University of Alberta

Development and Application of Mass Spectrometry for Bioanalysis of Proteins and Metabolites

by

Andrea G. De Souza

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

Spring 2011
Edmonton, Alberta

Examining Committee

Dr. Liang Li, Professor of Chemistry

Dr. Mark McDermott, Professor of Chemistry

Dr. Charles Lucy, Professor of Chemistry

Dr. Dennis Hall, Professor of Chemistry

Dr. Greg Goss, Professor of Biological Sciences

Dr. David Schriemer, Professor of Chemistry, University of Calgary

Abstract

The proteome, a hybrid term from *prote*in and gen*ome*, is the total set of proteins expressed by a genome in a cell, tissue, or organism at a given time. Proteomics, the study of the proteome, is concerned with a variety of topics such as identity, quantity, activity, structure, protein-protein interactions, and post-translational modifications. Mass spectrometry is an important tool used in bioanalysis, especially in proteomics research.

Recently, proteomics technologies have been applied to toxicology, giving rise to the field of toxicoproteomics. Fish are a valuable model for investigating the impact of toxins on aquatic environments and the zebrafish (*Danio rerio*) is an excellent candidate for proteomic studies. Gills are an attractive organ for toxicoproteomics, comprising >50% of the total surface area of fish and providing a direct route for the uptake of contaminants.

In this work, both qualitative and quantitative mass spectrometry techniques were applied in toxicoproteomics studies. We established the first large-scale proteome profile of a teleost fish tissue using two-dimensional liquid chromatography electrospray ionization tandem mass spectrometry in conjunction with a sequential protein solubilization method for protein fractionation and a precursor ion exclusion method for improving peptide and protein identification efficiency. The identified proteome exhibited excellent coverage of important biochemical pathways relevant to the function of the gill, as well as identifying numerous established and potential biomarkers of stress, disease, and environmental contamination.

With an established proteome profile, we utilized a stable isotope labelling strategy to quantify protein expression level changes in the gill upon exposure to two toxicants, naphthenic acids (NAs) and oil sands process water (OSPW), both of which are of environmental concern. Analysis of protein abundance changes in NAs-exposed zebrafish showed reproductive and immune response effects, and OSPW exposures exhibited expression level changes in a vast number of proteins crucial to the structural integrity and function of the gill, as well as proteins involved in immunoregulation and oxidative stress response.

In adhering to the use of mass spectrometry in bioanalysis, a method was developed by which urine samples were normalized in an attempt to account for dilution variation, thereby standardizing metabolite concentration measurements among samples.

Acknowledgements

The journey of my doctoral study has ended. I am eternally grateful to God for His continuing presence in my life. Through trials and tribulations, I gather strength from the knowledge that I can turn to Him for guidance and support.

I would like to thank my supervisor, Dr. Liang Li, for the opportunity to conduct research in his group and for his support. He was the primary reason that I chose to move from Vancouver to Edmonton to pursue my graduate studies, and I am grateful for the opportunities he has afforded me.

I would like to thank the people with whom I collaborated, Dr. Greg Goss and Dr. Tyson MacCormack. I learned a lot from both of you over the years and was constantly amazed by your vast biological knowledge. I am grateful to Greg not only for our research collaborations, but also for being a source of encouragement.

I would like to thank my thesis examining committee members, Dr. Dennis Hall, Dr. Charles Lucy, Dr. Mark McDermott, and Dr. David Schriemer, for taking the time to read my thesis and for their constructive comments.

It has been interesting working in the Li group and I would like to acknowledge the past and present members for contributing to my Li group experience.

Throughout my graduate studies, I utilized a number of resources in the Department of Chemistry. Thank you to the staff in the Mass Spectrometry

Facility, Biological Services, Machine Shop, Electronics Shop, Graduate Student Services, General Office, and Chemical Stores and Receiving.

