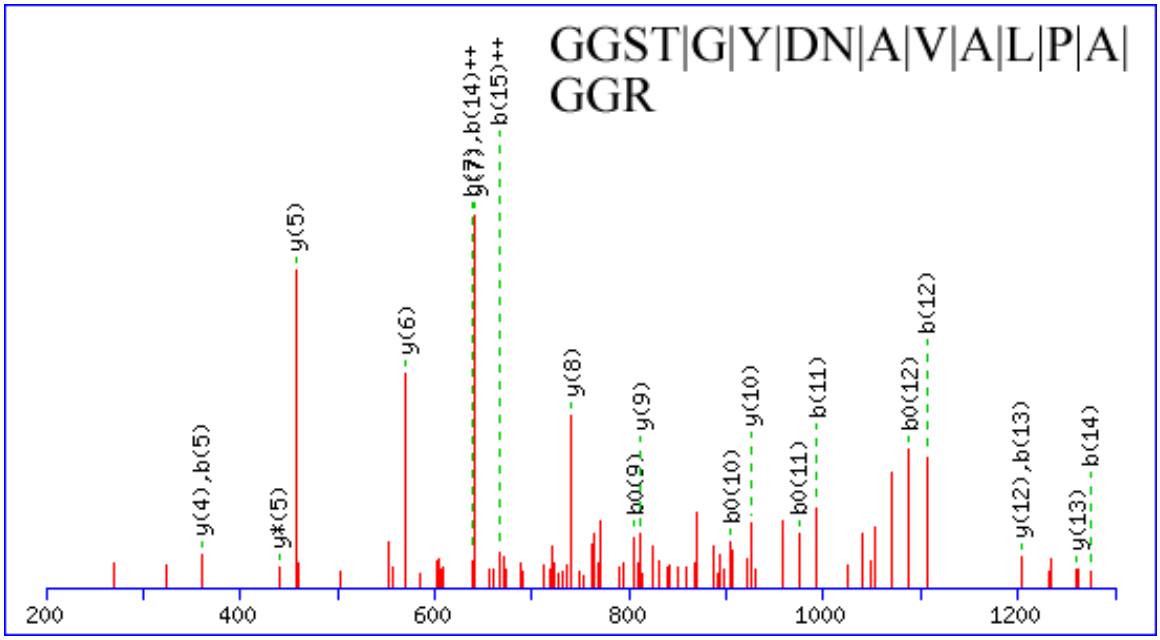
KAPAAYSSQKKVYLSAAPQCPFPSGDVRNQLLPAIREGLFDYVWVQFYNNEQCQ

YGANADALLARWNEWTQVTTNTIFLGLPAAASGAAPSGGYIPPDILTSQILPSIKS

SPK**YGGVMLWDR**FYDKQSGYSDAIKGSIN

(B)

Figure A-6. Tandem mass spectra of polypeptide (A) from Chitinase (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).



(A)

MAASLQSTATFLQSAKIATAPSRGSSHLRSTQAVGKSFGLETSSARLTCSFQSDFK

DFTGKCSDAVKIAGFALATSALVVSGASAEGAPKRLTYDEIQSKTYMEVKGTGT

ANHSPTIDGGSETFSFKPGKYAGKKFCFEPTSFTVKADSVSKNAPPEFQNTKLMT

RLTYTLDEIEGPFEVASDGSVNFKEEDGIDYAAVTVQLPGGERVPFLFTVKQLDA

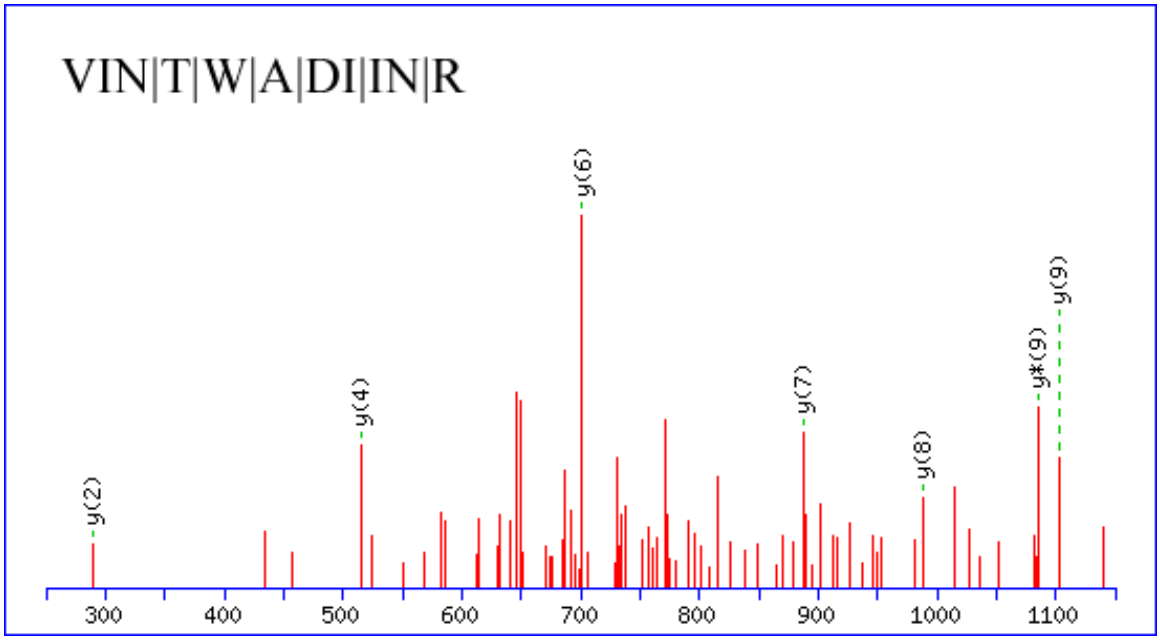
SGKPDSFTGKFLVPSYRGSSFLDPKGR**GGSTGYDNAVALPAGGR**GDEEELVKE

NVKNTAASVGEITLKVTKSKPETGEVIGVFESLQPSDTDLGAKVPKDVKIQGVW

YGQLE

(B)

Figure A-7. Tandem mass spectra of peptide (A) matched to 33 kDa Oxygen-evolving protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (B).



(A)

MTAILERRESESLWGRFCNWITSTENRLYIGWFGVLMIPTLLTATSVFIIAFIAAPP

VDIDGIREPVSGSLLYGNNIISGAIIPTSAAIGLHFYPIWEAASVDEWLYNGGPYELI

VLHFLLGVACYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGQGSFS

DGMPLGISGTFNFMIVFQAEHNILMHPFHMLGVAGVFGGSLFSAMHGSLVTSSLI

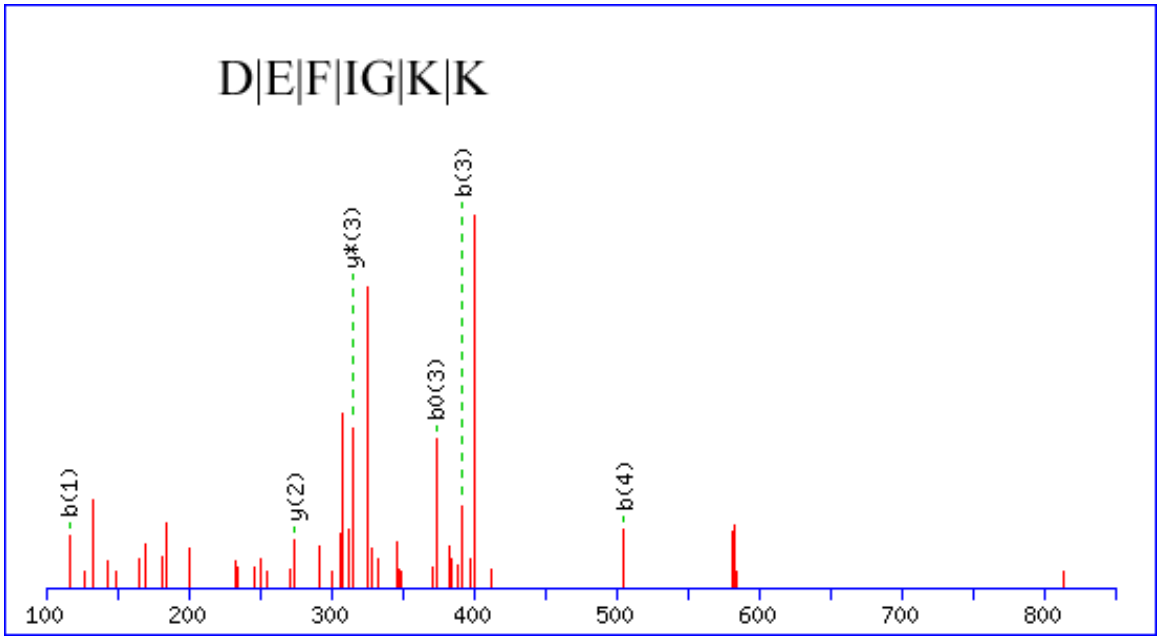
RETTENESANEGYRFGQEEETYNIVAAHGYFGRLIFQYASFNNSRSLHFFLAAWP

VVGIWFTALGISTMAFNLNGFNFNQSVVDSQGR**VINTWADIINR**ANLGMEVMH

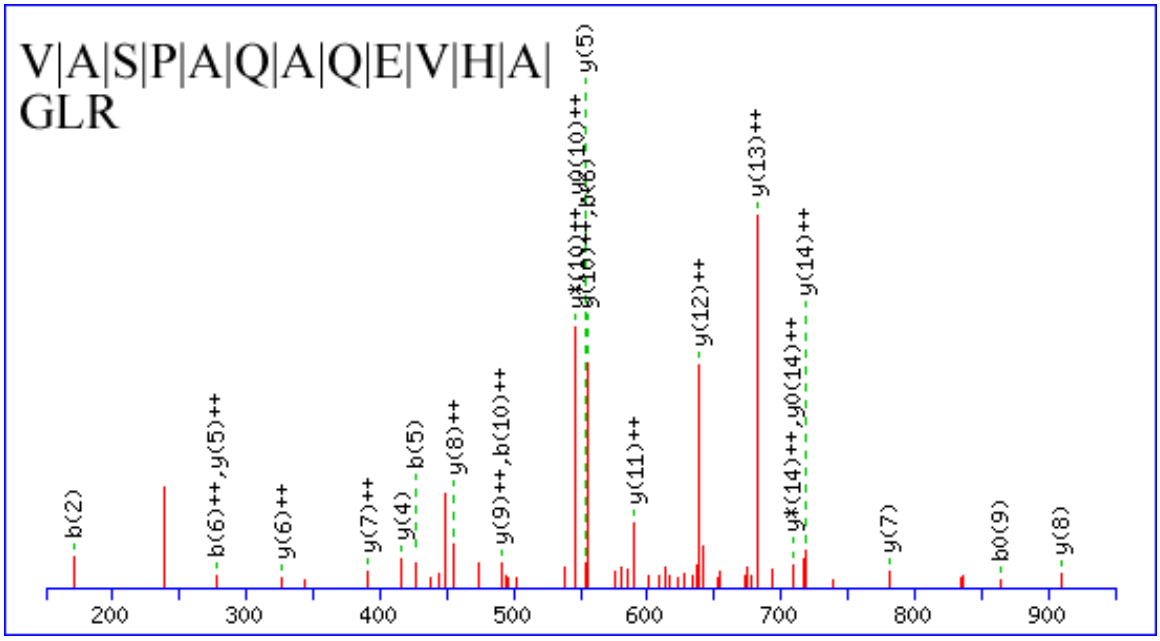
ERNAHNFPLDLAAIEAPSTNG

(B)

Figure A-8. Tandem mass spectra of peptide (A) matched to photosystem II protein D1 (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).

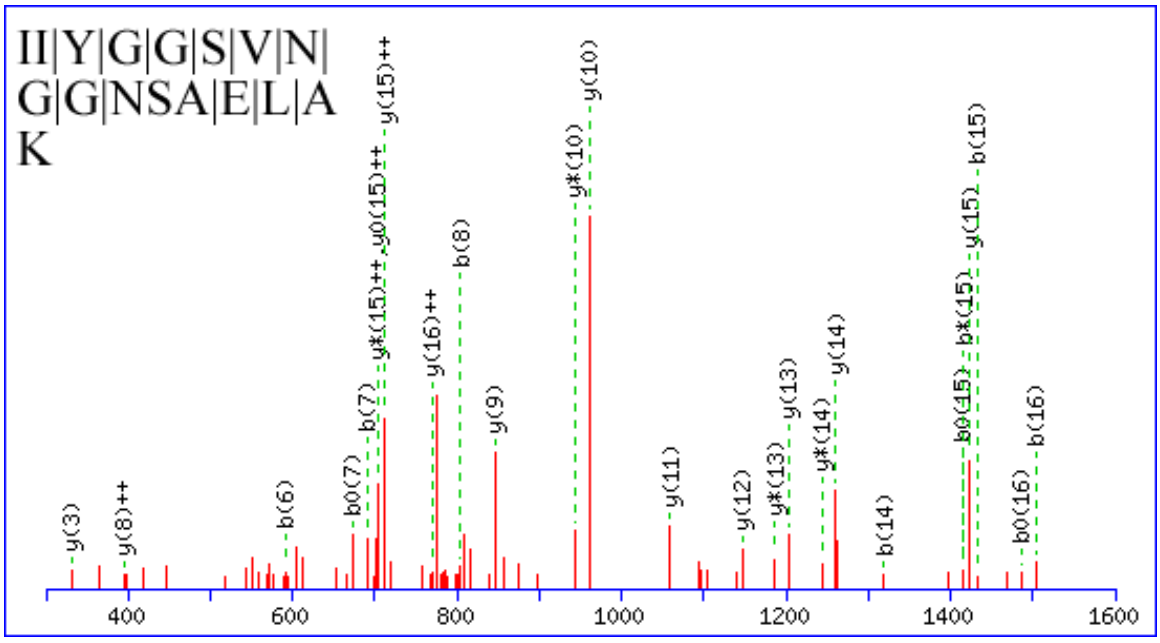


(A)



(B)

Figure A-9. Tandem mass spectra of peptides (A,B,C) matched to Triosephosphate Isomerase (TIM) (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (D).



(C)

MAATSLTAPPSFSGLRRISPKLDAAAVSSHQSFFHRVNSSTRLVSSSSSSHRSPRG

VVAMAGSGKFFVGGNWKCNGTKDSIAKLISDLNSATLEADVDVVVSPPFVYIDQ

VKSSLTDRIDISGQNSWVGKGGAFTGEISVEQLKDLGCKWVILGHSERRHVIGEK

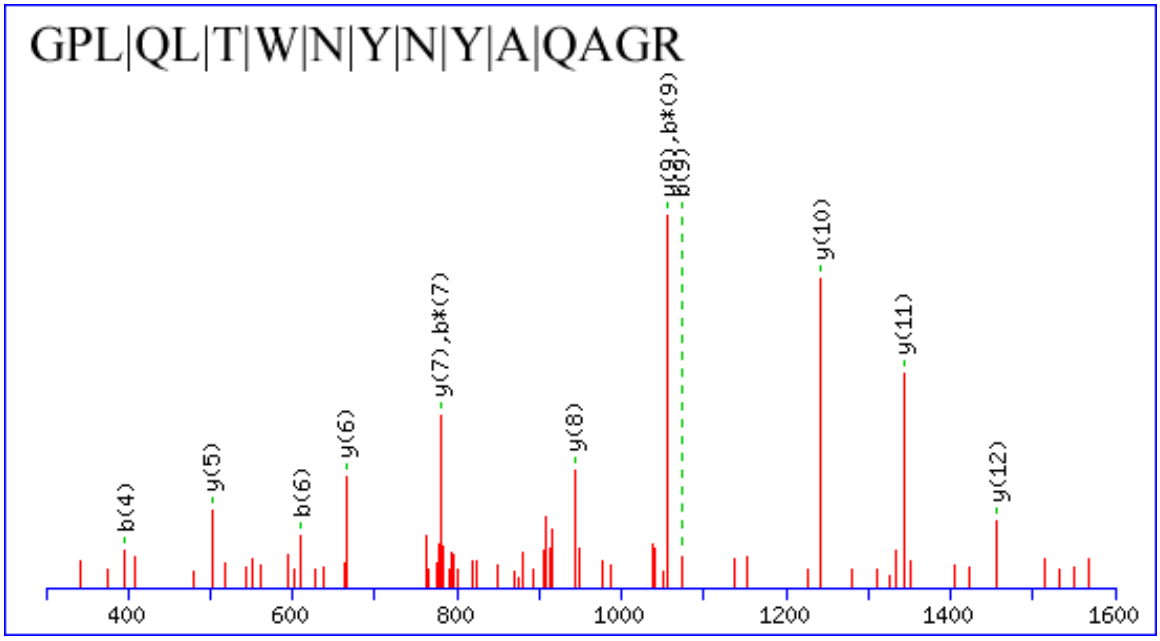
DEFIGKKAAYALSEGLGVIACIGEKLEEREAGKTFDVCFAQLKAFADAVPSWDN

IVVAYEPVWAIGTGK**VASPQQAQEVHVAVR**GWLKKNVSEEVASKTR**IIYGGSV**

NGGNSAELAKEEDIDGFLVGGASLKGPEFATIVNSVTSKKVAA

(D)

Figure A-9 (cont'd). Tandem mass spectra of peptides (A,B,C) matched to Triosephosphate Isomerase (TIM) (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (D).



(A)

MKYLWVFIVFSIAVLSLACSAQQCGRQAGNRRCPNNLCCSQFGYCGRTNEYCCT

GFGCQSNCRRCGVRTVGEDVVGDIGGIISKGMFNNILKHRDDDACEGKGFYTYE

AFVAAARSFPAFGSTGDDTTRKREIAAFLAQTSHETSGGRPSAPDGPYAWGYCF

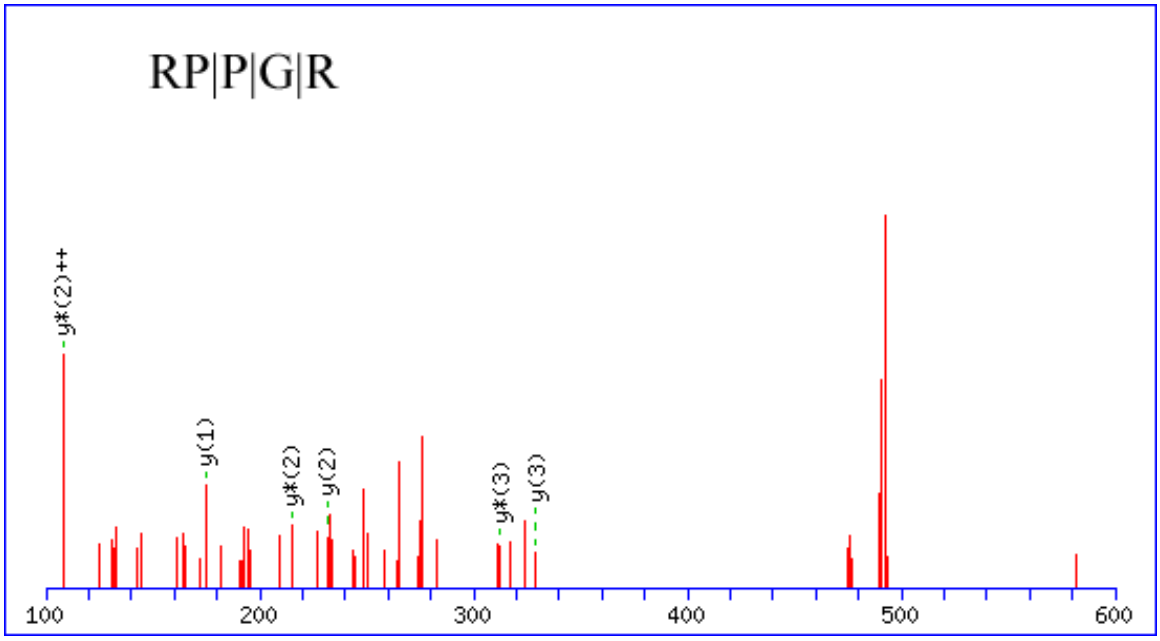
VKERNPPSKYCDTITPCPKSYYGR**GPLQLTWNYNYAQAGR**AIGVDLLNNPDLV

ATDAVTSFKTAIWFWMTAHSSKPSCHDVITGSWRPSASDNSVRHVPDYAVVTNI

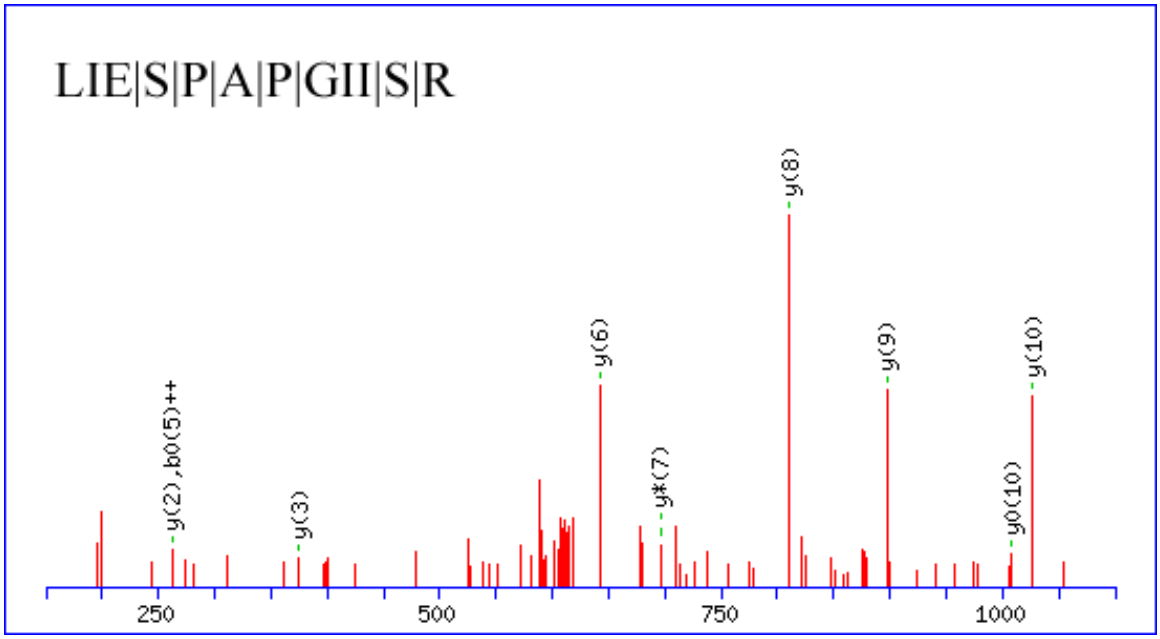
INGEIEYGKSRNPQVEDRIEFFKRYCQILGVSPGKF

(B)

Figure A-10. Tandem mass spectra of peptide (A) matched to hevein-like antimicrobial (*E. europaeus*) as indicated (bold) in the primary amino acid sequence (B).



(A)



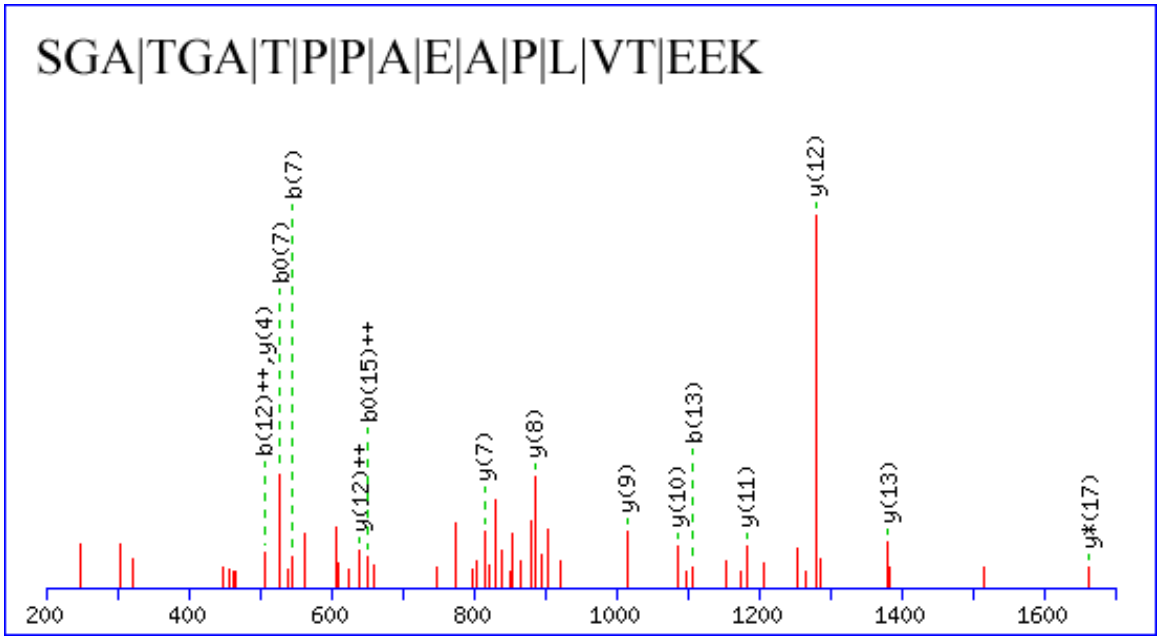
(B)

Figure A-11. Tandem mass spectra of peptides (A,B) matched to ATP synthase SF1 alpha subunit (*H. annuus*) as indicated (bold) in the primary amino acid sequence (C).

MVTIQADEISNIIRERIEQYNREVKIVNTGTVLQVGDGIARIHGLDEVMAGELVEF
QEGTIGIALNLESTNVGVVLMGDGLLIQEGSSVKATGKIAQIPVSEAYLGRVINLA
KPIDGRGEISSSEYR**LIESPAPGIISR**RSVYEPLQTGLIAIDSMIPIGRGQRELIIGDR
QTGKTAVATDTILNQQGQNVICVYVAIGQKASSVAQVVTTFQEKGAMEYTIVVA
ETADSPATLQYLAPYTGAALAEYFMYRERHTLIIYDDLSKQAQAYRQMSLLLR**R**
PPGREAYPGDVFYLHSRLLERAAKLSSLLGEGSMTALPIVETQSGDVSAYIPTNVI
SITDGQIFLSADLFNAGIRPAINVGISVSRVGSAAQIKAMKQVAGKLKLELAQFAE
LEAFAQFASDLDKATQNQLARGQRLRELLKQSQSAPLAVEEQILTIYTGTNGYLD
SLEVGQVRKFLVELRTYLKTNKPQFQEIISSTKIFTEEAEAILKEAIQEQRERFILQE
QAA

(C)

Figure A-11 (cont'd). Tandem mass spectra of peptides (A,B) matched to ATP synthase SF1 alpha subunit (*H. annuus*) as indicated (bold) in the primary amino acid sequence (C).



(A)

RGIPVRHGSRPTGKNMDSFMCNIKTVATGGHDETKTNTDQPSNTSTDLLSSAKL

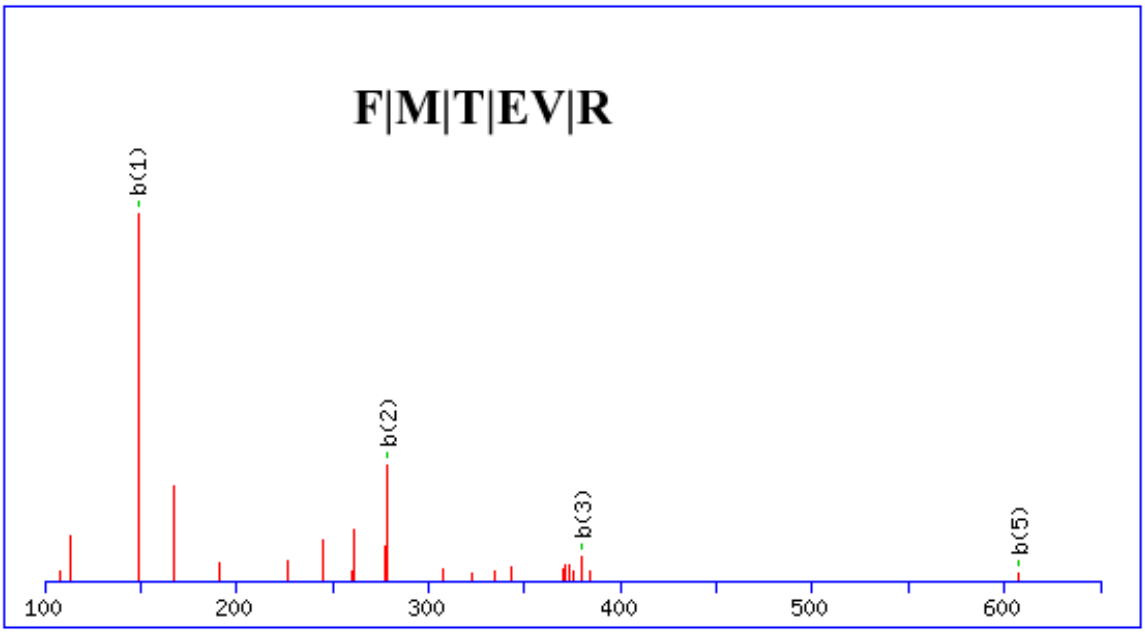
VAEAAQSAATNKTDQIDKQKVAGATADLLDSSKEYGKFDESQGVGQYIKQADD

YLHKYEK**SGATGATPPAEAPLVTEEK**KAEAPPGVEEKGGKDESESGIGAGDAI

KAAGSFFK

(B)

Figure A-12. Tandem mass spectra of peptide (A) matched to Drought-induced protein SDi-6 (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).



(A)

MGTPVEVSKLHFLLFPFMAHGHMIPTLDMAKLFATKGAKSTILTTPLNAKLFFEK

PIKNLNPGLEIDIQIFNFPCVELGLPEGCENVDFFTSNNNDDKNEMIVKFFFSTRFF

KDQLEKLLGTTRPDCLIADMFFPWATEAAGKFNVPRLVFHGTGYFSLCAGYCIG

VHKPQKRVASSSEPFVIPELPGNIVITEEQIIDGDGESDMGK**FMTEVR**ESEVKSSG

VVLNSFYELEHDYADFYKSCVQKRAWHIGPLSVYNRGFEEKAERGKKANIDEAE

CLKWLDSKKPNSVIYVSFGSVAFFKNEQLFEIAAGLEASGTSFIWVVRKTKEKEE

WLPEGFEERVKGKGMIIRGWAPQVLILDHQATCGFVTHCGWNSLLEGVAAGLP

MVTWPVAAEQFYNEKLVTQVLRTGVSVGAKKNVRTTGDFISREKVVKAVREVL

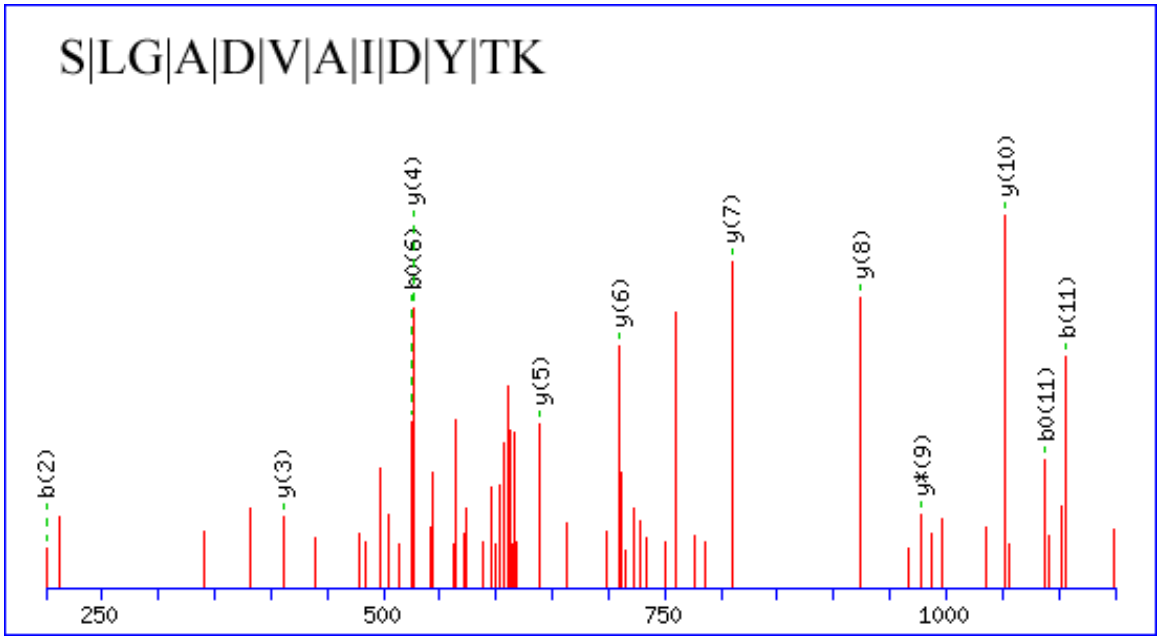
VGEEADERRERAKKLAEMAKAAVEGGSSFNDLNSFIEEFTS

(B)

Figure A-13. Tandem mass spectra of peptide (A) matched to Glucosyltransferase-like protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (B).

Table A-III. 1-D Sundance Stem Hydroponic Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	Accession no.	sp.	Protein Name	Score	Peptide (Coverage)	Peptides
39-47	33	gi|14532287	*H. annuus*	Quinone oxido-reductase-like phenylcoumaran benzylic ether	54	1 (4%)	SLGADVAIDYTK
39-47	33	gi|7578903	*T. heterophylla*	reductase homolog TH3	40	2 (3%)	VVILGDGNAK, DKVVILGDGNAK



(A)

MKAWKYDEYGSVDVLKLATDVAVPEIKDDQVLVKVVAAAVNPVDYKRRLGYF

KAIDSPLPIIPGFDVLGVVLKVGSQVKDLKEGDEVYGDINDKGLEGPTQFGTLAE

YTAVEERLLALKPKNLDFIQAAALPLAIETAYEGLERAKFSEGKTILVLNGAGGV

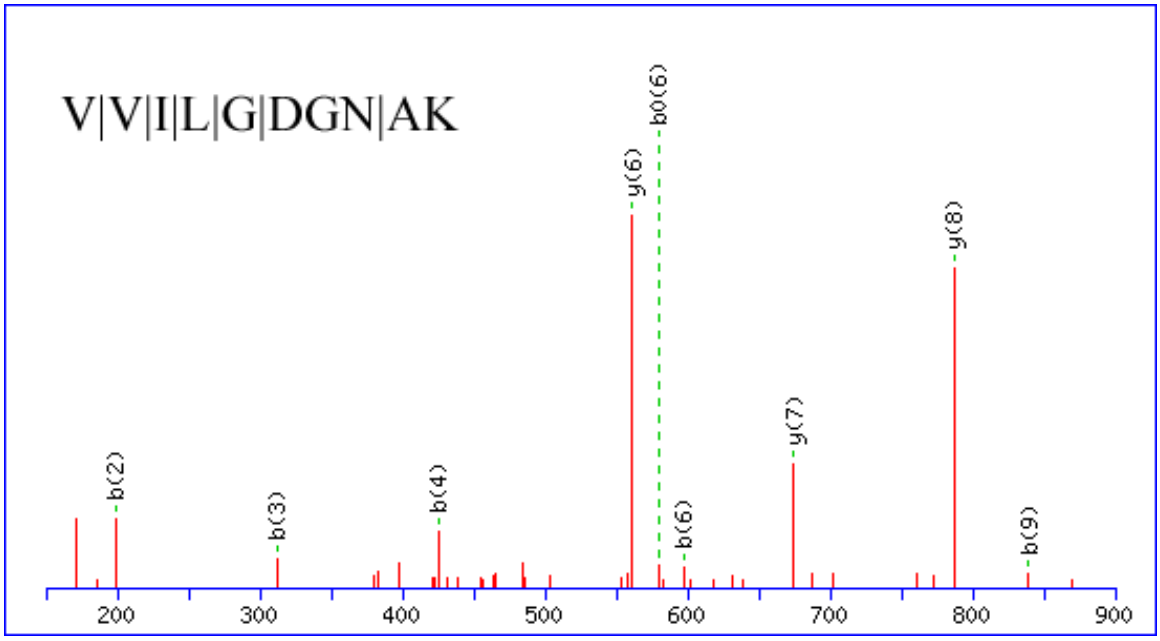
GSFIIQLAKHVYGASKVAATSSTGKLELLK**SLGADVAIDYTK**ENFEDLPDKYDV

VYDAIGQPEKALKAVNETGVAVSITGPIPPPGFSFILTSDGSILTKLNPYLESGKIKS

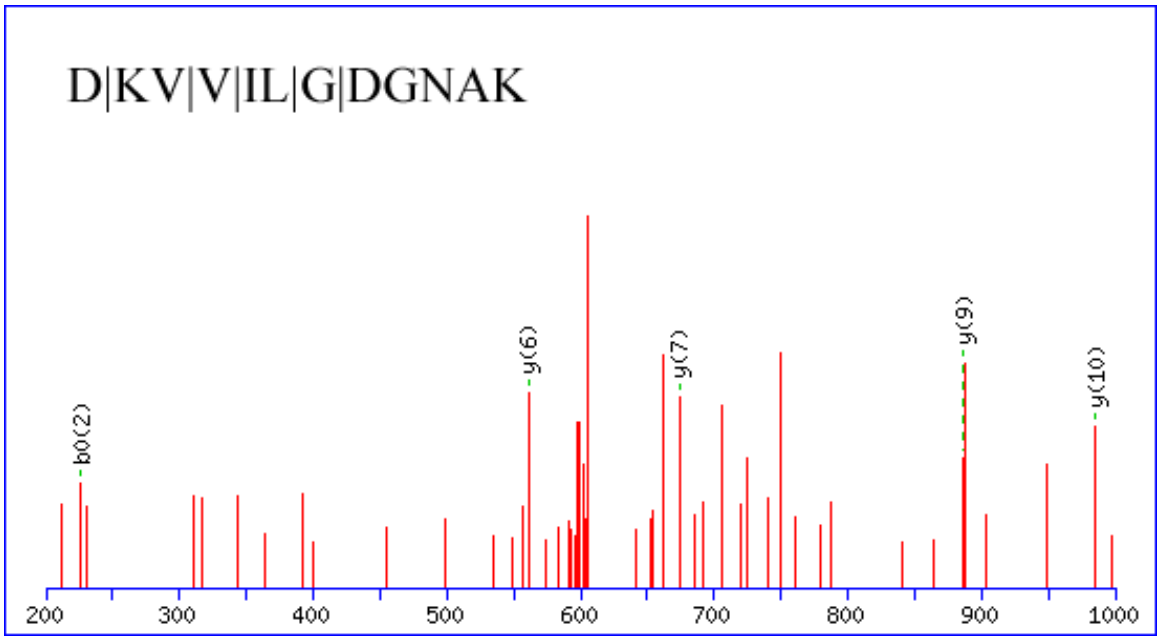
DNDPKSPFPFDKLNEAYAYLESTGPTGKVVIYPIP

(B)

Figure A-14. Tandem mass spectra of peptide (A) matched to Quinone oxido-reductase-like protein (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).



(A)



(B)

Figure A-15. Tandem mass spectra of peptide (A,B) matched to Phenylcoumaran benzylic ether reductase homolog (*T. heterophylla*) as indicated (bold) in the primary amino acid sequence (C).

MGSKSKILIIGATGYIGRQVAKASLALSHPTFLLVRDSPASSKPEKAQLLDSFKAS

GANILKGSLEDHASLVEAVKKVDVVISTVGGEQIANQFNIIKAIKEVGTIKRFLPSE

FGNDVDNVHAVEPAKSVFELKAQVRRAIEAESIPYTYVSSNCFAGYFLPSFAQAG

LTSPPR**DKVVILGDGNAK**AVYVKEEDIGTFAIKAADDPRTLNKTLYLRLPANTL

SFNELVALWEKKIGKTLEKVYVPEEHVVKLIAETPFPANIVIAIGHSIFVKGDQTN

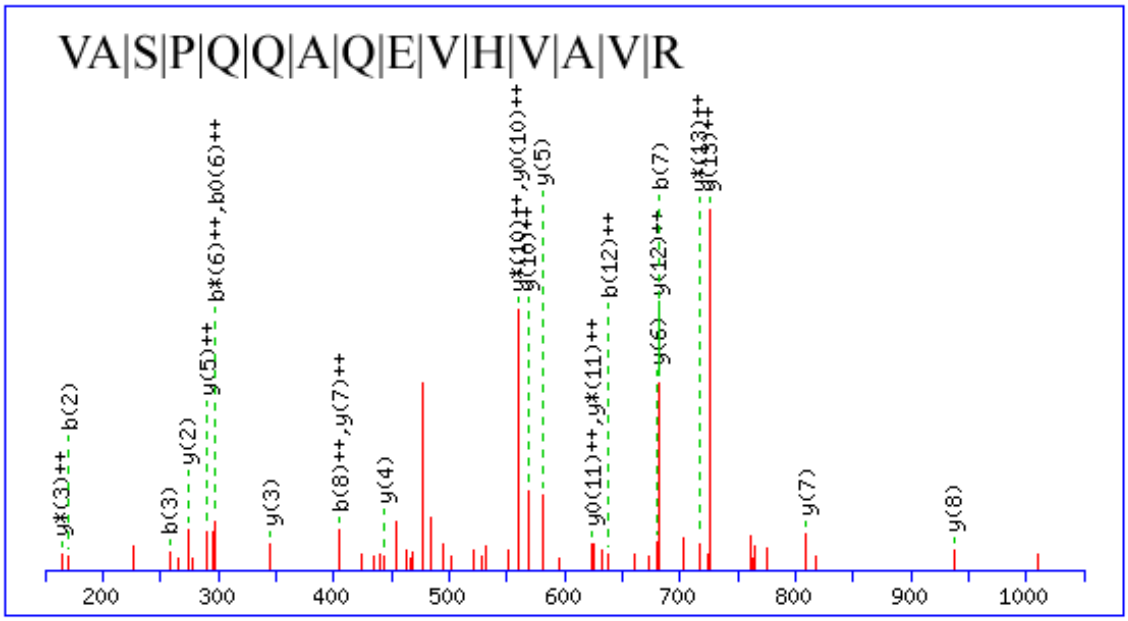
FDIGPDGVEGSLLYPDVKYTTVDEYLSAFV

(C)

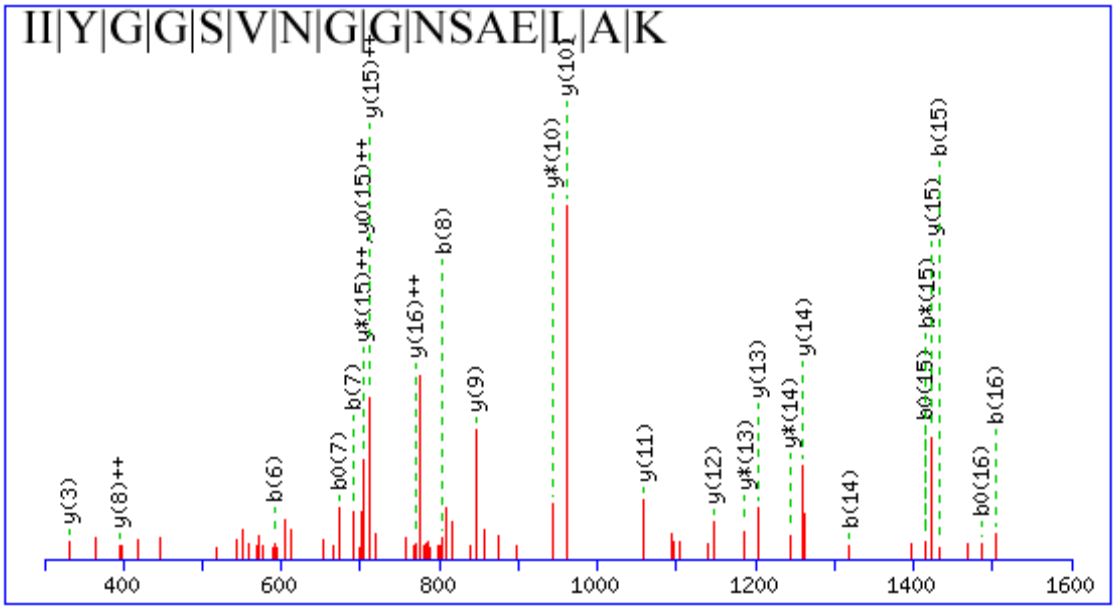
Figure A-15 (cont'd). Tandem mass spectra of peptide (A,B) matched to Phenylcoumaran benzylic ether reductase homolog (*T. heterophylla*) as indicated (bold) in the primary amino acid sequence (C).