Last, but not least, to my family and friends, thank you for making my Ph.D. graduate experience bearable.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

^{o}C	degree Celsius
%	percent
µg	microgram
µL	microliter
µm	micrometer
µmol	micromole
$^{13}CD_2O$	isotopically-enriched formaldehyde
2D	two-dimensional
2MEGA	N-terminal dimethylation after lysine guanidinylation
Å	Angstrom
ACN	acetonitrile
AMP	adenosine monophosphate
ATP	adenosine triphosphate
AUC	area under the curve
BCA	bicinchoninic acid
Ca	calcium
CCME	Canadian Council of Ministers of the Environment
CH_2O	formaldehyde
CID	collision-induced dissociation
CV	correlation of variation
d	distance
Da	Dalton

DAVID	Database for Annotation, Visualization, and Integrated Discovery
DC	direct current
DRE	dynamic range enhancement
DTT	dithiothreitol
EDC	endocrine disrupting chemical
EIC	extracted ion chromatogram
ESI	electrospray ionization
EtOH	ethanol
FA	formic acid
FKBP3	FK506 binding protein 3
G	gauge
g	gram
g	gravity (relative centrifugal force)
GA	geometric average
GC	gas chromatography
GnRH	gonadotropin-releasing hormone
GO	gene ontology
h	hour
HCl	hydrochloric acid
HPLC	high performance liquid chromatography
HSP	heat shock protein
IAA	iodoacetamide
ICP	inductively coupled plasma

ID	identity
IU	international unit
K	potassium
KCl	potassium chloride
KEGG	Kyoto Encyclopedia of Genes and Genomes
KH_2PO_4	potassium dihydrogen phosphate
L	litre
LC	liquid chromatography
LC_{50}	lethal concentration 50%
m	mass
m/z	mass-to-charge
M	molar
MALDI	matrix-assisted laser desorption ionization
MAPK	mitogen-activated protein kinase
MCP	multichannel plate
MeOH	methanol
MHC	major histocompatibility complex
min	minute
mL	milliliter
mm	millimeter
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MS/MS	tandem mass spectrometry

Na	sodium
NA	naphthenic acid
NaCl	sodium chloride
NADPH	nicotinamide adenine dinucleotide phosphate
$NaHCO_3$	sodium bicarbonate
NH_4HCO_3	ammonium bicarbonate
nL	nanoliter
nm	nanometer
NMR	nuclear magnetic resonance
OSPW	oil sands process water
P_i	inorganic phosphate
PAGE	polyacrylamide gel electrophoresis
PAH	polycyclic aromatic hydrocarbons
PBS	phosphate-buffered saline
PE	polyethylene
pH	potential of hydrogen
PI	protease inhibitor
PIE	precursor ion exclusion
PPAR	peroxisome proliferator-activated receptor
Q	quadrupole
q	charge
QIT	quadrupole ion trap
R^2	square of the correlation coefficient

RF	radiofrequency
SCX	strong cation exchange
SD_{geo}	geometric standard deviation
SDS	sodium dodecyl sulphate
t	time
TCA	tricarboxylic acid
TFA	trifluoroacetic acid
TMS	tricaine methanesulfonate
ToF or Tof	time-of-flight
U	DC voltage
UDP	uridine diphosphate
UPLC	ultra performance liquid chromatography
UV	ultraviolet
V	voltage
v	velocity
VEGF	vascular endothelial growth factor
v/v	volume/volume
w/w	weight/weight
ZGC	Zebrafish Genome Collection

Chapter 1

Introduction

1.1 Proteomics and Mass Spectrometry

The proteome, a hybrid term from *prote*in and gen*ome*, is the total set of proteins expressed by a genome in a cell, tissue, or organism at a given time. Proteomics, the study of the proteome, is traditionally concerned with a variety of topics such as identity (proteome mapping), post-translational modifications, quantity, activity (functional proteomics), three-dimensional structure (structural proteomics), and protein-protein interactions. Application of proteomics approaches to toxicology research, also known as toxicoproteomics, focuses on elucidating mechanisms of action and identifying protein biomarkers of toxicity by studying changes in protein expression upon exposure to toxicants or as a result of adverse drug effects. A proteome profile provides a base from which postulations on biological function can be inferred by mapping identified proteins to known biological pathways, as was done in Chapter 2, and can be applied to understanding quantitative protein expression changes, as was done in Chapters 3 and 4.

Traditionally, peptides were sequenced using Edman degradation, where the amino-terminal residue was labelled, cleaved from the rest of the peptide, and subsequently detected by liquid chromatography ultraviolet (LC/UV) spectroscopy.[1] This process continued in a stepwise manner to determine the sequence of the peptide. With advances in mass spectrometry (MS), the time-consuming Edman degradation process was replaced with mass spectrometry-

based protein identification strategies, such as two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). In 2D-PAGE, proteins are separated in the first dimension based on their isoelectric points, followed by electrophoresis in the presence of sodium dodecyl sulphate (SDS) to further separate proteins based on their molecular mass in the second dimension. The proteins can then be extracted from the gel, digested, and analyzed by MS for further analysis by database searches. Despite its success, 2D-PAGE has its limitations including identification of low-abundance proteins, membrane proteins, and proteins with extremes in isoelectric points and molecular weight. The limitations of 2D-PAGE led to the development of an alternative solution-based proteomics approach, commonly referred to as shotgun proteomics.