Table A-IV. 1-D Sundance Leaf Hydroponic (soluble fraction) Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	Accession no.	sp.	Protein Name	Score	Peptide (Coverage)	Peptides
25-35	33	gi|15226479	*Arabidopsis*	TIM	49	3 (10%)	VASPQQAQEVHVAVR, IIYGGSVNGGNSAELAK, TRIIYGGSVNGGNSAELAK
25-35	21	gi|256120	*L. sativa*	TRIOSEPHOSPHATE ISOMERASE	79	2(14%)	VASPAQAQEVHAGLR, VAYALSQGLKVIACVGETLEQR

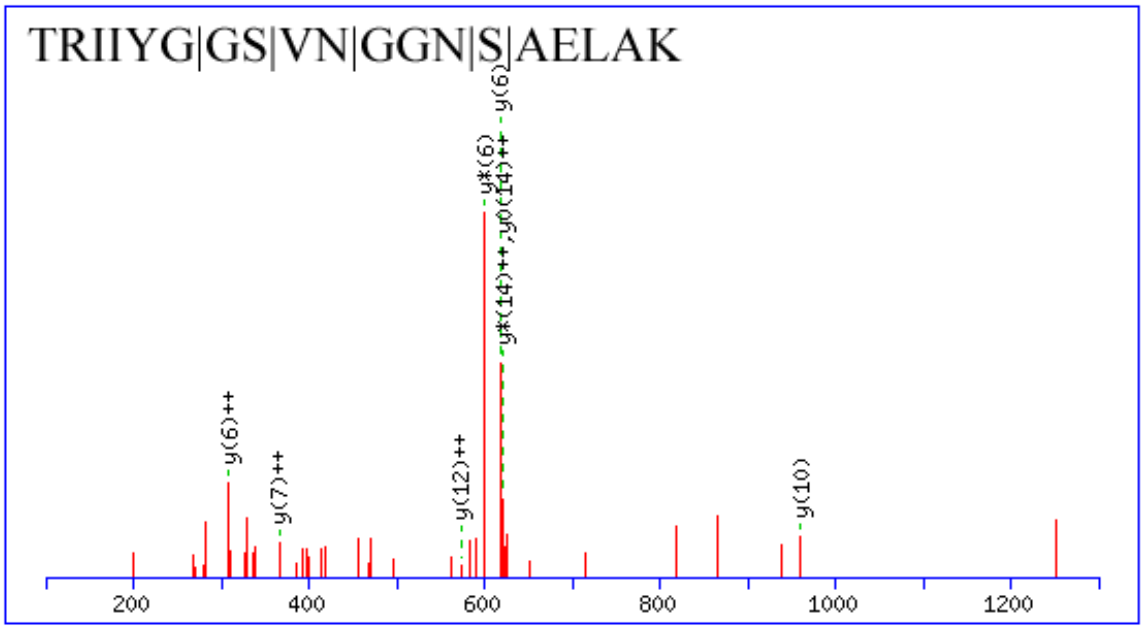


(A)



(B)

Figure A-16. Tandem mass spectra of peptides (A,B,C) matched to Triosephosphate Isomerase (TIM) (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (D).



(C)

MAATSLTAPPSFSGLRRISPKLDAAAVSSHQSFFHRVNSSTRLVSSSSSSHRSPRG

VVAMAGSGKFFVGGNWKCNGTKDSIAKLISDLNSATLEADVDVVVSPPFVYIDQ

VKSSLTDRIDISGQNSWVGKGGAFTGEISVEQLKDLGCKWVILGHSERRHVIGEK

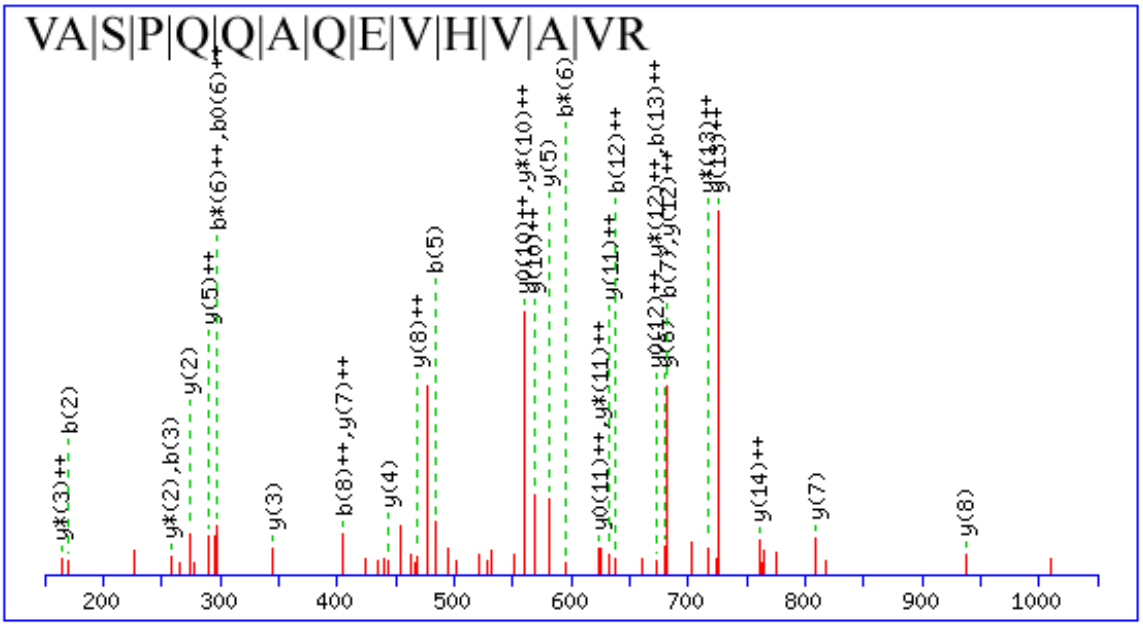
DEFIGKKAAYALSEGLGVIACIGEKLEEREAGKTFDVCFAQLKAFADAVPSWDNI

VVAYEPVWAIGTGK**VASPQQAQEVHVAVR**GWLKKNVSEEVASK**TRIIYGGSV**

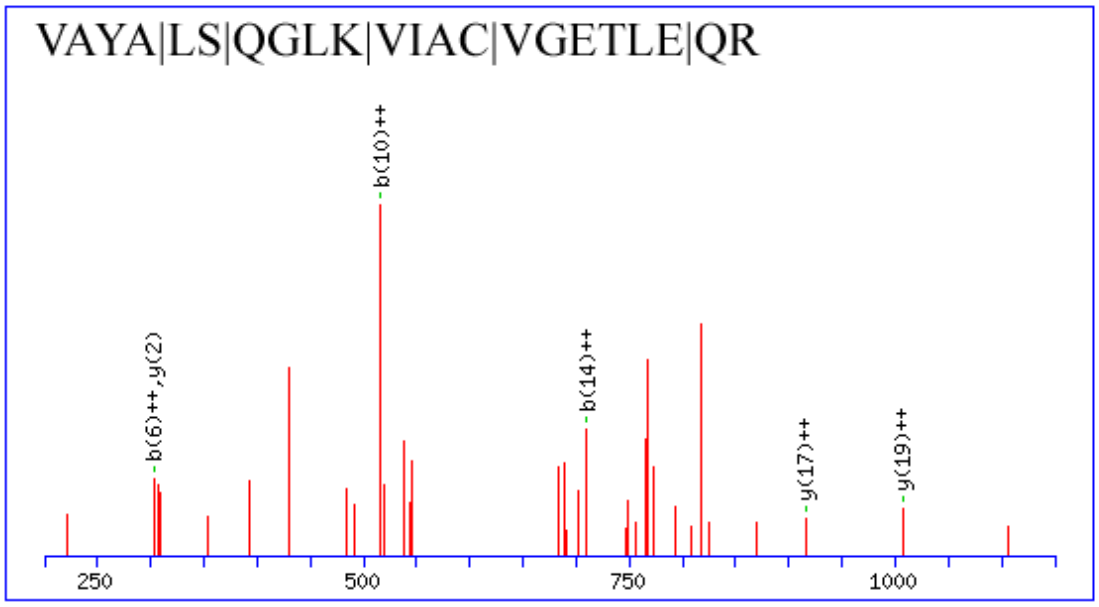
NGGNSAELAKEEDIDGFLVGGASLKGPEFATIVNSVTSKKVAA

(C)

Figure A-16 (cont'd). Tandem mass spectra of peptides (A,B,C) matched to Triosephosphate Isomerase (TIM) (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (D).



(A)



(B)

Figure A-17. Tandem mass spectra of peptides (A,B) matched to Triosephosphate Isomerase (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (C).

IQVAAQNCWVKKGGAFTGEVSAEMLANLGVPWVILGHSERRALLNETNEFVGD

K**VAYALSQGLKVIACVGETLEQR**EAGTTMEVVAAQTKAIADKISSWDNVVLA

YEPVWAIGTGK**VASPAQAQEVHAGLR**KWFCDNVSAEVSASTRIIYGGSVSGSN

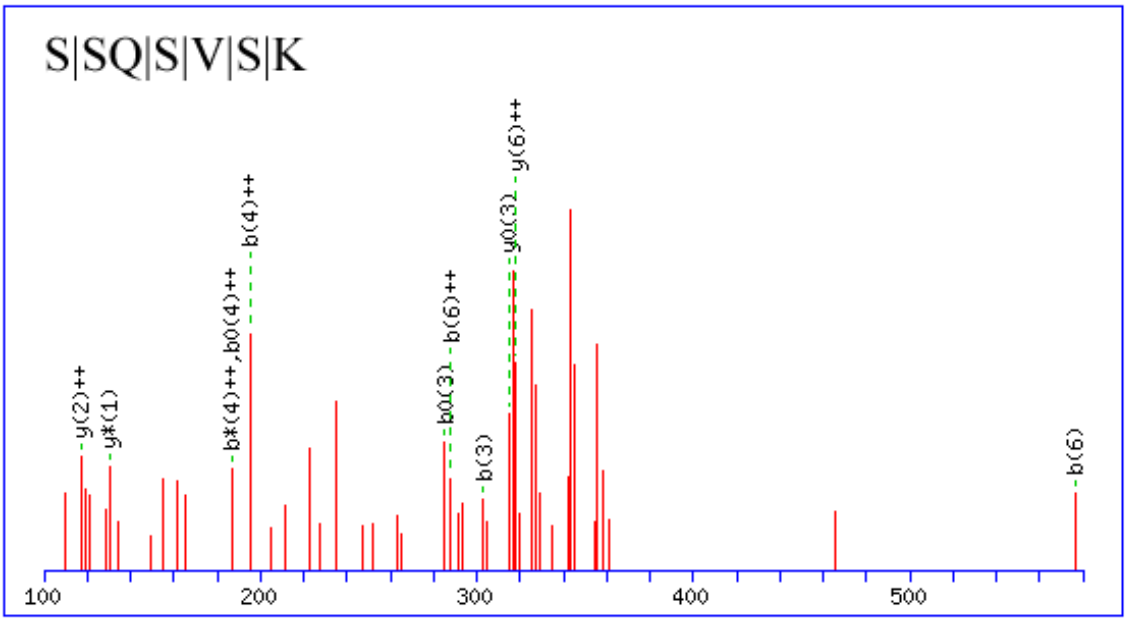
CKELGGQTDVDGFLVGGASLKPEFIDIIKAAEVKKSA

(C)

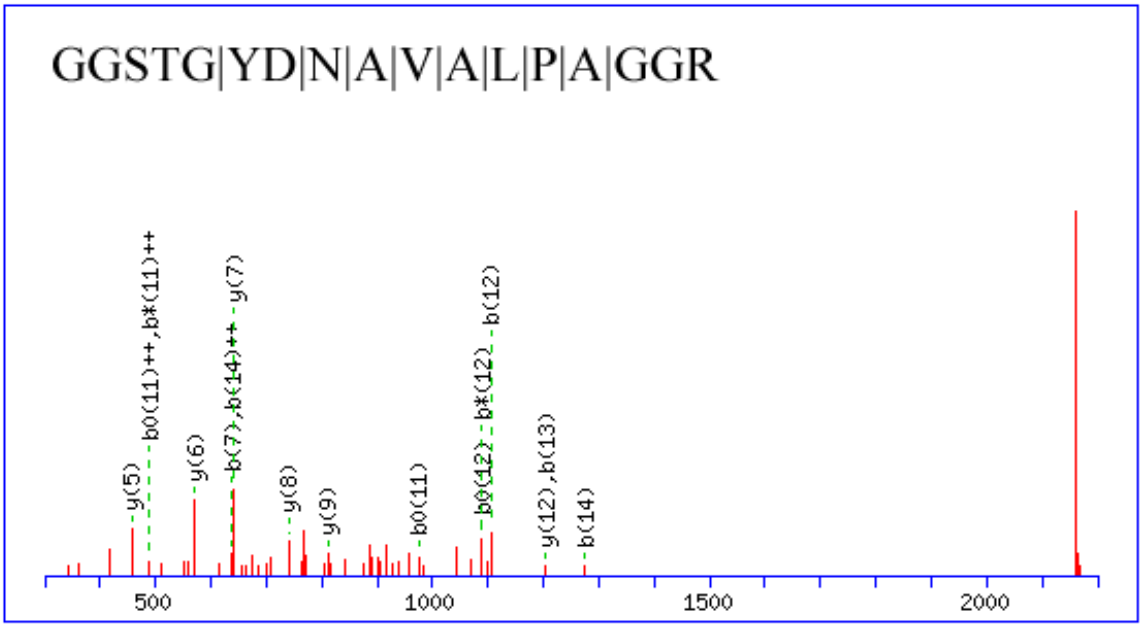
Figure A-17 (cont'd). Tandem mass spectra of peptides (A,B) matched to Triosephosphate Isomerase (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (C).

Table A-V. 1-D Sundance Leaf Soil 25-35 kDa EDTA samples Cd, Cr, Ni, Pb

EDTA	Theroetical Mass	Accession no.	sp.	Protein Name	score	Peptide (Coverage)	Peptides
ctrl	35	gi|131385	*Arabidopsis*	OEE1	25	2 (6%)	SSQSVKSK, GGSTGYDNAVALPAGGR
0.1 EDTA	32	gi|126723930	*H. annuus*	Chitinase	59	1 (2%)	YGGVMLWDR
0.3 EDTA	32	gi|126723930	*H. annuus*	Chitinase	70	1 (2%)	YGGVMLWDR



(A)



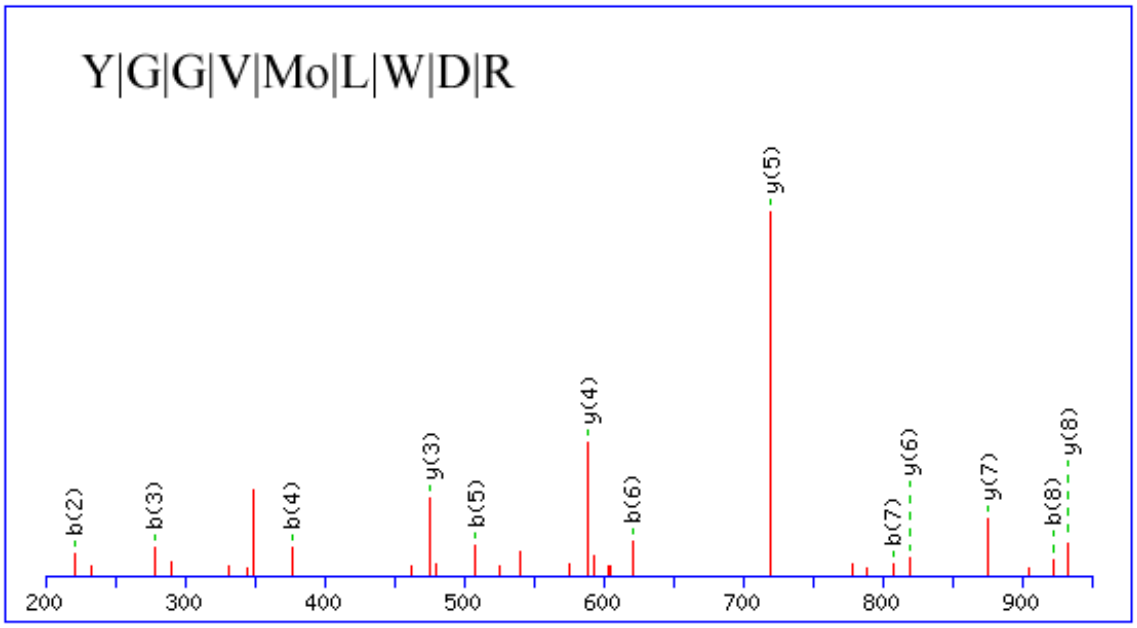
(B)

Figure A-18. Tandem mass spectra of peptide (A,B) matched to Oxygen-evolving enhancer protein 1, chloroplast precursor (OEE1) (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (C).

MAASLQAAATLMQPTKVGGVSARNNLQLR**SSQSVSK**AFGLEPSASRLSCSLQTD
LKDFAQKCTDAAKIAGFALATSALVVSGANAEGVPKRLTFDEIQSKTYMEVKGT
GTANQCPTIDGGVDSFAFKPGKYNAKKFCLEPTSFTVKAEGVSKNSAPDFQKTK
LMTRLTYTLDEIEGPFEVSPDGTVKFEEKDGIDYAAVTVQLPGGERVPFLFTIKQL
VASGKPESFSVDFLVPSYRGSSFLDPKGR**GGSTGYDNAVALPAGGR**GDEEELQ
KENVKNTASLTGKITFTVTKSNPQTGEVIGVFESIQPSDTDLGAKTPKDVKIQGIW
YAQLES

(C)

Figure A-18 (cont'd). Tandem mass spectra of peptide (A,B) matched to Oxygen-evolving enhancer protein 1, chloroplast precursor (OEE1) (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (C).



(A)

MEFTHPALLLLLFITVFSFLKPSTAAGIATYWGQQSDDTEGTLEAACATGNYQFV

NIAFLSTFGNNQQPVLNLAAHCDPASTCSRYSSQIKACQAQNVKVFLSIGGQRGS

YSLSSPQDAQQVADFLWNTYLGGQPATRPLGDAVLDGIDFDIEQGTDQFWSDLA

KAPAAYSSQKKVYLSAAPQCPFPSGDVRNQLLPAIREGLFDYVWVQFYNNEQCQ

YGANADALLARWNEWTQVTTNTIFLGLPAAASGAAPSGGYIPPDILTSQILPSIKS

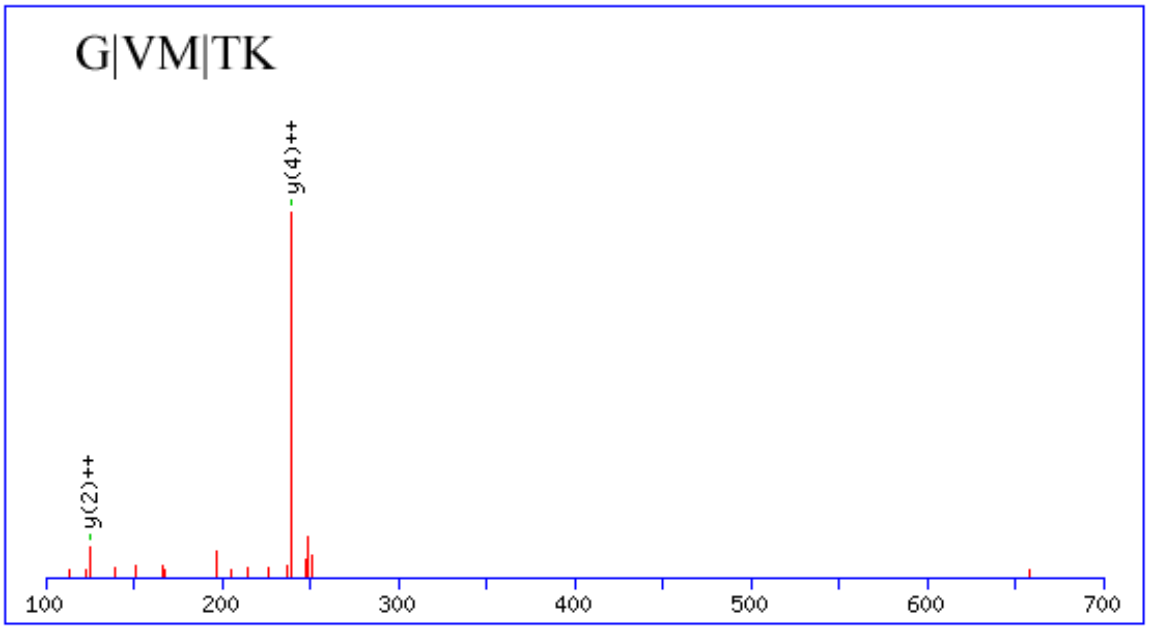
SPK**YGGVMLWDR**FYDKQSGYSDAIKGSIN

(B)

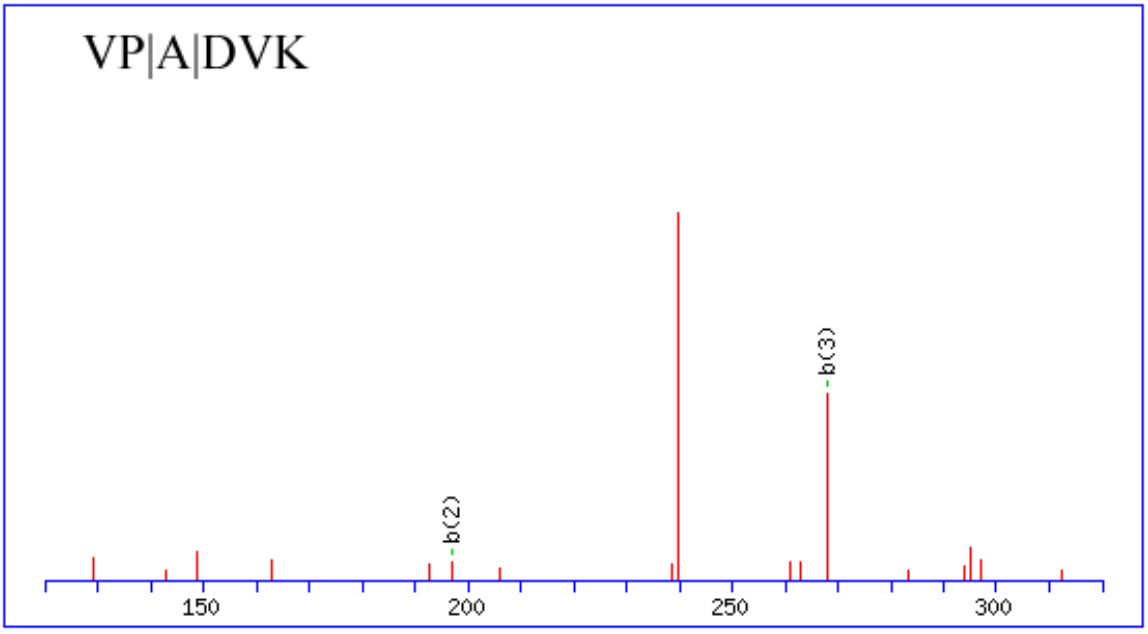
Figure A-19. Tandem mass spectra of polypeptide (A) from Chitinase (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).

Table A-VI. 1-D Teddy Bear Leaf Soil 25-35 kDa EDTA samples Cd, Cr, Ni, Pb

Marker Region kDa	Theroetical Mass	Accession no.	sp.	Protein	score	Peptide (Coverage)	Peptides
25 - 35	28	gi|2098793	*H. annuus*	HSP70-related protein	19	2 (4%)	GVMTK, VPADVK



(A)



(B)

Figure A-20. Tandem mass spectra of peptides (A,B) from HSP70-related protein (*H. annuus*) (C) and HSP70 (*N. benthamiana*) (D) as indicated (bold) in the primary amino acid sequence.

DLGVQGGILSGEGGDETKDILLLDVAPLTLGIETVG**GVMTK**LIPRNTVIPTKKSQ

VFTTYQDQQTVVSIKVYEGERSLTKDCRLLGTFDLTGIPPAPRGTPQIEVTFEVDA

NGILNVKAEDKASGKSEKITITNEKGRLSQEEIERMVREAEEFAEEDKKVKEKID

ARNALETYVYNMKNQINDKDKLADKLESDEKEKIEAATKEALDRLDDNQSAEK

EEYDEKLKEVEAVCNPIVTAVYQRSGGAPGGGAESTEDDDEHDEL

(C)

GVMTKIIPRNTTLPTSKSEVFSTAADGQTSVEINVLQGEREFVKDNKSLGRFRLD

GIPPAPRGVPQIEVKFDIDANGILSVTATDKGTGKKQDITITGASTLPKDEVDRMV

QEAEKFAKEDKEKRDSIDTKNQAESVVYQTEKQLKELGDK**VPADVK**DKVESKL

RELKDAISGGSTQTIKDTMAALNQEVMQLGQSLYSQPDAAAGPGPAGASGSRGK

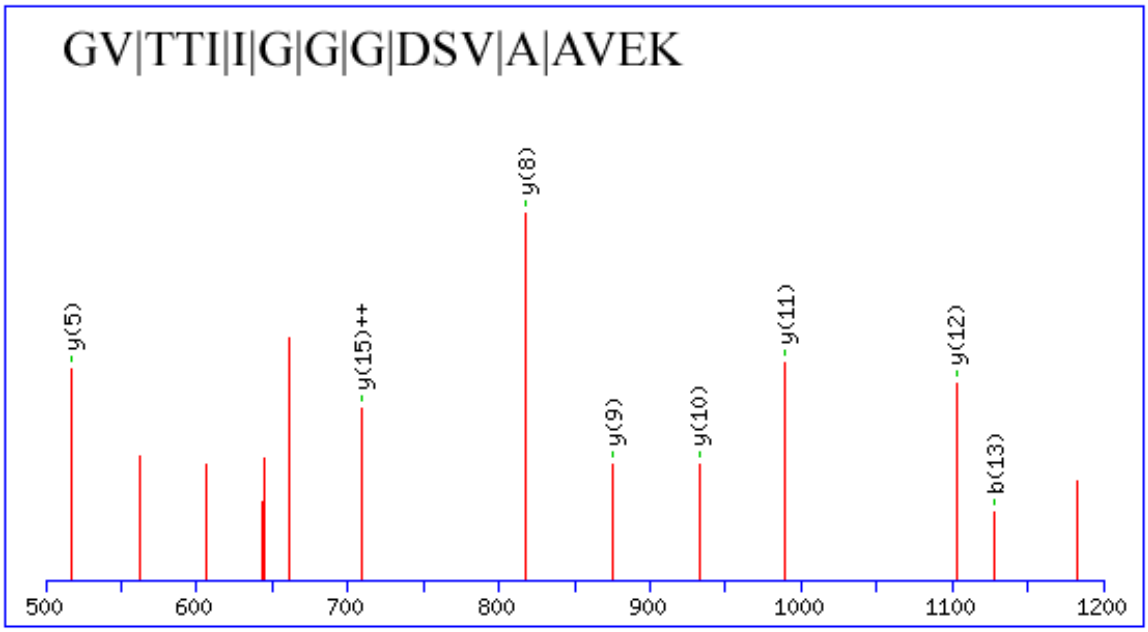
DDGDGEVIDADFNHEFWIRYVTRLQLRDR

(D)

Figure A-20. Tandem mass spectra of peptide (A,B) from HSP70-related protein (*H. annuus*) (C) and HSP70 (*N. benthamiana*) (D) as indicated (bold) in the primary amino acid sequence.

Table A-VII. 1-D Teddy Bear Leaf Soil Cr only

Marker Region kDa	Theroetical Mass	Accession no.	Sp.	Protein	Score	Peptide (Coverage)	Peptides
40	42	gi|121485004	*H. annuus*	cytosolic phosphoglycerate kinase	36	1 (5%)	GVTTIIGGGDSVAAVEK
200+	35	gi|94502504	*H. annuus*	Cytochrome f	60	2 (10%)	NILVIGPVPGQK, QVVDIIPPGPELLVSEGESIK



(A)

MATKKSVSSLTEGDLKGKRVFVRVDLNVPLDDTFKITDDTRIRAAVPTIKYLMS

HGARVILSSHLGRPKGVTPKFSLKPLVPRLSELLGIEVKMADDCIGPDVEKLVAEI

PEGGVLLLENVRFYKEEEKNDPEFAKKLASLADLYVNDAFGTAHRAHASTEGV

AKHLKPAVAGFLVQKELDYLVGAVSNPKKPFAAIVGGSKVSSKIGVIESLLEKVD

ILVLGGGMIFTFYKAQGLAVGSSLVEEDKLDLATTLLEKAKSKGVSLLLPSDVVI

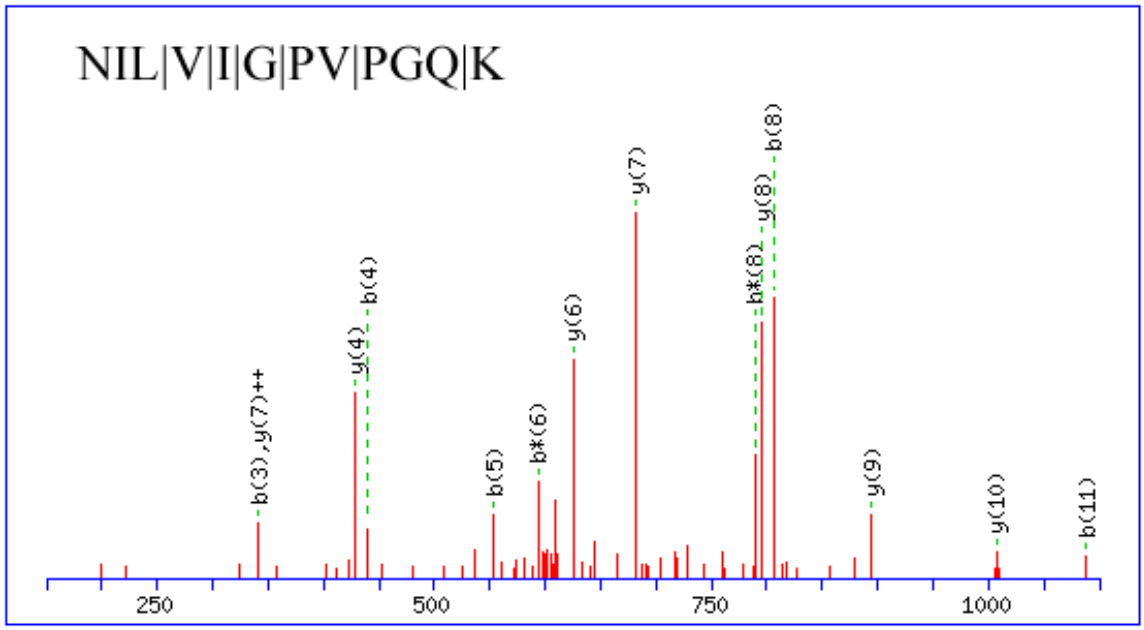
ADKFAADANSKIVPASSIPDGWMGLDIGPDSIKSFSEALDTTKTVIWNGPMGVFE

FDKFAVGTEAIAKKLAELSGK**GVTTIIGGGDSVAAVEK**VGLADKMSHISTGGG

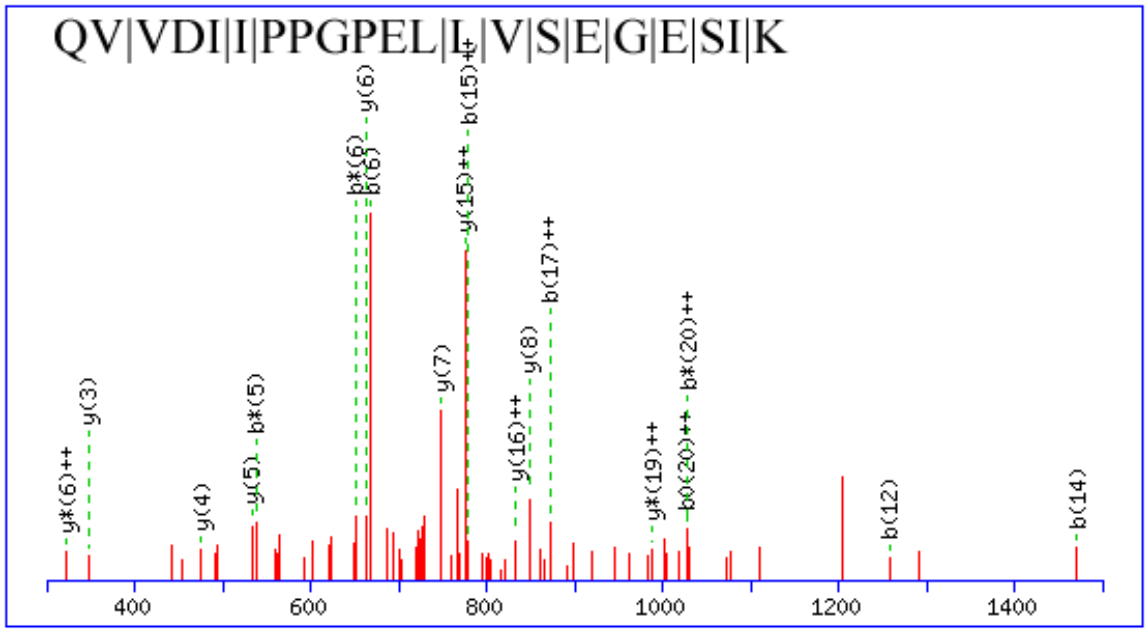
ASLELLEGKPLPGVPALDEA

(B)

Figure A-21. Tandem mass spectra of peptide (A) matched to cytosolic phosphoglycerate kinase (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).



(A)



(B)

Figure A-22. Tandem mass spectra of peptides (A,B) from cytochrome f (*H. annuus*) as indicated (bold) in the primary amino acid sequence (C).

MQTRNTFSWIKEQITRSISASLMIYIITRTSISNAYPIFAQQGYENPREATGRIVCAN

CHLANKPVDIEVPQAVLPDTVFEAVVRIPYDTQLKQVLANGKKGGLNVGAVLIL

PEGFELAPPDRISPEIKEKMGNLSFQSYRPNQK**NILVIGPVPGQK**YSEITFPILSPD

PATKKDIHFLKYPIYVGGNRGRGQIYPDGSKSNNTVYNATASGIVSKILRKEKGG

YEITIADASDGR**QVVDIIPPGPELLVSEGESIK**LDQPLTSNPNVGGFGQGDAEIVL

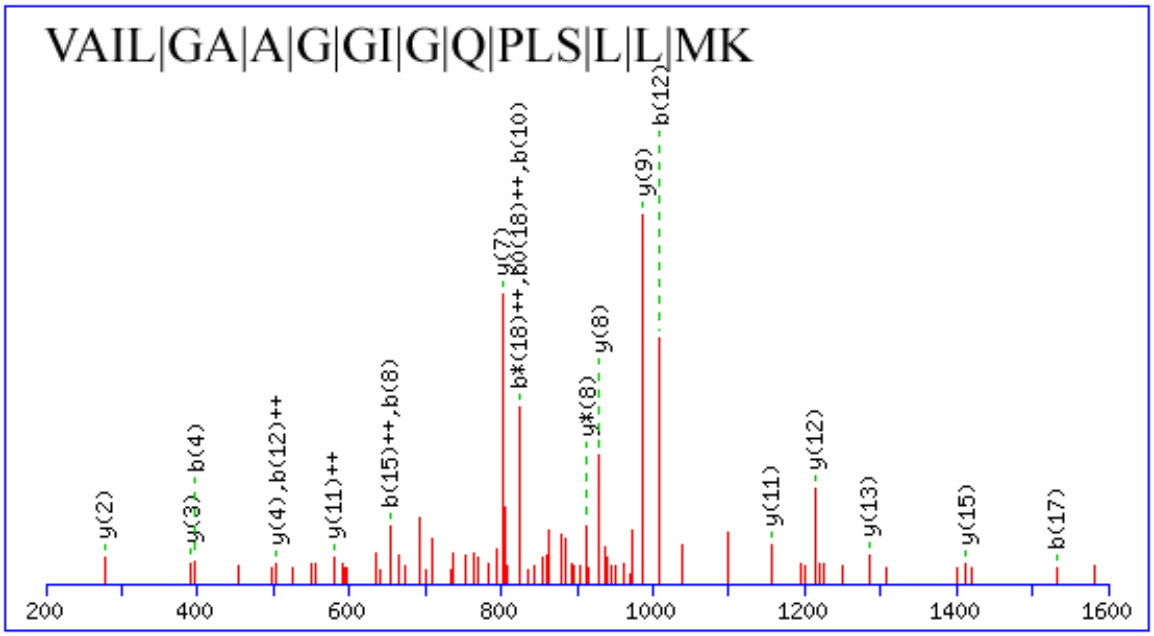
QDPLRVQGLLFFFAAVILAQIFLVLKKKQFEKVQLSEMNF

(C)

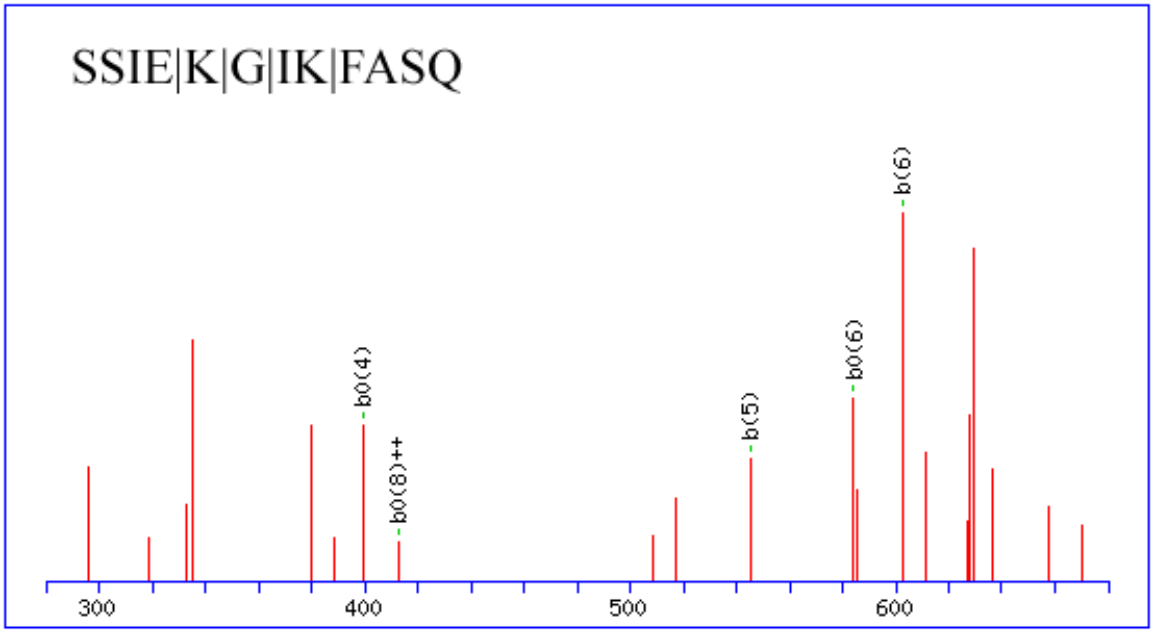
Figure A-22 (cont'd). Tandem mass spectra of peptides (A,B) from cytochrome f (*H. annuus*) as indicated (bold) in the primary amino acid sequence (C).

Table A-VIII. 1-D Teddy Bear Leaf Hydroponic Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	Accession no.	Sp.	Protein	Score	Peptide (Coverage)	Peptides
40	36	gi|2827080	*M. sativa*	malate dehydrogenase precursur chloroplast	113	2 (8%)	VAILGAAGGIGQPLSLLMK, SSIEKGIKFASQ
40	40	gi|6899972	*C. annuum*	ferredoxin-NADP+ oxidoreductase precursor	80	2 (6%)	EPYIGR, LYSIASSALGDFGDSK



(A)



(B)

Figure A-23. Tandem mass spectra of peptides (A,B) from malate dehydrogenase precursor (*M. sativa*) as indicated (bold) in the primary amino acid sequence (C).

MMRPSILRSVKSAVSRSITRRGYATEPVPERK**VAILGAAGGIGQPLSLLMK**LNP

LVSTLSLYDIAGTPGVAADVSHINSRSQVTGYAGEDELGKALEGADVVIIPAGVP

RKPGMTRDDLFNINAGIVKSLATAISKYCPHALVNMISNPVNSTVPIAAEVFKKA

GTYDEKRLFGVTTLDVVRAKTFYAGKANVPVAEVNVPVIGGHAGVTILPLFSQA

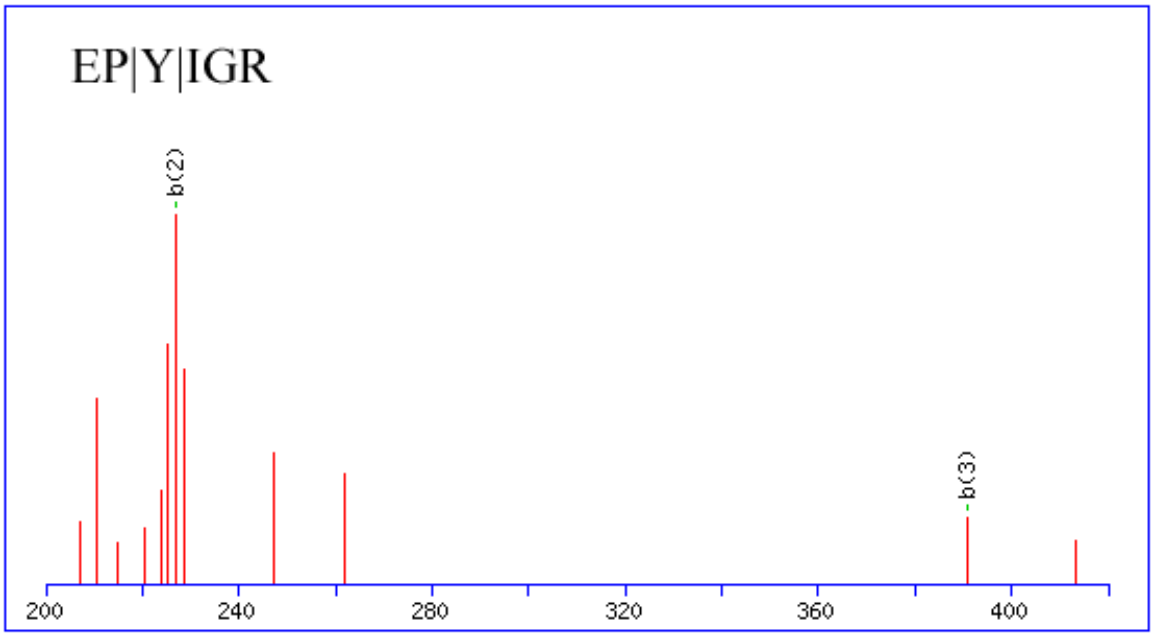
SPQANLDDDVIKALTARTQDGGTEVVTAKAGKGSATLSMAYAGAIFADACLKG

LNGVPDVVECSYVQSNLIAELPFFASKVRIGKNGVEEILGLGSLSDFEKQGLENLK

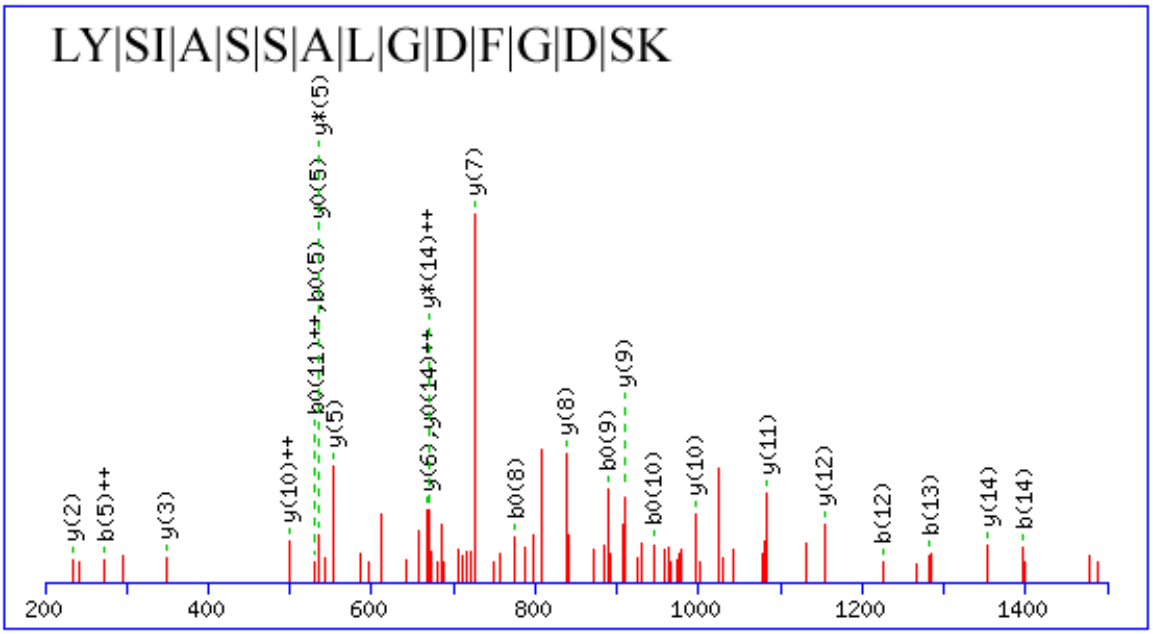
SELK**SSIEKGIKFASQ**

(C)

Figure A-23 (cont'd). Tandem mass spectra of peptides (A,B) from malate dehydrogenase precursor (*M. sativa*) as indicated (bold) in the primary amino acid sequence (C).



(A)



(B)

Figure A-24. Tandem mass spectra of peptides (A,B) from chloroplast ferredoxin-NADP+ oxidoreductase precursor (*C. annuum*) as indicated (bold) in the primary amino acid sequence (C).