In a typical shotgun proteomics workflow, a protein sample (e.g. cell lysate, tissue extract, subcellular fraction) is digested into peptides, separated by multidimensional LC, and analyzed by tandem mass spectrometry (MS/MS) with subsequent database searching. Typically, in the multidimensional LC separation, the digested peptide mixture is first separated on the basis of charge by strong cation exchange (SCX), followed by separation on the basis of hydrophobicity by reversed phase LC. The separated peptide solution is introduced into the mass spectrometer via electrospray ionization and peptides are fragmented in the mass spectrometer. The resulting MS/MS spectra can then be searched in databases that use algorithms to identify the peptide sequences and match them to proteins. There are reviews in the literature describing the various ionization and peptide fragmentation techniques,[2, 3] mass analyzers,[4, 5] and databases[6] used in shotgun

proteomics. Consequently, I will focus only on those techniques used in the proteomics studies in this thesis.

1.1.1 Electrospray Ionization

The analysis of biomolecules by mass spectrometry was revolutionized in the 1980s with the introduction of soft ionization techniques, whereby large, polar, and thermally labile molecules could be ionized and vapourized. This achievement was recognized in 2002 with the awarding of the Nobel Prize in chemistry to Koichi Tanaka and John Fenn for their development of the soft ionization techniques matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI), respectively, for analyses of biological macromolecules by MS. ESI was the ionization technique used for the work presented in this thesis.

The ESI process is composed of four major events: nebulization, desolvation, fission, and production of gas phase ions. A solution is infused through a thin capillary to which a potential is applied, resulting in the formation of charged droplets (i.e. nebulization). Solvent evaporation (i.e. desolvation) from the charged droplets is aided by a heated drying gas, such as nitrogen. As the charged droplets shrink in size due to desolvation, the surface Coulombic forces of the ions increase and eventually exceed the surface tension of the solvent, resulting in smaller droplets (i.e. fission). Fission continues until an ion desorbs from a droplet (i.e. ion evaporation model)[7, 8] or solvent is completely evaporated (i.e. charge residue model).[9] The formation of gas phase ions during ESI is illustrated in Figure 1.1.