MATAVTAAVSLPSSKSTSFPTRTSIISPEKINFNKVPLYYRNVSGGSKLVTIRAQVT

TEAPAKVEKISKKQDEGVVVNKFRPK**EPYIGR**CLLNTKITGDDAPGETWHMVFS

TEGEIPYREGQSIGVIADGVDANGKPHKLR**LYSIASSALGDFGDSK**TVSLCVKRL

VYTNDKGEEVKGVCSNFLCDLKPGADVKITGPVGKEMLMPKDPNATVIMLGTG

TGIAPFRSFLWKMFFEKHDDYKFNGLAWLFLGVPTSSSLLYKEEFEKMKEKAPE

NFRLDFAVSREQTNEKGEKMYIQTRMAQYAEELWTLLKKDNTFVYMCGLKGM

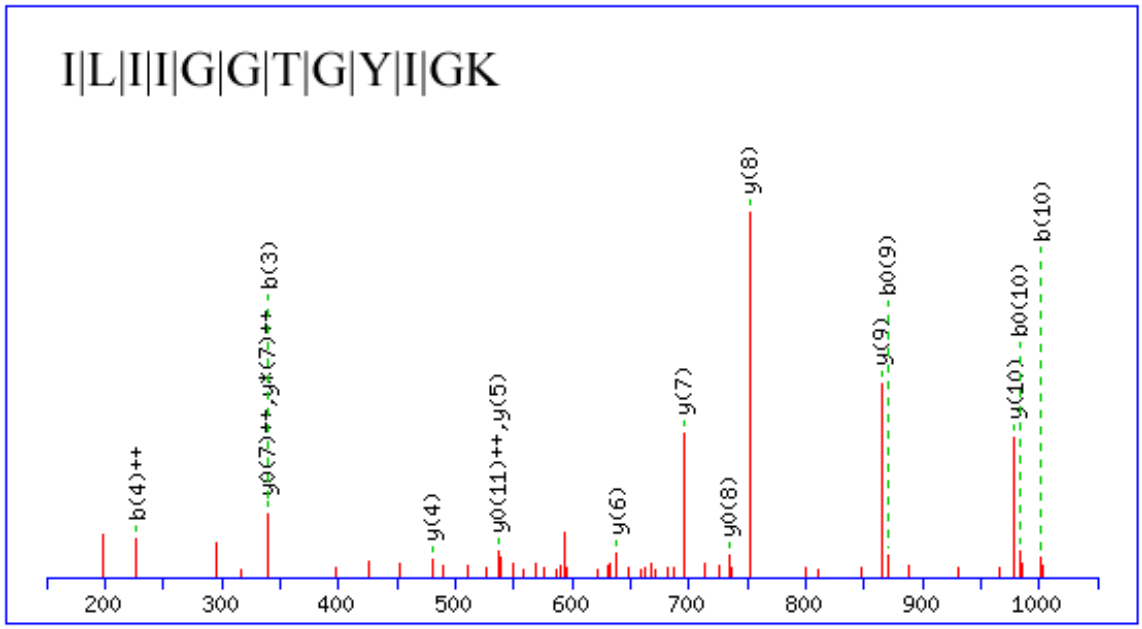
EQGIDDIMSSLAAKEGIDWADYKKQLKKAEQWNVEVY

(C)

Figure A-24 (cont'd). Tandem mass spectra of peptides (A,B) from chloroplast ferredoxin-NADP+ oxidoreductase precursor (*C. annuum*) as indicated (bold) in the primary amino acid sequence (C).

Table A-IX. 1-D Teddy Bear Stem Hydroponic Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	Accession no.	Sp.	Protein	Score	Peptide (Coverage)	Peptides
40	34	gi|507823	*Arabidopsis*	NADPH oxidoreductase	65	1 (3%)	ILIIGGTGYLGK
40	36	gi|2827080	*M. sativa*	malate dehydrogenase putative	48	3 (15%)	VAILGAAGGIGQPLSLLMK, ALEGADVVIIPAGVPR, DDLFNINAGIVK



(A)

MVVSEKSK**ILIIGGTGYIGK**YLVETSAKSGHPTFALIRESTLKNPEKSKLIDTFKS

YGVTLLFGDISNQESLLKAIKQVDVVISTVGGQQFTDQVNIIKAIKEAGNIKRFLPS

EFGFDVDHARAIEPAASLFALKVRIRRMIEAEGIPYTYVICNWFADFFLPNLGQLE

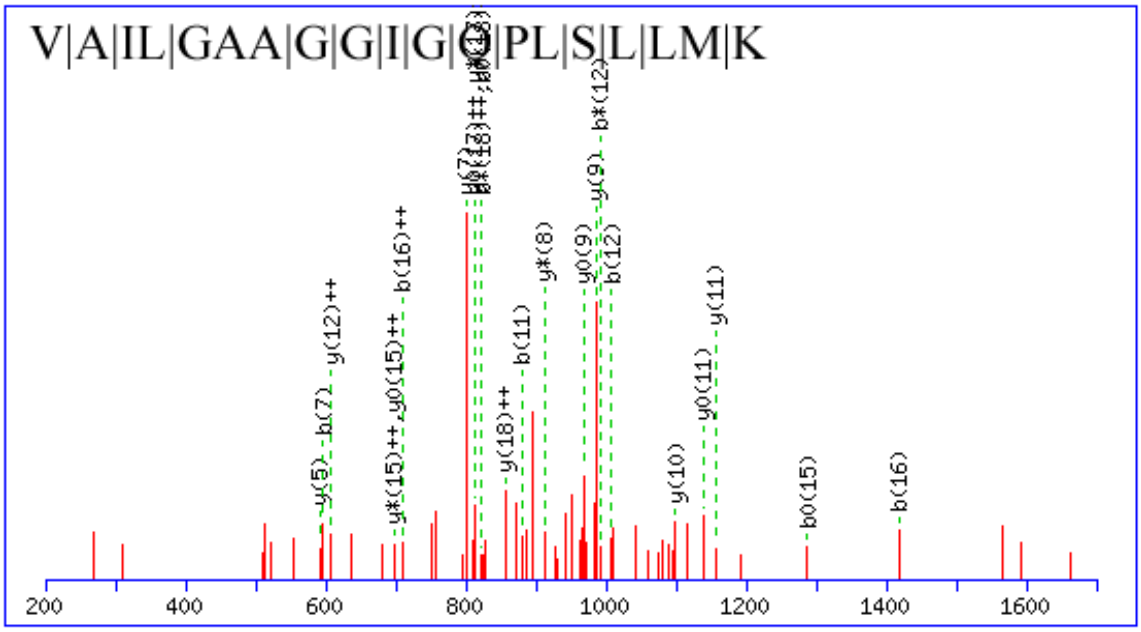
AKTPPRDKVVIFGDGNPKAIYVKEEDIATYTIEAVDDPRTLNKTLHMRPPANILSF

NEIVSLWEDKIGKTLEKLYLSEEDILQIVQEGPLPLRTNLAICHSVFVNGDSANFE

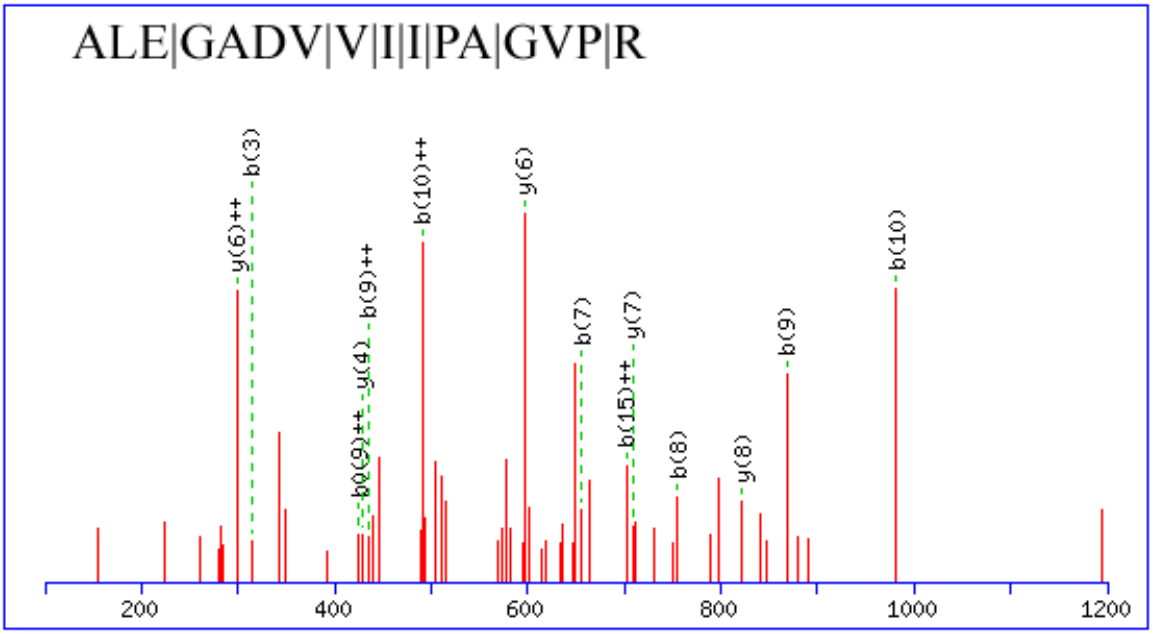
VQPPTGVEATELYPKVKYTTVDEFYNKFV

(B)

Figure A-25. Tandem mass spectra of peptides (A) matched to NADPH oxidoreductase, putative (*Arabidopsis)* as indicated (bold) in the primary amino acid sequence (B).

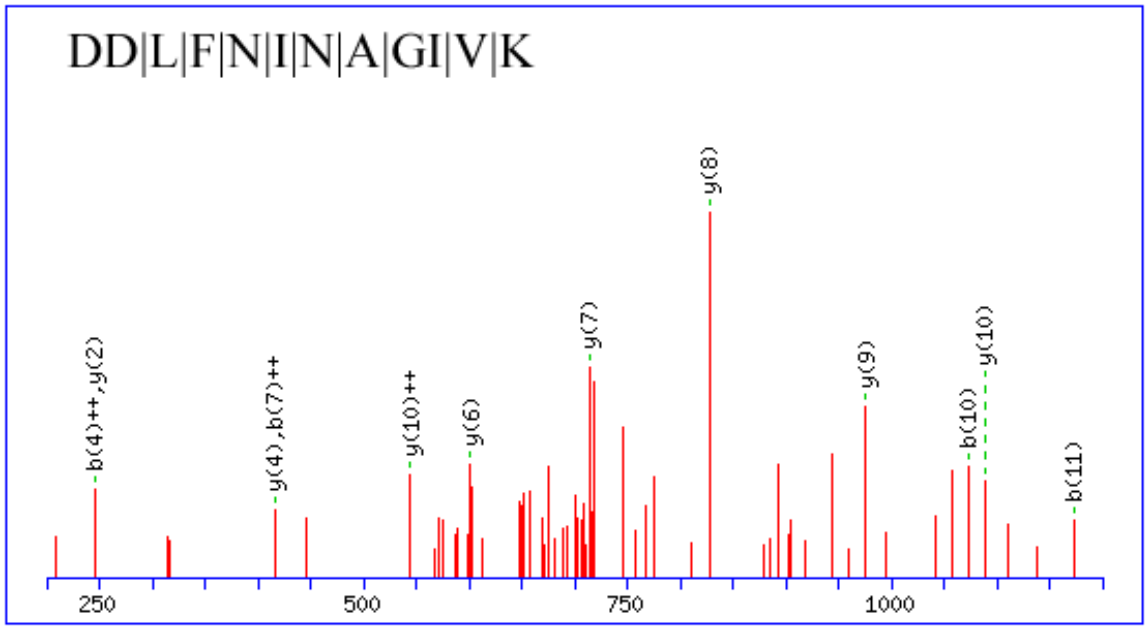


(A)



(B)

Figure A-26. Tandem mass spectra of peptides (A,B,C) matched to malate dehydrogenase precurser (*M. sativa*) as indicated (bold) in the primary amino acid sequence (D).



(C)

MMRPSILRSVKSAVSRSITRRGYATEPVPERK**VAILGAAGGIGQPLSLLMK**LNP

LVSTLSLYDIAGTPGVAADVSHINSRSQVTGYAGEDELGK**ALEGADVVIIPAGVP**

RKPGMTR**DDLFNINAGIVK**SLATAISKYCPHALVNMISNPVNSTVPIAAEVFKKA

GTYDEKRLFGVTTLDVVRAKTFYAGKANVPVAEVNVPVIGGHAGVTILPLFSQA

SPQANLDDDVIKALTARTQDGGTEVVTAKAGKGSATLSMAYAGAIFADACLKG

LNGVPDVVECSYVQSNLIAELPFFASKVRIGKNGVEEILGLGSLSDFEKQGLENLK

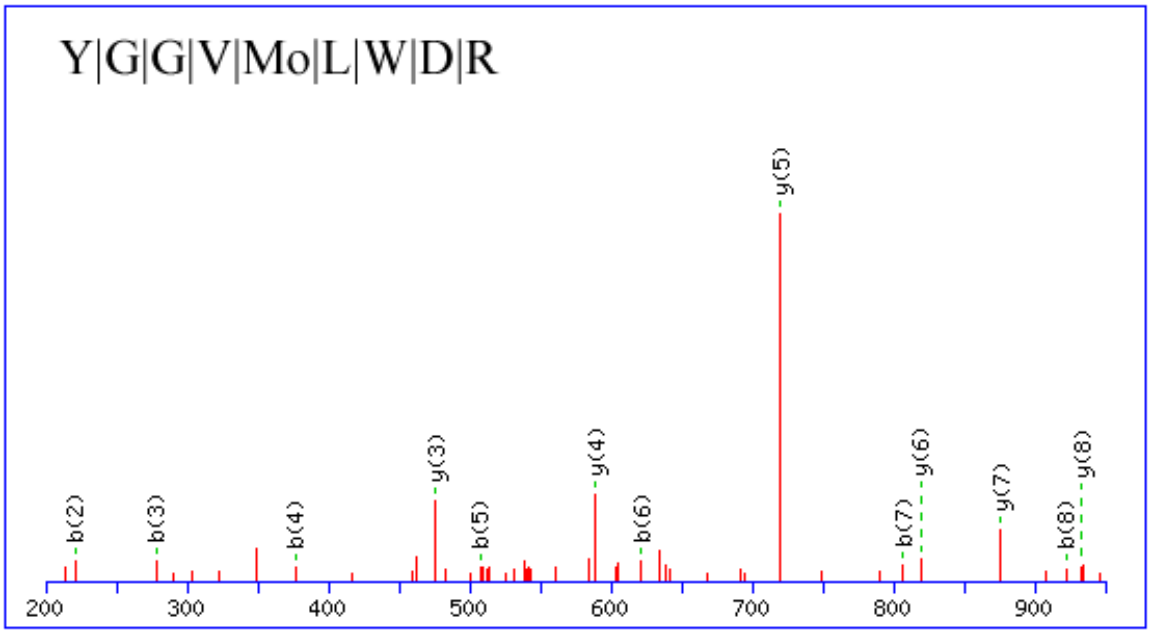
SELKSSIEKGIKFASQ

(D)

Figure A-26 (cont'd). Tandem mass spectra of peptides (A,B,C) matched to malate dehydrogenase precurser (*M. sativa*) as indicated (bold) in the primary amino acid sequence (D).

Table A-X. 1-D Teddy Bear Leaf Hydroponic (Soluble fraction) Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	Accession no.	Sp.	Protein	Score	Peptide (Coverage)	Peptides
25-35	32	gi|126723930	*H. annuus*	Chitinase	49	1 (3%)	YGGVMLWDR
115-125	106	gi|3415117	*Arabidopsis*	villin 3	28	1(2%)	AAALAALTSAFNSSSGR

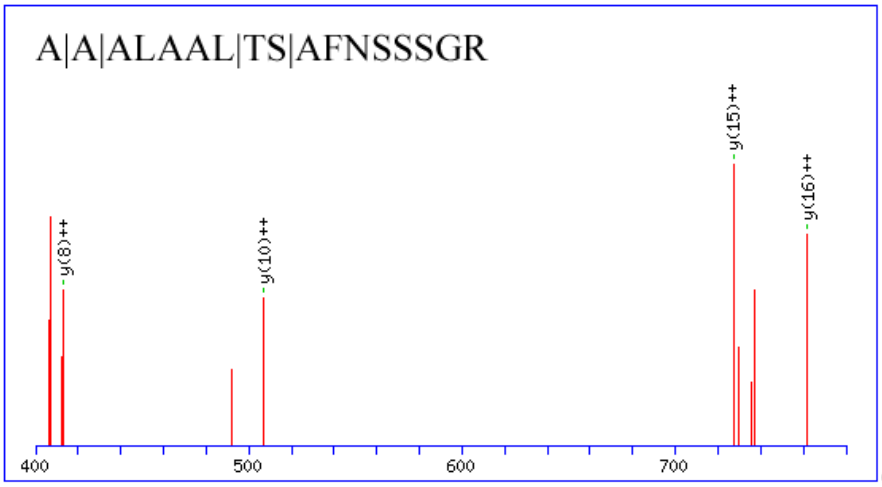


(A)

MEFTHPALLLLLFITVFSFLKPSTAAGIATYWGQQSDDTEGTLEAACATGNYQFV
NIAFLSTFGNNQQPVLNLAAHCDPASTCSRYSSQIKACQAQNVKVFLSIGGQRGS
YSLSSPQDAQQVADFLWNTYLGGQPATRPLGDAVLDGIDFDIEQGTDQFWSDLA
KAPAAYSSQKKVYLSAAPQCPFPSGDVRNQLLPAIREGLFDYVWVQFYNNEQCQ
YGANADALLARWNEWTQVTTNTIFLGLPAAASGAAPSGGYIPPDILTSQILPSIKS
SPK**YGGVMLWDR**FYDKQSGYSDAIKGSIN

(B)

Figure A-27. Tandem mass spectra of polypeptide (A) from Chitinase (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).



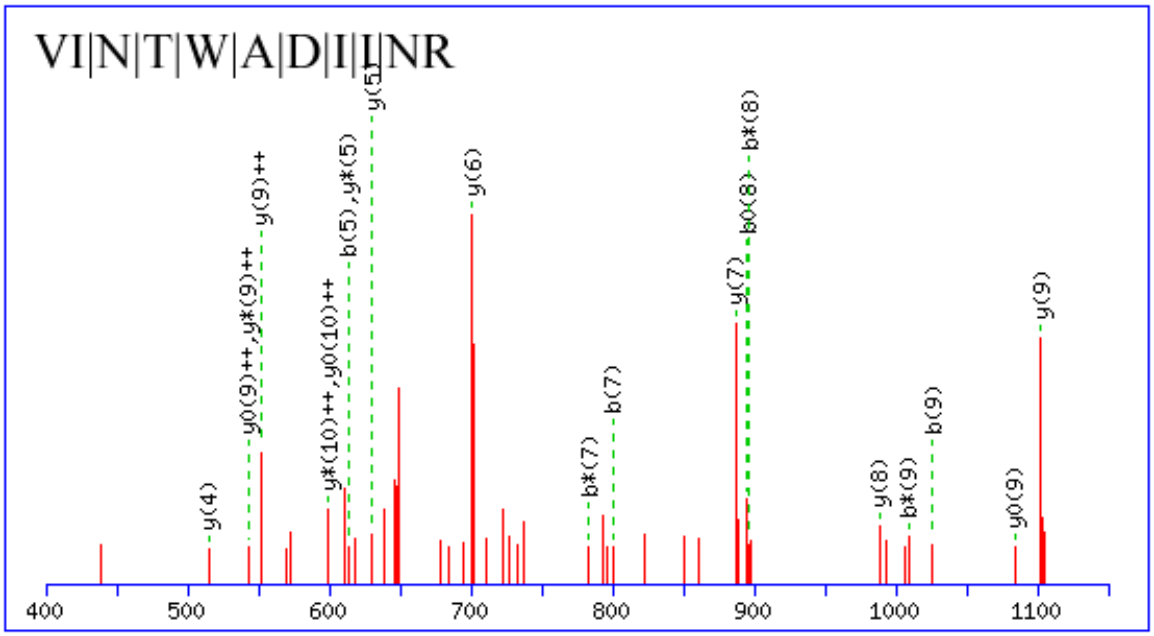
(A)

MSGSTKVLDPAFQGVGQKPGTEIWRIENFEPVPVPKSEHGKFYMGDTYIVLQTTQNKGGAYLFDI
HFWIGKDTSQDEAGTAAVKTVELDAALGGRAVQYREIQGHESDKFLSYFKPCIIPLEGGVASGFKK
PEEEEFETRLYTCKGKRAVHLKQVPFARSSLNHDDVFILDTKEKIYQFNGANSNIQERAKALVVIQ
YLKDKFHEGTSDVAIVDDGKLDTESDSGEFWVLFGGFAPIARKVASEDEIIPETTPPKLYSIADGQV
ESIDGDLSKSMLENNKCYLLDCGSEIFIWVGRVTQVEERKTAIQAAEDFVASENRPKATRITRVIQG
YEPHSFKSNFDSWPSGSATPANEEGRGKVAALLKQQGVGLKGLSKSTPVNEDIPPLLEGGGKLEV
WYIDANSKTVLSKDHVGKLYSGDCYLVLYTYHSGERKEDYFLCCWFGKNSNQEDQETAVRLAST
MTNSLKGRPVQARIFEGKEPPQFVALFQHMVVLKGGLSSGYKNSMTEKGSSGETYTPESIALIQVS
GTGVHNNKALQVEAVATSLNSYDCFLLQSGTSMFLWVGNHSTHEQQELAAKVAEFLKPGTTIKH
AKEGTESSSFWFALGGKQNFTSKKVSSETVRDPHLFSFSFNRGKFQVEEIHNFDQDDLLTEEMHLL
DTHAEVFVWVGQCVDPKEKQTAFEIGQRYINLAGSLEGLSPKVPLYKITEGNEPCFFTTYFSWDST
KATVQGNSYQKKAALLLGTHHVVEDQSSSGNQGPRQR**AAALAALTSAFNSSSGR**TSSPSRDRSN
GSQGGPRQRAEALAALTSAFNSSPSSKSPPRRSGLTSQASQRAAAVAALSQVLTAEKKKSPDTSPS
AEAKDEETSFSEVEATEEATEAKEEEEVSPAAEASAEEAKPKQDDSEVETTGVTFTYERLQAKSEK
PVTGIDFKRREAYLSEVEFKTVFGMEKESFYKLPGWKQDLLKKKFNLF **(B)**

Figure A-28. Tandem mass spectra of peptide (A) matched to villin 3 (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (B).