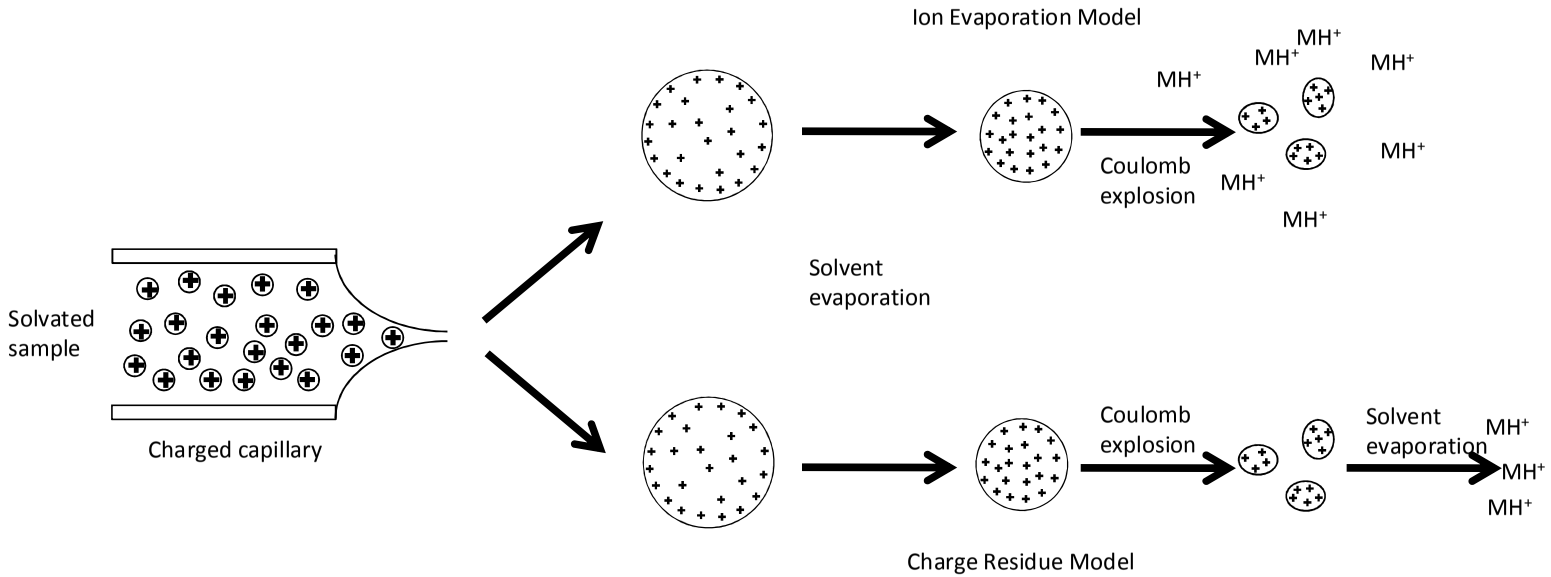

Figure 1.1. The electrospray ionization process resulting in the formation of gas phase ions.

1.1.2 Peptide Ion Fragmentation and Database Searching

Of the existing methods for peptide ion fragmentation,[10-12] the most common method involves collision of the peptide of interest with an inert gas in a process referred to as collision-induced dissociation (CID). The kinetic energy of the collisions is converted to internal energy, resulting in fragmentation of the precursor ion into product ions. Ion fragmentation typically occurs across the peptide backbone at the C-C, C-N, or N-C bond. The resulting product ions are classified depending on which bond is broken and where the charge is retained. If the collision results in charge retention on the N-terminal fragment, the product ion is classified as either an *a*, *b*, or *c* ion. If the charge is retained on the C-terminal fragment, the product ion is classified as either an *x*, *y*, or *z* ion. Figure 1.2 illustrates the nomenclature of peptide ion fragmentation. Low energy CID spectra are dominated by *b* and *y* ions, which are further classified as b_n or y_n ions to reflect the amino acid residues position on the N-terminus or C-terminus fragment of the amide bond. Furthermore, if the peptide sequence contains arginine, lysine, asparagine or glutamine, a loss of ammonia (-17 Da) can be observed and is denoted as a b^* or y^* ion. A loss of water (-18 Da) in the presence of serine, threonine, aspartic acid or glutamic acid is denoted as a b^o or y^o ion.

The dataset of acquired MS/MS spectra is searched through a database to assign peptide and protein identities. There are several database searching algorithms available for assigning peptide and protein identities,[13-16] such as MASCOT.[16] MASCOT uses a probability-based scoring algorithm[17] to assess the probability that a match between the observed mass spectrum and a theoretical

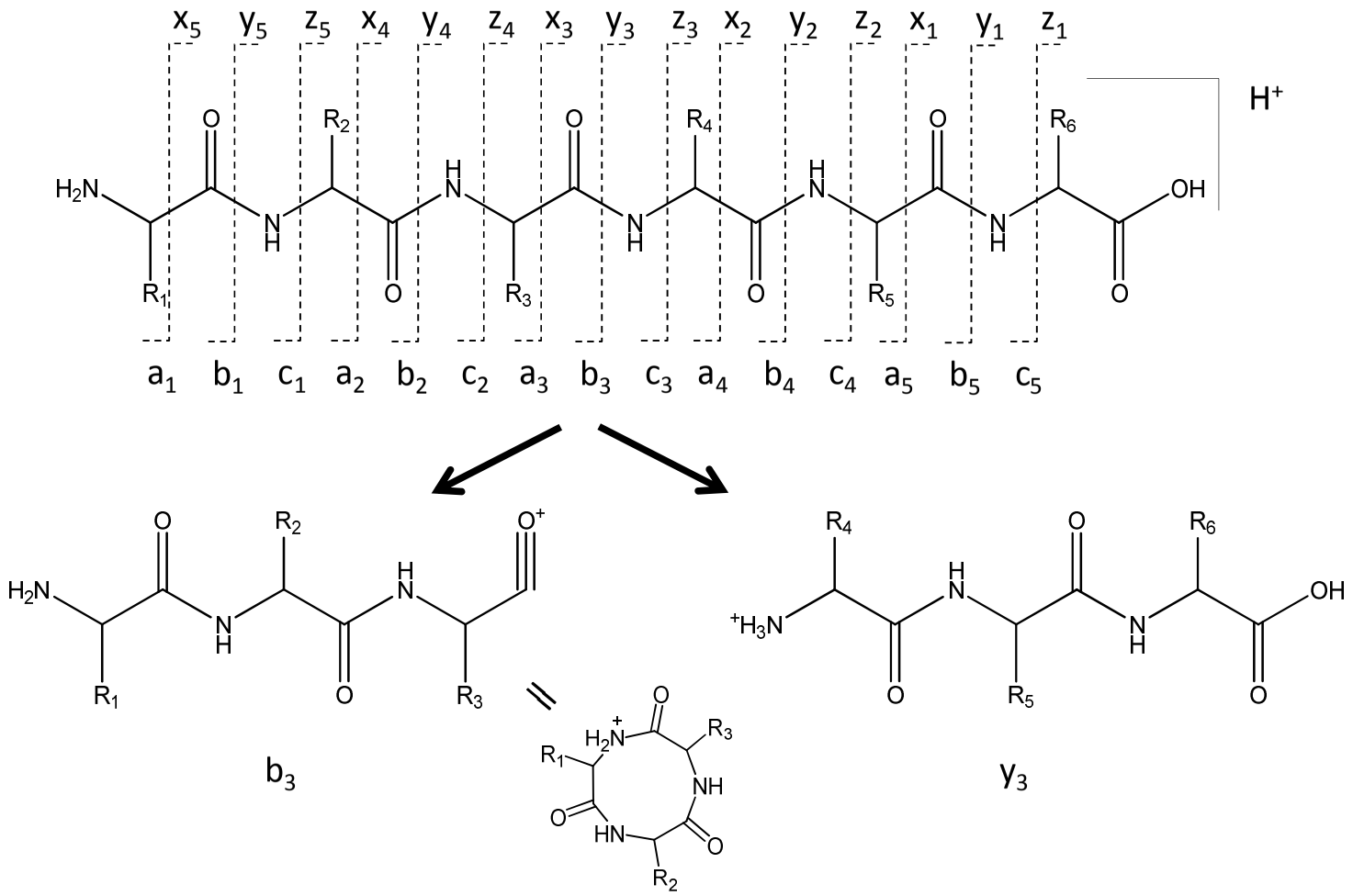

Figure 1.2. Peptide fragmentation nomenclature.

mass spectrum is due to chance.

1.1.3 The Quadrupole Mass Analyzer

The quadrupole (Q) mass analyzer consists of four axial rods to which a direct current (DC) potential is applied and to which a radiofrequency (RF) potential, 180 degrees out of phase, is applied to each pair of rods, resulting in an electric field through which only a certain m/z range passes for a given combination of potentials (Figure 1.3). Ions outside the m/z range hit the rods and are expelled. The potential applied to the rods is of the form: $\Phi_o = \boldsymbol{U} + \boldsymbol{V}\cos\boldsymbol{\omega t}$, where U is the DC voltage and $V\cos\omega t$ is the RF potential of frequency $\omega/2\pi$. The RF potential is applied in two waveforms, + and -, to two pairs of opposing rods, and the polarity of the rods is switched, resulting in a fluctuation in the trajectory of an ion in the electric field. Varying U, V, and ω will transmit ions with different m/z toward the detector. The versatility of quadrupoles is that they function not only as mass analyzers, but also as ion transfer optics (RF only mode) and collision cells.

1.1.4 The Time-of-Flight Mass Analyzer

The time-of-flight (ToF) mass analyzer consists of a field-free tube through which ions drift after being accelerated by an electric field. Since potential energy is equal to qV, where q is the charge and V is the voltage, and kinetic energy is equal to $\frac{1}{2}mv^2$, where m is mass and v is velocity, then