Table A-XI. 1-D Teddy Bear Leaf Hydroponic (Transmembrane fraction) Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	Accession no.	Sp.	Protein	Score	Peptide (Coverage)	Peptides
35	38	gi|94502471	*H. annuus*	photosystem II protein D1	28	1 (3%)	VINTWADIINR,



(A)

MTAILERRESESLWGRFCNWITSTENRLYIGWFGVLMIPTLLTATSVFIIAFIAAPP

VDIDGIREPVSGSLLYGNNIISGAIIPTSAAIGLHFYPIWEAASVDEWLYNGGPYELI

VLHFLLGVACYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGQGSFS

DGMPLGISGTFNFMIVFQAEHNILMHPFHMLGVAGVFGGSLFSAMHGSLVTSSLI

RETTENESANEGYRFGQEEETYNIVAAHGYFGRLIFQYASFNNSRSLHFFLAAWP

VVGIWFTALGISTMAFNLNGFNFNQSVVDSQGR**VINTWADIINR**ANLGMEVMH

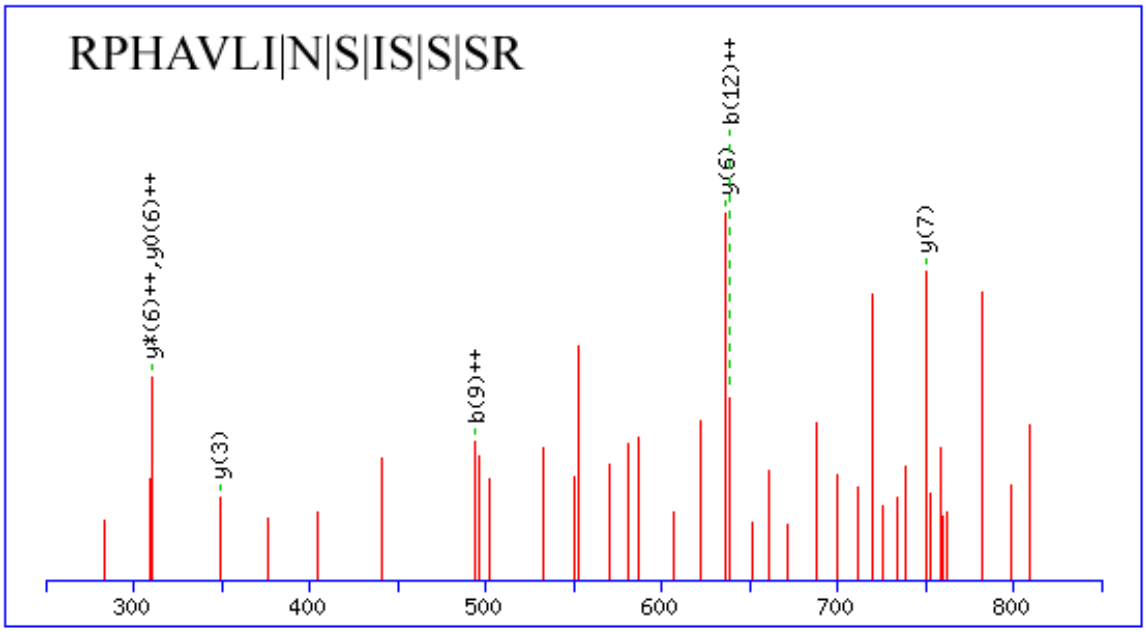
ERNAHNFPLDLAAIEAPSTNG

(B)

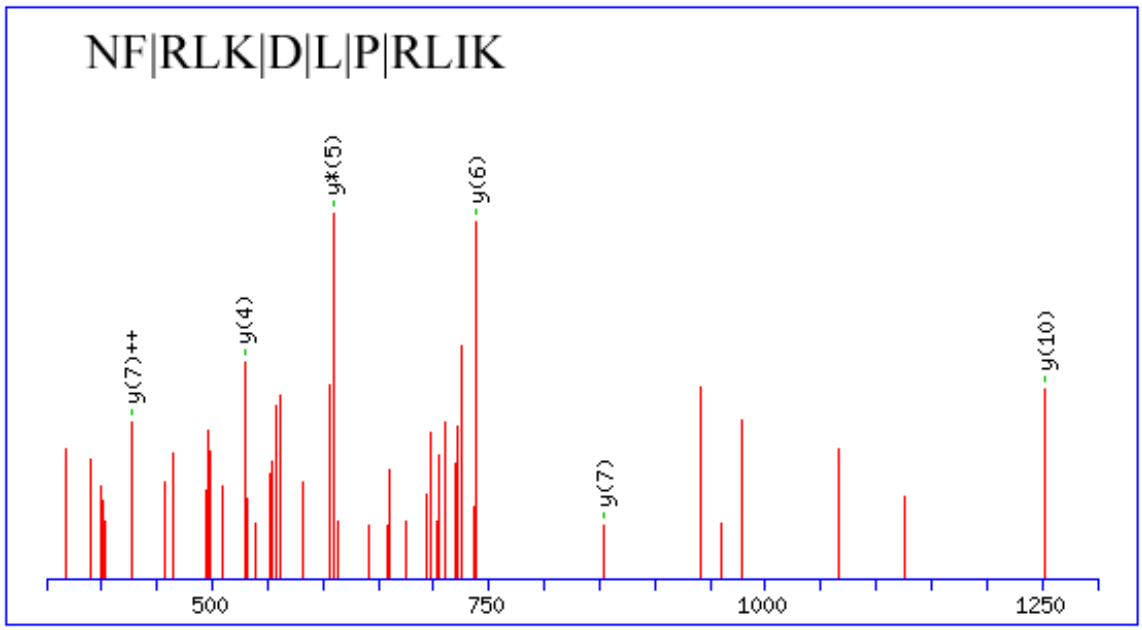
Figure A-29. Tandem mass spectra of peptide (A) matched to photosystem II protein D1 (*H. annuus*) as indicated (bold) in the primary amino acid sequence (B).

Table A-XII. 2-D Teddy Bear Leaf Hydroponic Cd, Cr, Ni, Pb

Marker Region kDa	Theoretical Mass kDa	Accession no.	sp	Protein	Score	Peptide (Coverage)	Peptides
25-35	31	gi|92884215	*M. truncatula*	Cytokinin-O-glucosyltransferase	15	3 (8%)	RPHAVLINSISSSR, NFRLKDLPRLIK, NPNDLTIRFNTEVADKCLR
25-30	24.5	gi|38605916	*O. sativa*	ABA/WDS stress induced	17	2 (4%)	KHHLFG, AQKPVHDASKR
25-30	34	gi|51963656	*O. sativa*	NAC Domain protein	18	2 (4%)	GRGLLVGMK, GEKTPWMHEYR
45	42	gi|15221544	*Arabidopsis*	similar to 1-aminocyclopropane-1-carboxylic acid oxidase	16	3 (4%)	DFQDAIK, VYGPIK, GLGGTSALLDFRV
45	41	gi|20372648	*Rosa hybrid cultivar*	1-aminocyclopropane-1-carboxylic acid oxidase	27	1 (2%)	AIANFMSKAR

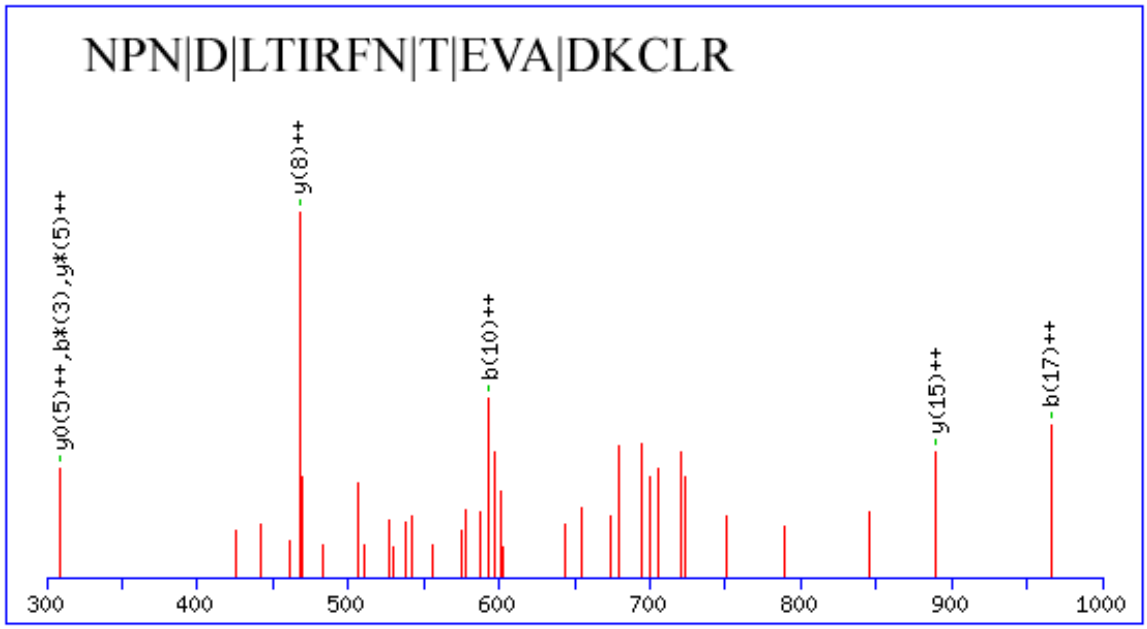


(A)



(B)

Figure A-30. Tandem mass spectra of peptides (A,B,C) matched to Cytokinin-O-glucosyltransferase (*M. truncatula*) as indicated (bold) in the primary amino acid sequence (D).



(C)

MTNFEEK**RPHAVLINSISSSR**GENAFDGFTDFNFETIPDGLTPKDGDGDISQDLHS

LGESIITNFHHFFDELLAKLQDSATAGLIPPVTCLVSDCYMPFTVAAAEEHALPVV

LFSPSRACYFLACLLSPKMYQNSQLPFKDESDLTNEYLDTKVDWIPGLK**NFRLK**

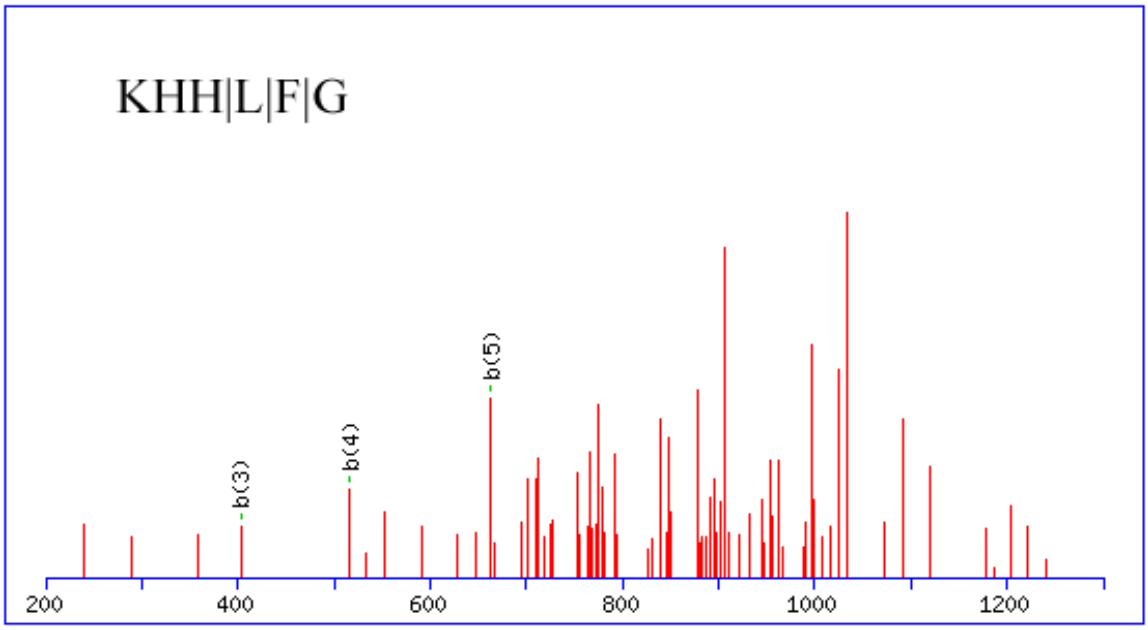
DLPRLIKTK**NPNDLTIRFNTEVADKCLR**ASGMVFNTSNELESDVMNAFYSMFP

SLYTIGPLASFVNQSPQNHLTSLDCNLWKEDTKCLEWLESKEPGSVVYVNFGSIT

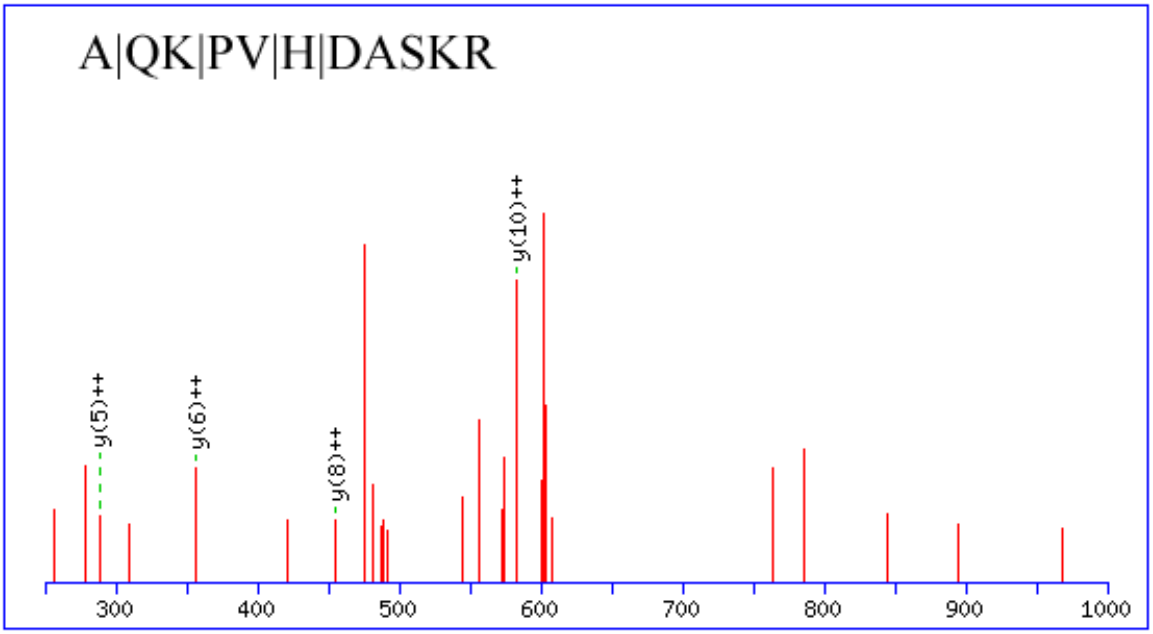
VKSQEKP

(D)

Figure A-30 (cont'd). Tandem mass spectra of peptides (A,B,C) matched to Cytokinin-O-glucosyltransferase (*M. truncatula*) as indicated (bold) in the primary amino acid sequence (D).



(A)



(B)

Figure A-31. Tandem mass spectra of peptides (A,B) matched to ABA/WDS stress induced (*O. sativa*) as indicated (bold) in the primary amino acid sequence (C).

MADEYGRGGYGRSGAGAGDDYESGGYNRSSSGGADEYAAGRSGR**AQKPVXD**

ASKRFTKSRRRATTYGXRRRRVNKSGPRASDSGXNNRSGANRSTAATSPARRV

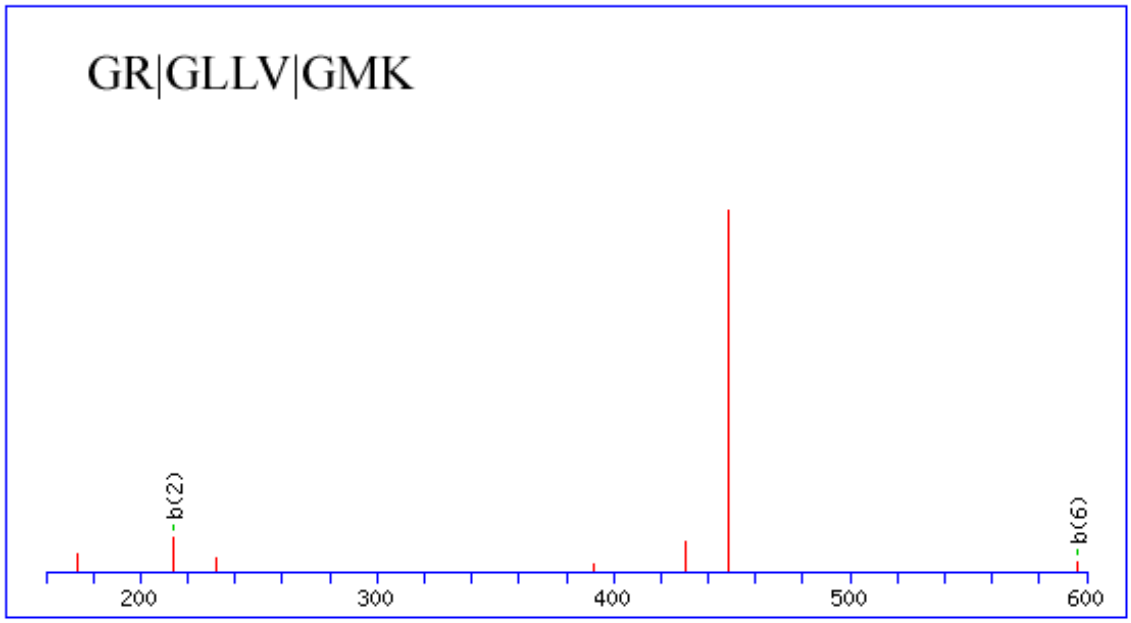
QRRXGAEADEEYVDGLSSRAPGEVQEGGEGAQEQGAPREVGPXRRXFAMYERH

QAKKDPENAQRHRIEEGVAAAAALGSGGFAFHEHHDKKEAKQAAKDAEEEAEE

ESGSGARGGEGKK**KHHLFG**

(C)

Figure A-31(cont'd). Tandem mass spectra of peptides (A,B) matched to ABA/WDS stress induced (*O. sativa*) as indicated (bold) in the primary amino acid sequence(C).



(A)

MSESEVSVINQLEEEETRLELPPGFRFHPTDEEVVTHYLTRKAQDRSFSCVVIADV

NLNNCEPWDLPSKAKMGEKEWFFFCHKDRKYPTGMRTNRATASGYWKATGKD

KEIFR**GRGLLVGMK**KTLVFYMGRAPR**GEKTPWVMHEYR**LDGKLPPNLPRSA

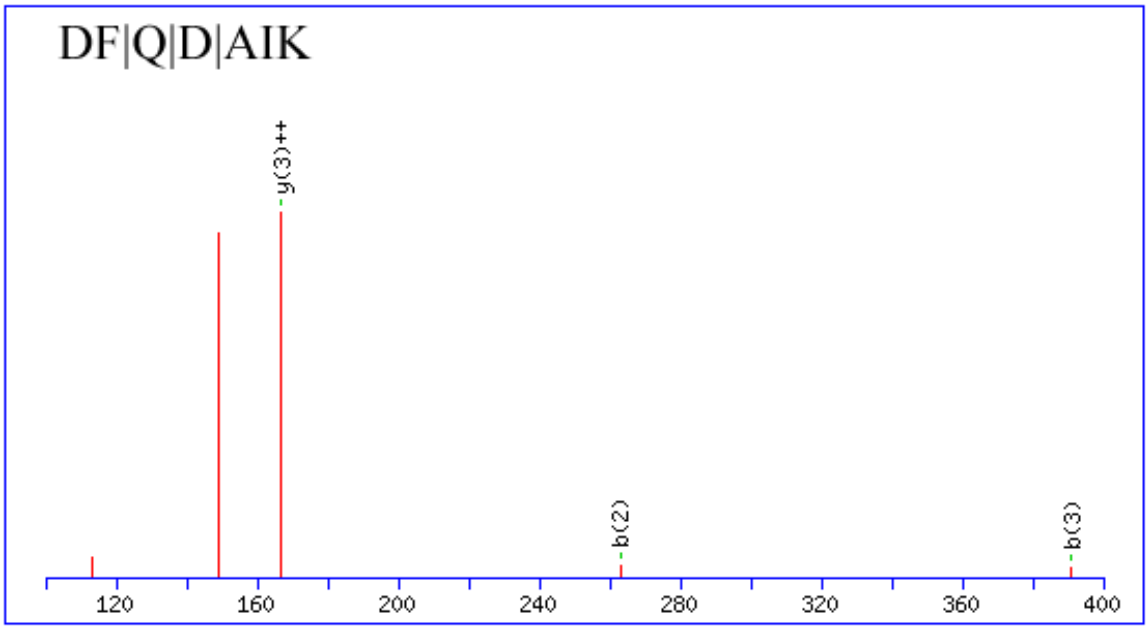
KEEWAVCRVFNKDLAAKIAQMPPPPFPRNDSFDLDLDDFLHLDADLPPLIDDPFA

STSTLKTEPPPPANLMHNHYGYFSLPASATNYNHSSGAMADQAIRRFCKAEAST

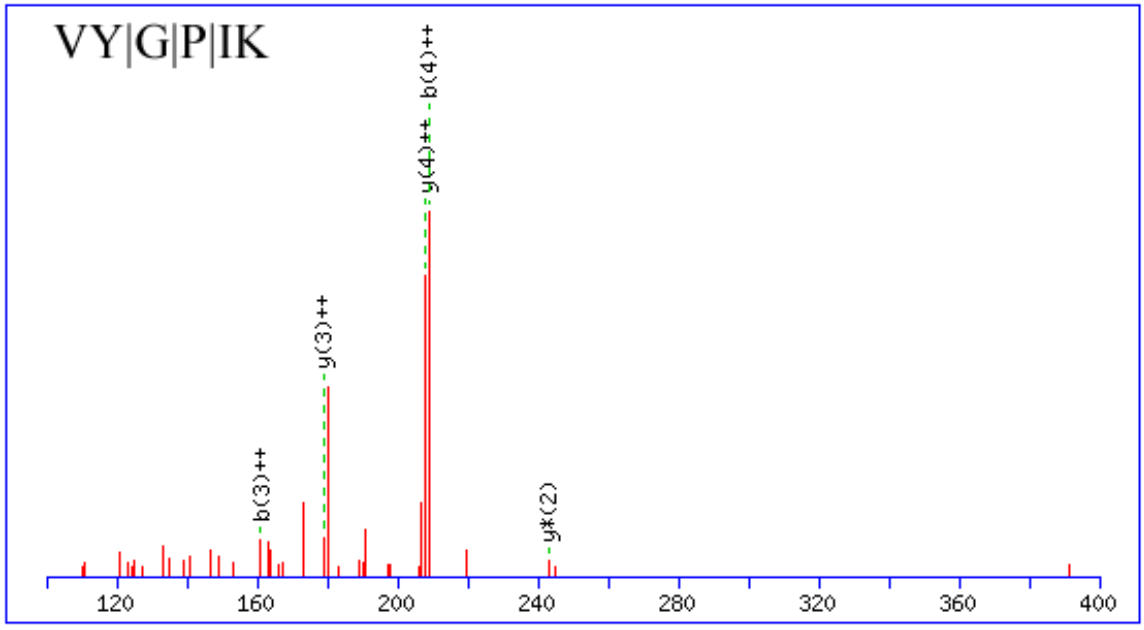
ACFSGADADVDPVVDELLSFPDSITDYSYIWKA

(B)

Figure A-32. Tandem mass spectra of peptide (A) matched to NAC Domain protein (*O. sativa*) as indicated (bold) in the primary amino acid sequence (B).