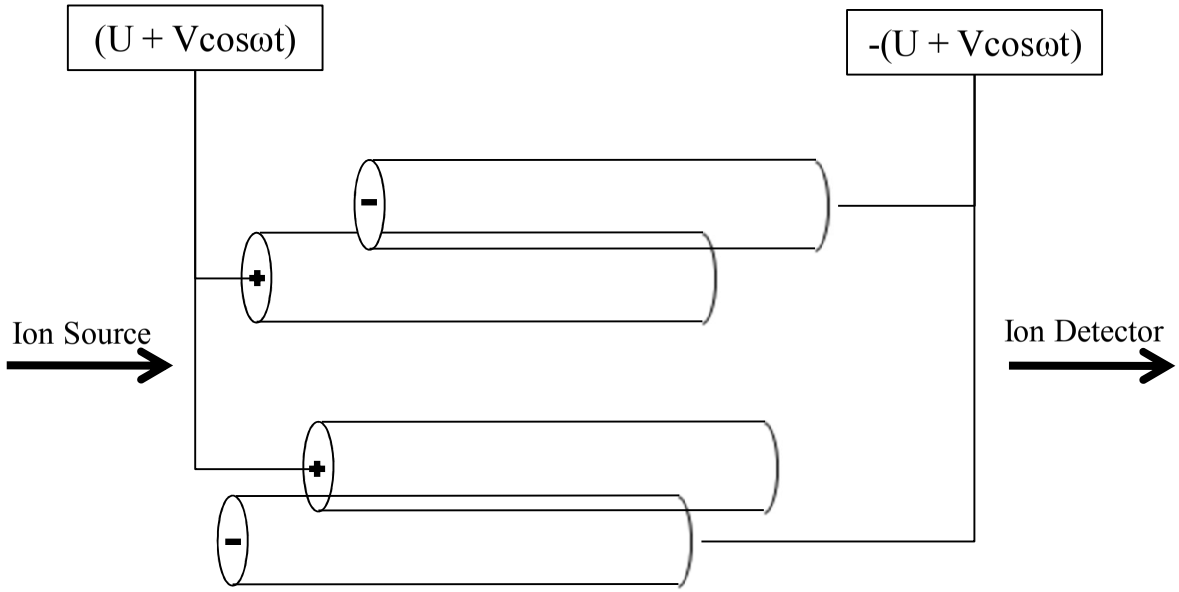

Figure 1.3. Quadrupole mass analyzer.

$qV = \frac{1}{2}mv^2$. Keeping in mind that velocity is equal to distance (d) over time (t), this equation can be rearranged to $\frac{m}{q} = 2V\left(\frac{t}{d}\right)^2$. This equation demonstrates that ions of higher mass-to-charge ratios ($\frac{m}{q}$) will take longer to travel through the flight tube than ions of lower mass-to-charge ratios.

Traditional time-of-flight analyzers suffer from the drawback that ions of the same mass-to-charge ratio can reach the detector at different times due to initial energy distribution. A reflectron time-of-flight analyzer compensates for this initial energy distribution through the use of a series of electric grids of increasing potential. Depending on their kinetic energy, ions of the same mass-to-charge ratio will penetrate the electric field at different depths; ions with greater kinetic energy will penetrate the field deeper and spend more time in the reflectron than ions of lower kinetic energy. The end result is both ions reaching the detector simultaneously.

1.1.5 Hybrid Quadrupole Orthogonal Time-of-Flight Mass Spectrometer

For the work presented in this thesis, the Waters Q-Tof Premier™ mass spectrometer was used and combines the scanning capabilities of a quadrupole with the resolving power of a time-of-flight. Figure 1.4 shows a schematic of this hybrid mass analyzer. In the Q-Tof Premier™, the NanoLockSpray interface generates gas phase ions which are directed to the quadrupole mass analyzer with the quadrupole acting as an ion focussing device. In Tof MS acquisition, the