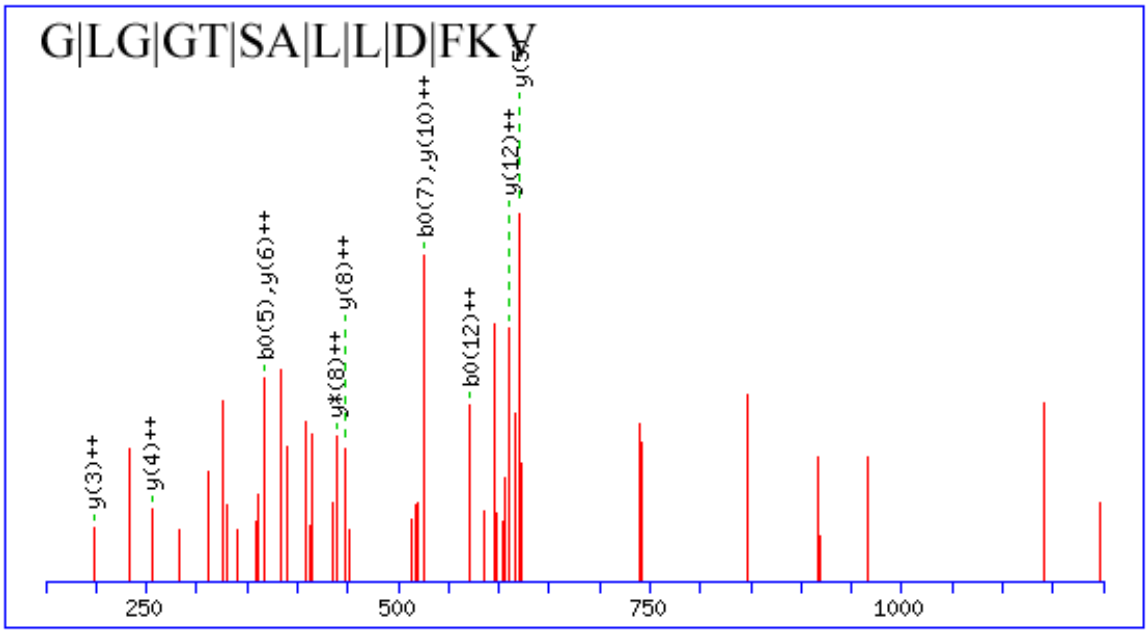


(A)



(B)

Figure A-33. Tandem mass spectra of peptides (A,B,C) matched to similar to 1-aminocyclopropane-1-carboxylic acid oxidase (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (D).



(C)

MESTKIAPSFDRASELKAFDETKTGVKGLVDSGISKIPRIFHHSSVELANPKPLPSD

LLHLKTIPTIDLGGR**DFQDAIK**HKNAIEGIKEAAAKWGFFQVINHGVSLELLEKM

KDGVRDFHEQPPEVRKDLYSRDFGRKFIYLSNFDLYTAAAANWRDTFYCYMAP

DPPEPQDLPEICRDVMMEYSKQVMILGEFLFELLSEALGLNPNHLKDMECLKGL

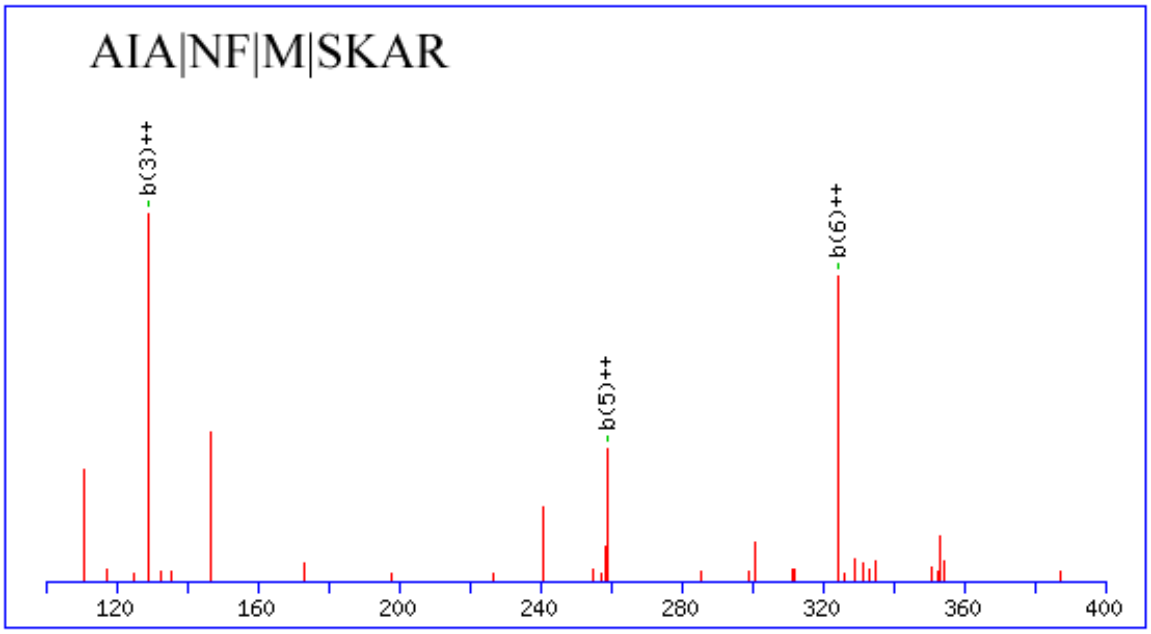
RMLCHYFPPCPEPDLTFGTSKHSDGSFLTVLLPDNIEGLQVCREGYWFDVPHVPG

ALIINIGDLLQLITNDKFISLKHRVLANRATRARVSVACFFHTHVKPNPR**VYGPIK**

ELVSEENPPKYRETTIRDYATYFNGK**GLGGTSALLDFKV**

(D)

Figure A-33 (cont'd). Tandem mass spectra of peptides (A,B,C) matched to similar to 1-aminocyclopropane-1-carboxylic acid oxidase (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (D).



(A)

MGLAENQLSFDVIEEWIKKNPKASICTPEGVAEFKNVANFQDYHGFPDFRK**AIAN**

FMSKARGGRVKFDPDRVVMAGGATGANETVMFCLADPGDAFLVPSPYYPAFY

RDLGWRTGVKIVPVDCDSSNNFQVTRAALEAAYEKARNNNINIKGLIITNPSNPL

GTTVDRDTLTSLVKFINDKNIHLVCDEIYAGTVFSCPKFTCVTEVLQDVKNCNRD

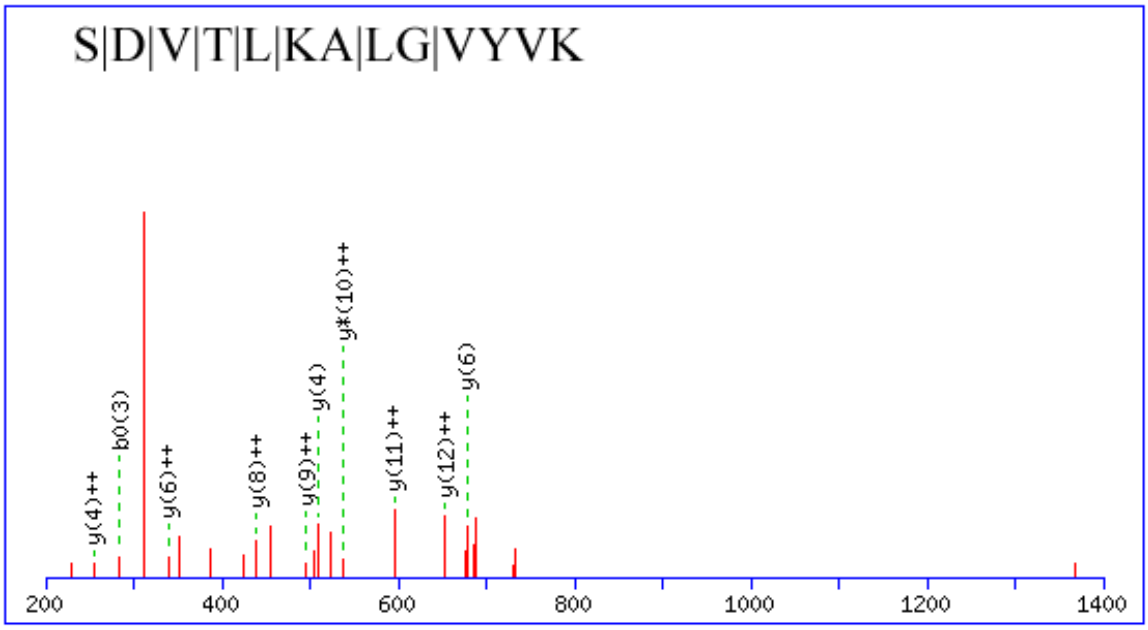
LIHIVYSLSKDMGVPGLRIGIVYSYNDAVVNCARKMSSFGLVSSQTQHMLSAML

LDDEFVSNFLETSSKRLARRHRFFTTGLEEVGINCLKGNAGLYCWMDLRKLLKD

QTFEAEMVLWRMIINEVKLNVSPGSSFRCVEPGWFRVC

(B)

Figure A-34. Tandem mass spectra of peptide (A) from 1-aminocyclopropane-1-carboxylic acid oxidase (*Rosa hybrid cultivar*) as indicated (bold) in the primary amino acid sequence (B).



(A)

MAVVMKVGAWGAPDGSPQDINAESRPQRLESITIYSVESPGVCGIKGFSFKYVD

QQGSSVKSAIWGSNSGNPNTIEMREGEQLKLVGGTFDNEGIGSLTLETNTTKHKT

YGYPVQGGEFSLPLPQGKGELVAFFGR**SDVTLKALGVYVK**GSPAKVGKWGAK

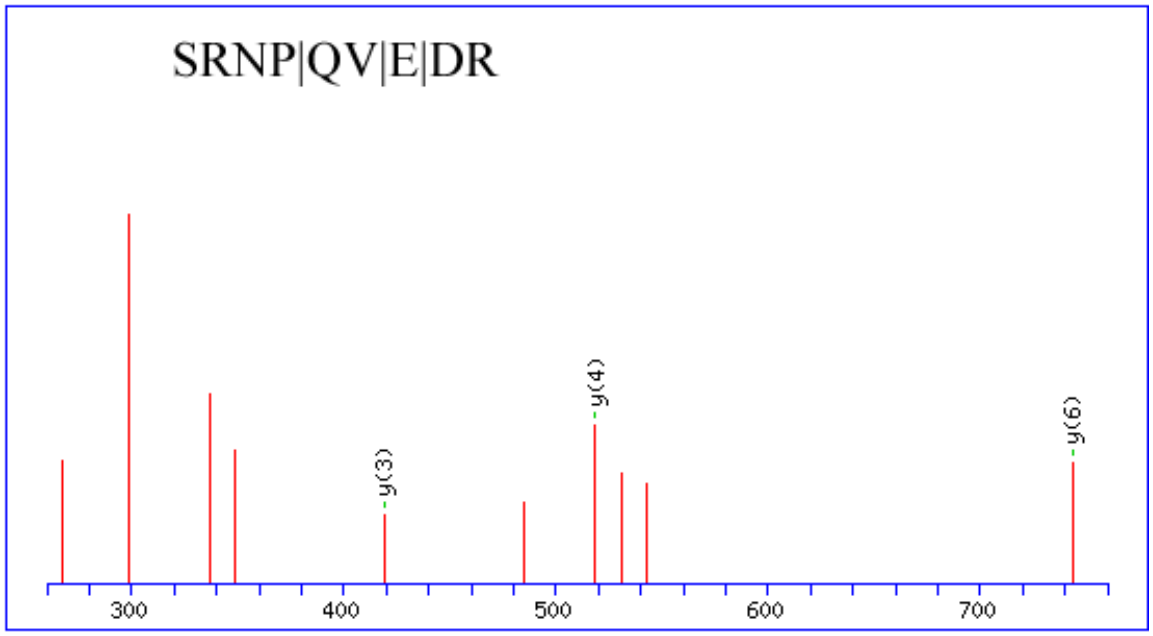
SGAPRDIRPDADPCKLESFTIHSSERIHGFSFTYLTKSGQAISVPLWGKKAGEEHTI

FMNQSEYVSSITGAYDTYGITFLNFDTNQGASHTFGRNPSAGTKTFSVPLPDNGA

LDDAAAVAFFGSSGDSLVAIGAYVGVAPE

(B)

Figure A-35. Tandem mass spectra of a polypeptide (A) matched to Jacalin-like protein LEM2 (*H. vulgare*) as indicated (bold) in the primary amino acid sequence (B).



(A)

MKYLWVFIVFSIAVLSHACSAQQCGRQAGNRRCANNLCCSQYGYCGRTNEYCC

TSQGCQSQCRRCGVRTVGEIVVGDIGGIISKGMFNNILKHRDDDACEGKGFYTYE

AFVAAARSFPAFGSTGDDATRKREIAAFLAQTSHETSAGWPSAPDGPYAWGYCF

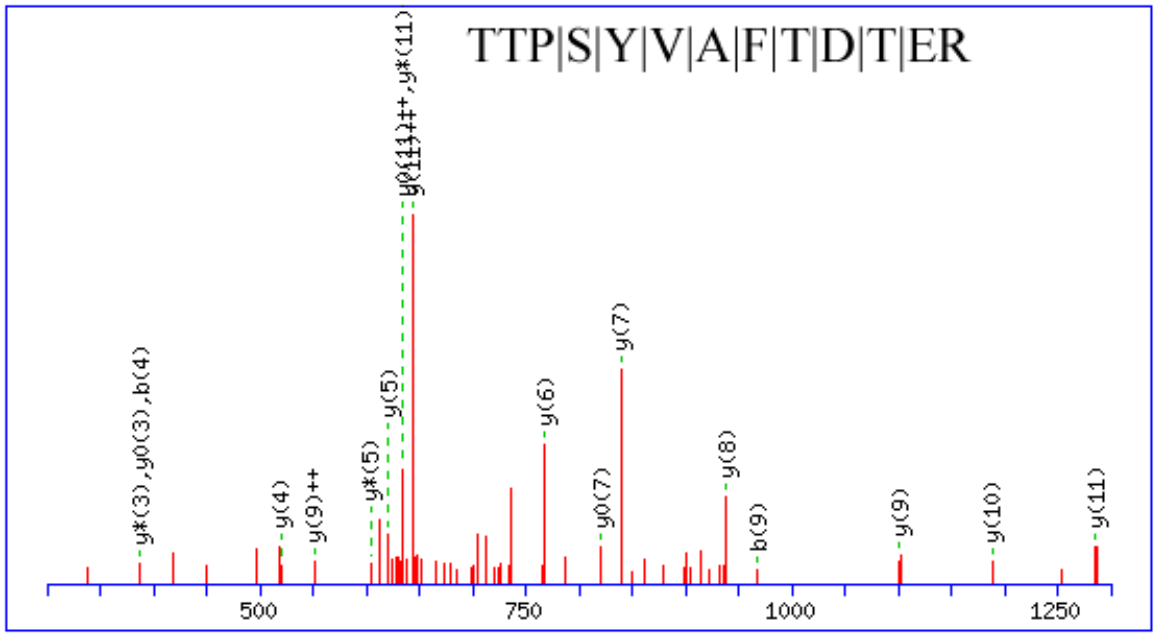
VRERNPPSKYCDTTTPCPKSYYGRGPIQLTWNYNYEQAGRAIGADLLNNPDLVA

TDAVISFKTAIWFWMTAQSSKPSCHDVITGSWRPSASDNSVCHVPDYAVVTNIIS

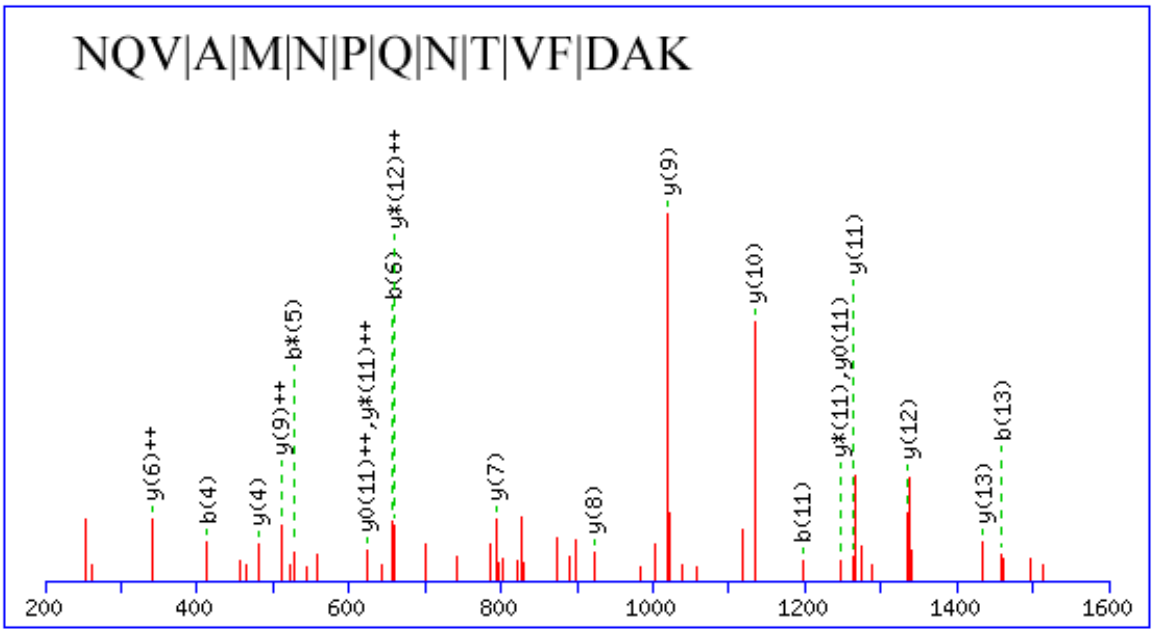
GEIEYGK**SRNPQVEDR**IEFFKRYCQILGVSPGKCYEERTFVSGLMMETI

(B)

Figure A-36. Tandem mass spectra of polypeptide (A) matched to hevein-like antimicrobial protein (*E. europaeus*) as indicated (bold) in the primary amino acid sequence (B).

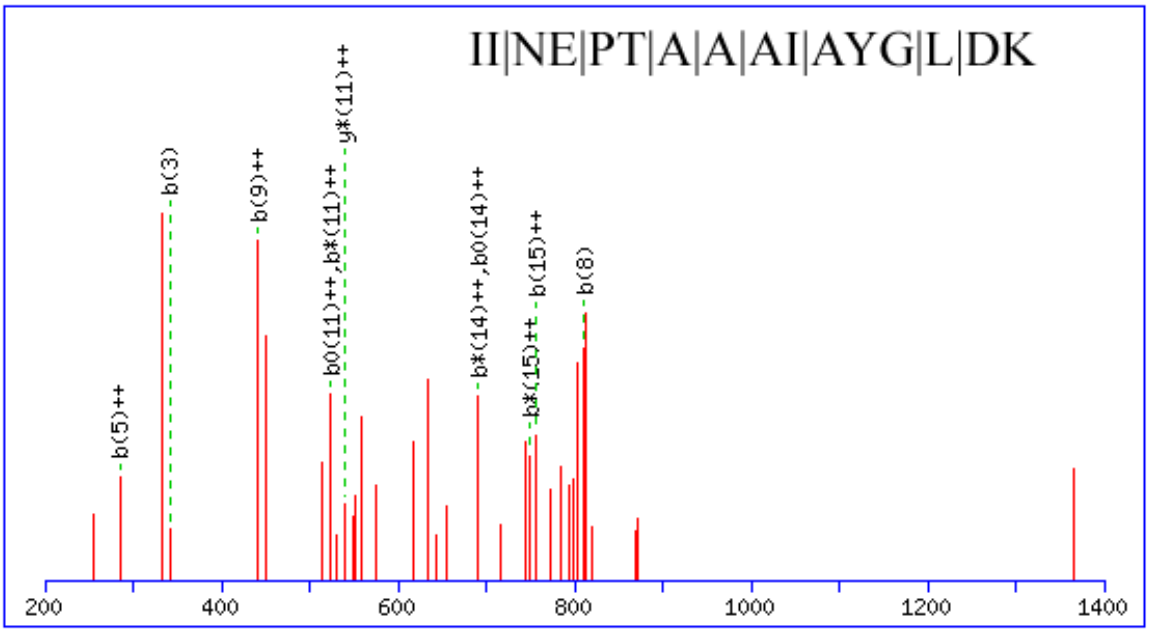


(A)

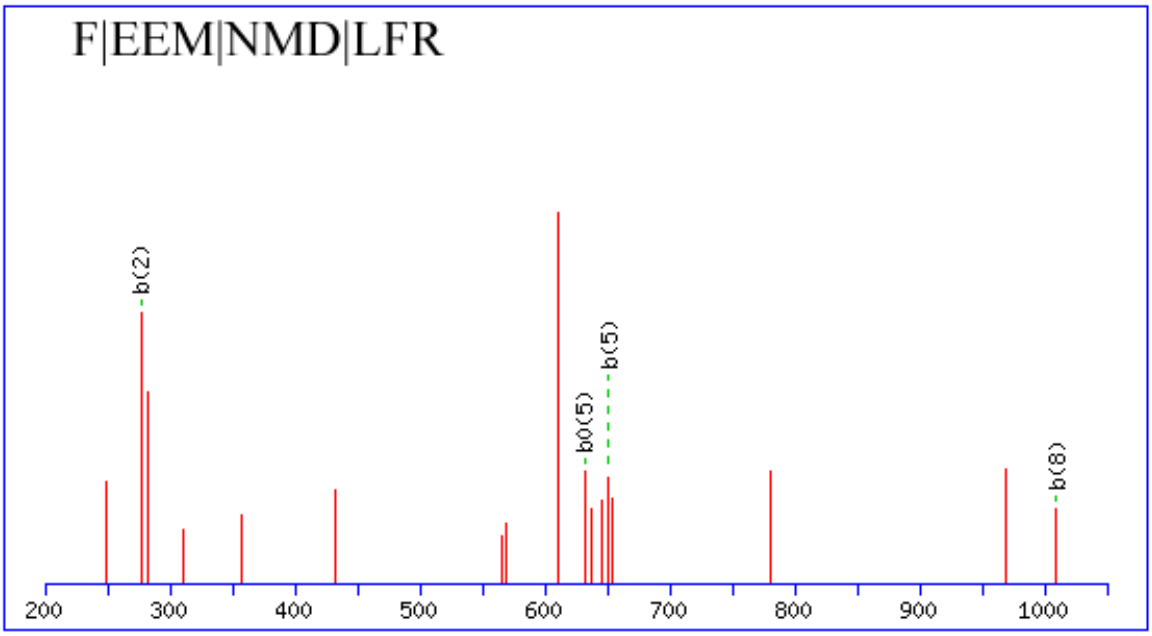


(B)

Figure A-37. Tandem mass spectra of polypeptides (A,B,C,D,E,F,G) matched to HSP 70 protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (H).

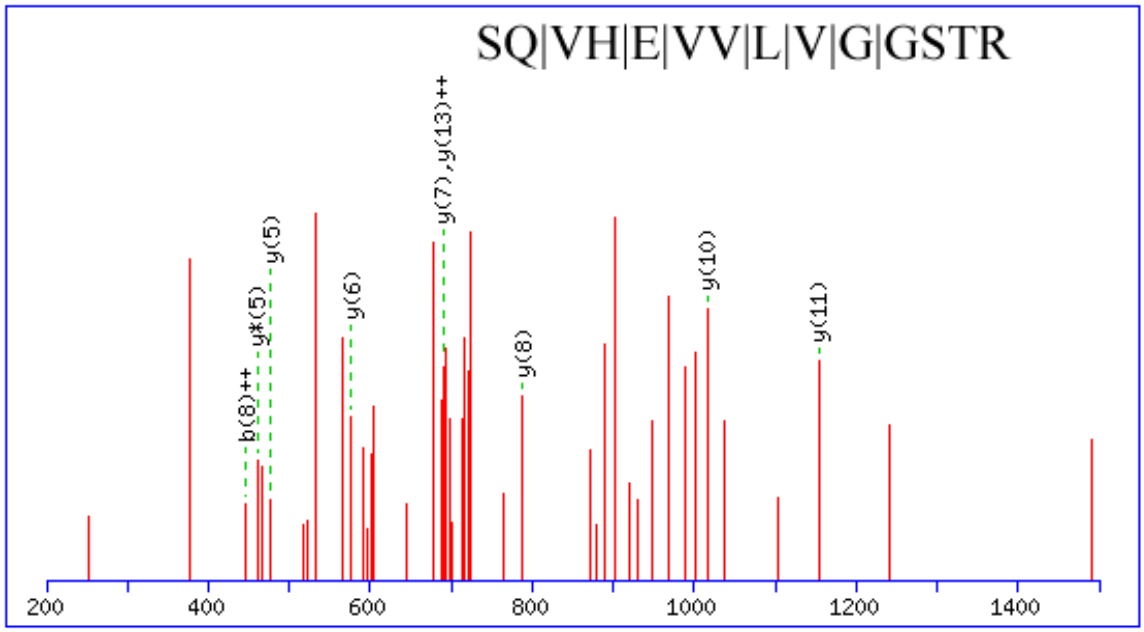


(C)

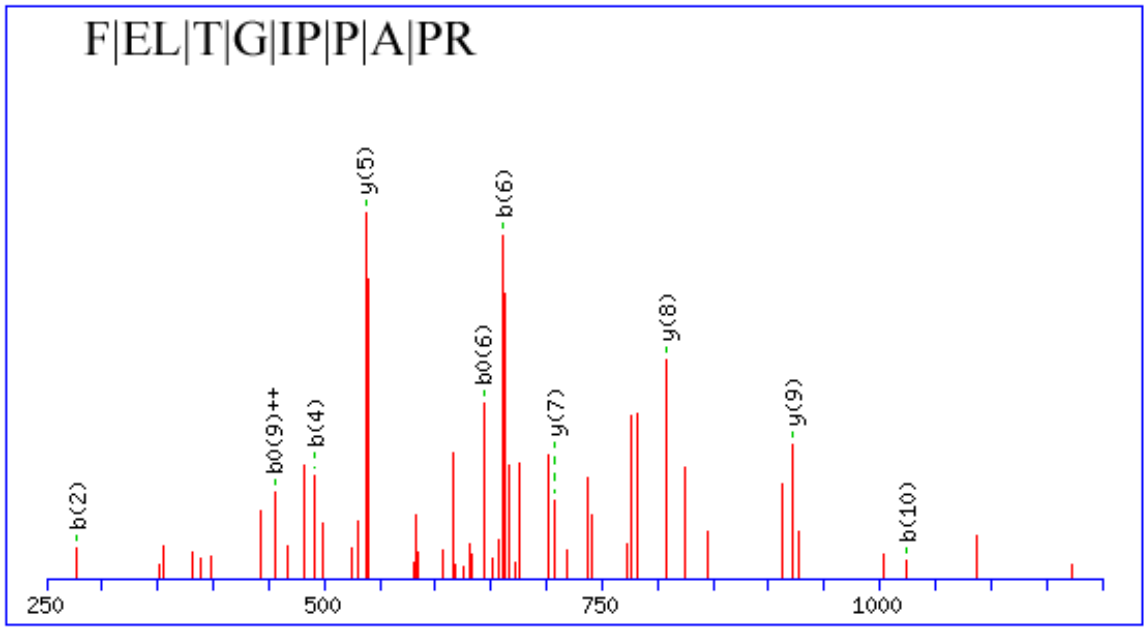


(D)

Figure A-37 (cont'd). Tandem mass spectra of polypeptides (A,B,C,D,E,F,G) matched to HSP 70 protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (H).

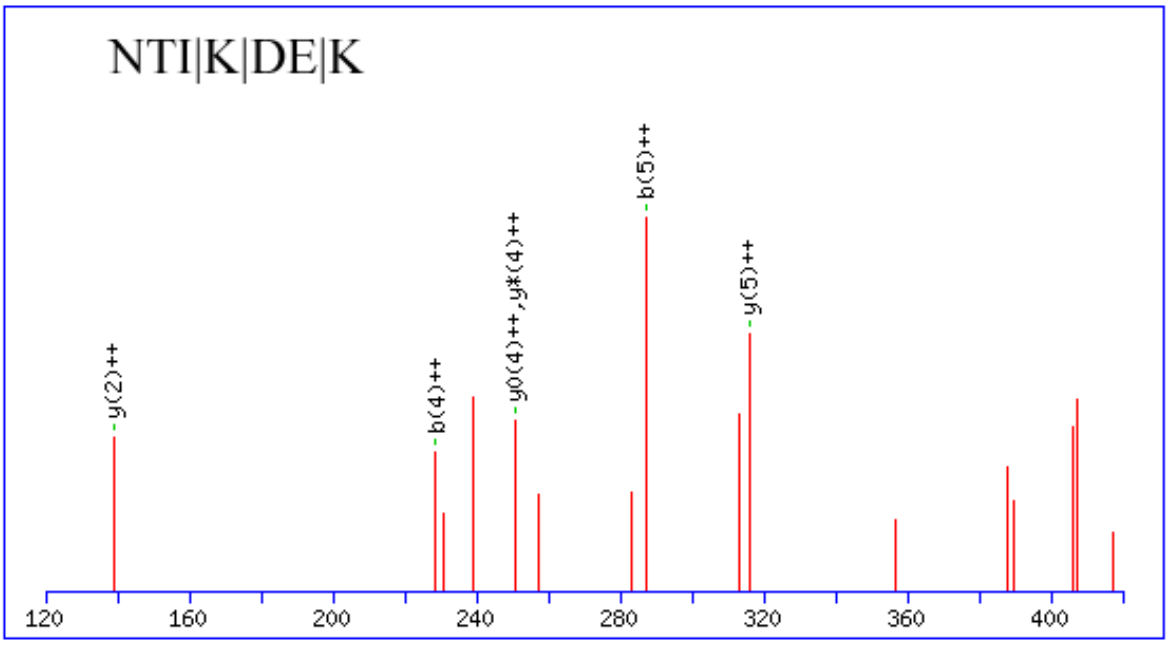


(E)



(F)

Figure A-37 (cont'd). Tandem mass spectra of polypeptides (A,B,C,D,E,F,G) matched to HSP 70 protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (H).



(G)

MATKSEKAIGIDLGTTYSCVGVWMNDRVEIIPNDQGNR**TTPSYVAFTDTER**LIGDAAK**NQVALN**

PQNTVFDAKRLIGRKFSDPSVQSDILHWPFKVVSGPGEKPMIVVSYKNEEKQFSPEEISSMVLVKM

KEVAEAFLGRTVKNAVVTVPAYFNDSQRQATKDAGAISGLNVLR**IINEPTAAAIAYGLDK**KGTK

AGEKNVLIFDLGGGTFDVSLLTIEEGVFEVKATAGDTHLGGEDFDNRLVNHFVAEFRRKHKKDIA

GNARALRRLRTACERAKRTLSSTAQTTIEIDSLHEGIDFYATISRAR**FEEMNMDLFR**KCMDPVEK

VLKDAKLDK**SSVHDVVLVGGSTR**IPKIQQLLQDFFNGKELCKSINPDEAVAYGAAVQAAILTGEG

SEKVQDLLLLDVAPLSLGLETAGGVMTVLIPRNTTVPCKKEQVFSTYADNQPGVLIQVYEGERAR

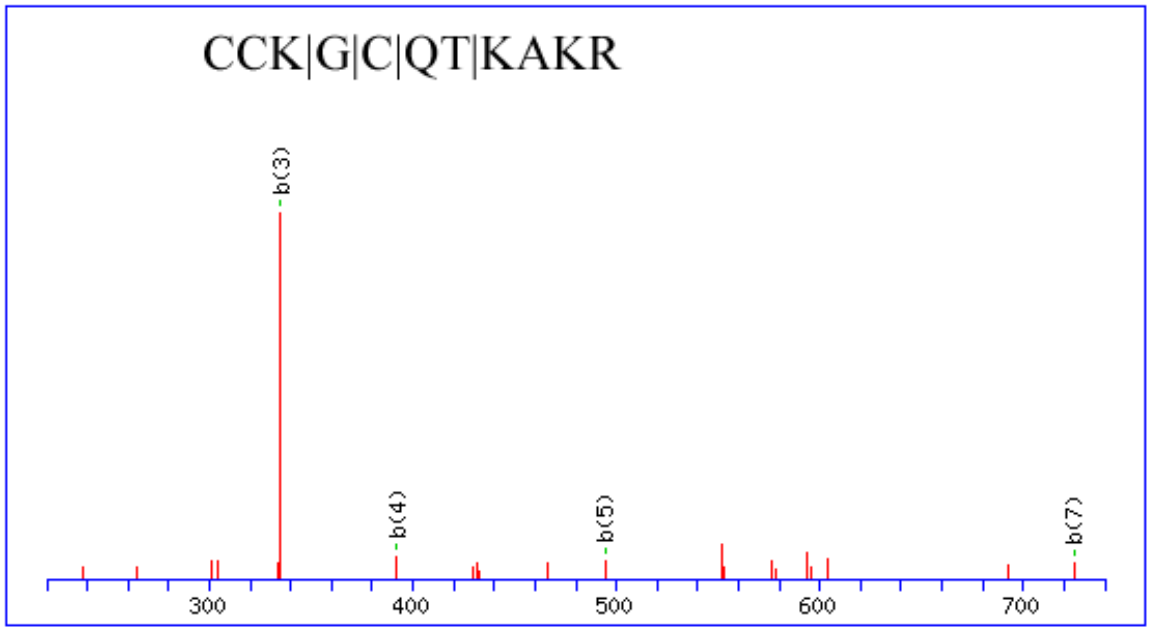
TRDNNLLGT**FELKGIPPAPR**GVPQINVCFDIDANGILNVSAEDKTAGVKNQITITNDKGRLSKEEIE

KMVQDAEKYKAEDEQVKKKVEAKNSLENYAYNMR**NTIKDEK**LAQKLTQEDKQKIEKAIDETIE

WIEGNQLAEVDEFEYKLKELEGICNPIISKMYQGGAAAGGMPTDGDFSSSGAAGGPKIEEVD

(H)

Figure A-37 (cont'd). Tandem mass spectra of polypeptides (A,B,C,D,E,F,G) matched to HSP 70 protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (H).



(A)

MISYVDIQQMTLHLIMKIPSTISREHGKLYKFRRIQQYIKTSMDSFLKVTYLIFIVV

FYIRVSSKGAYAMANLSPPVRTCVLRVGIK**CCKGCQTKAKR**KLLNVSGVSTVE

YNAEQGLLTVTGDANPTTLLHKLTKWGKKAELVSFLGDNYSSHVPRTPEQNQN

KTMEKKKKKPTKCCLLMCFGNKRSKNTKIEPMAIPNWQYRGVGNENGNARPFI

NAAMSPPMVYPPPQAVPGFTTPIPYPPPSFFPGRPPPPYTGAGMFQSAPPQSPPYFP

VPNPRLHYPHH

(B)

Figure A-38. Tandem mass spectra of polypeptide (A) matched to metal ion binding protein (*Arabidopsis*) as indicated (bold) in the primary amino acid sequence (B).

APPENDIX B

ADDITIONAL FIGURES FOR CH II

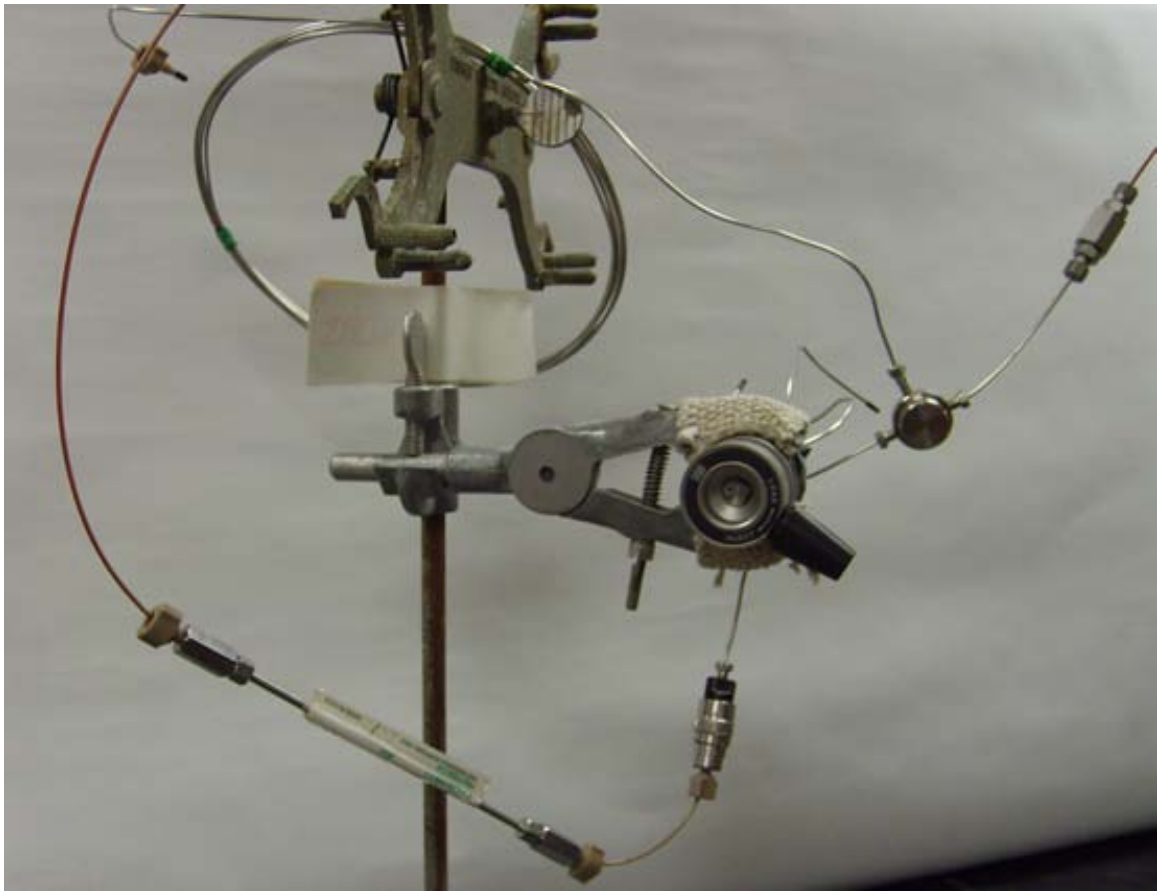


Figure B-1. Chromatographic apparatus designed for whole protein separation.

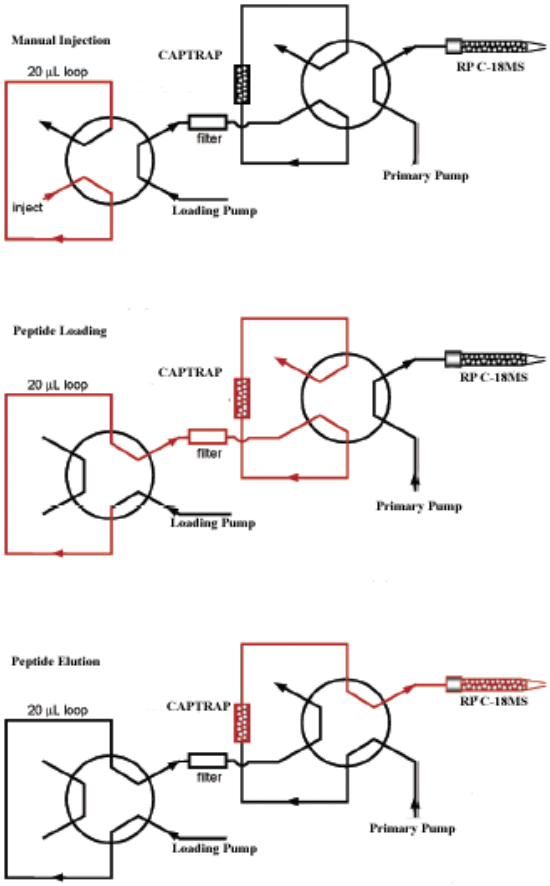


Figure B-2. Column switch diagram of offline cap-trap enrichment and online sample introduction.

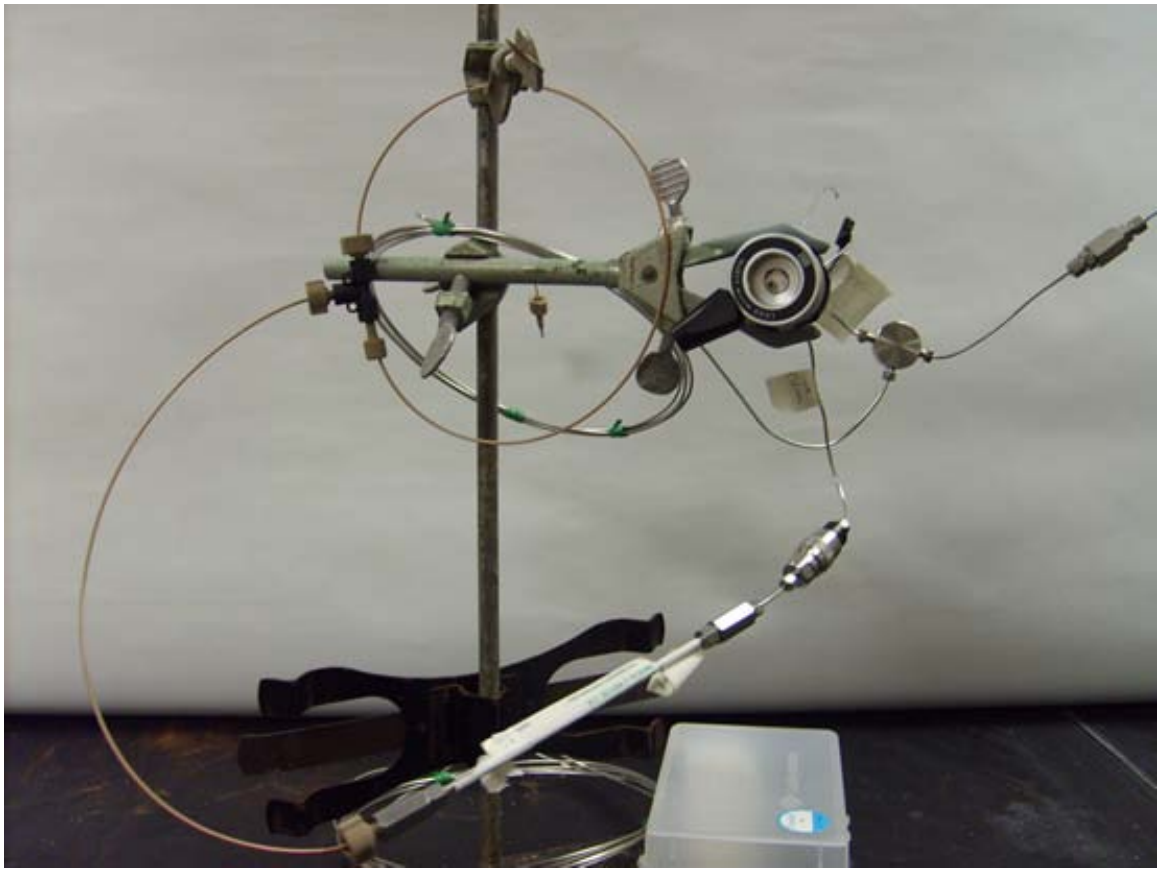


Figure B-3. Chromatographic apparatus designed for small molecule/peptide separation with post column split.

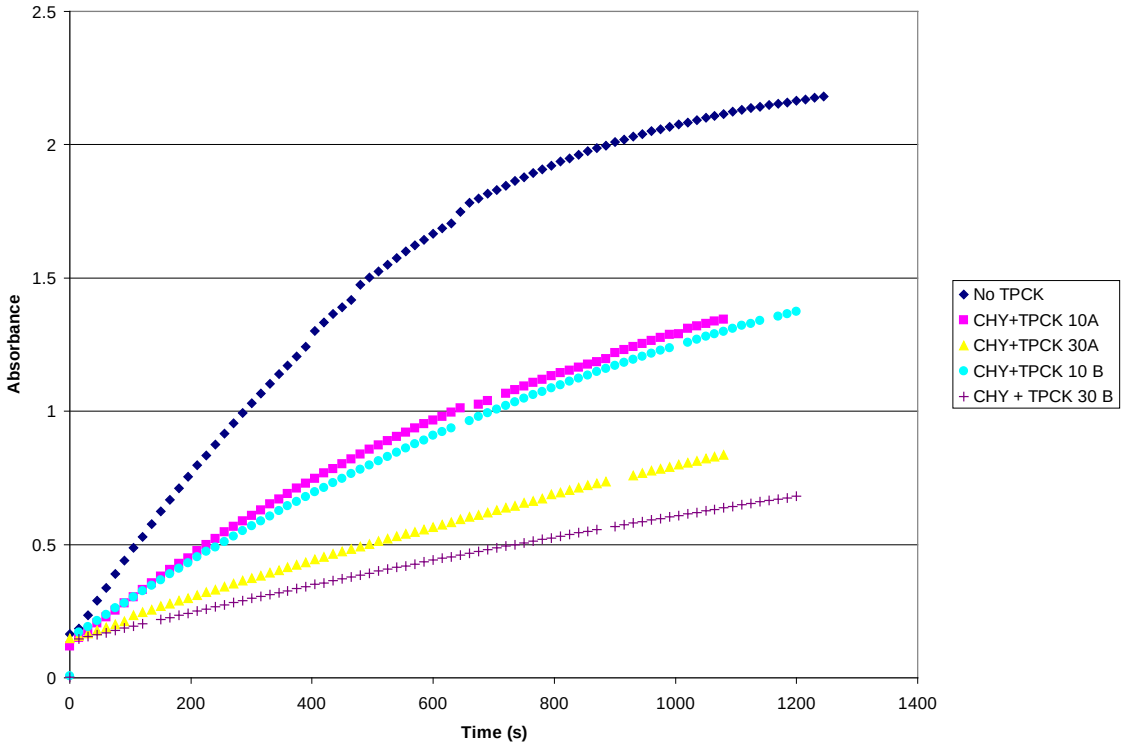


Figure B-4. UV spectrum analysis of CHY and CHY-TPCK samples enzymatic activity by reaction with BTEE observed at 256 nm.