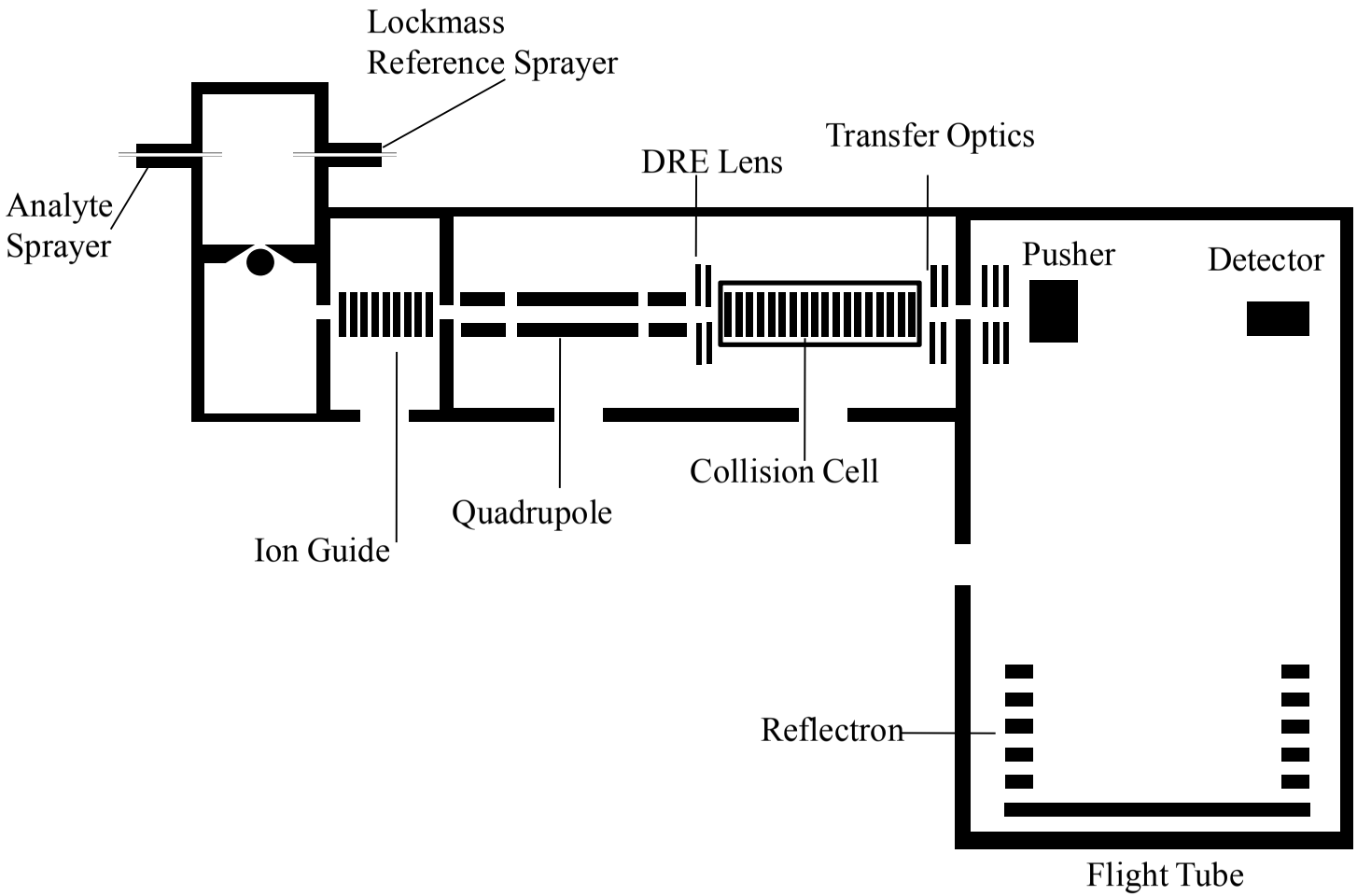

Figure 1.4. Schematic of the Waters ESI Q-ToF.

quadrupole operates in RF only mode and the ions pass through the quadrupole and are detected by a multichannel plate (MCP) detector at the end of the flight in the ToF. In ToF MS/MS acquisition, the DC voltage of the quadrupole is on and only select ions are fragmented in the collision cell and the resulting product ions detected at the end of flight in the ToF.

1.2 Toxicoproteomics and Aquatic Toxicology

The discipline of toxicoproteomics emerged from advances in global protein expression technologies and their application in toxicology. There is increasing concern about the adverse effects resulting from environmental exposure to manufactured chemicals and other anthropogenic pollutants; fish are often used to study these effects in the aquatic environment. Early work focussed almost exclusively on acute lethality assays and data from these studies were used in environmental regulation.[18] As recognition of the importance of sublethal effects became more apparent, research on the pathological response of exposure in early development and reproduction became more common, particularly research involving fish exposure to endocrine disrupting chemicals (EDCs).[19, 20] Increasing emphasis on elucidating mechanisms of action and biomarker discovery has led to toxicoproteomics applications in aquatic toxicology.[21]

Our interest was in utilizing proteomic technologies to study exposure to toxicants on gills of zebrafish.

1.2.1 Zebrafish as a Vertebrate Model

The zebrafish (*Danio rerio*) is a freshwater teleost that has gained popularity as a vertebrate model since it was first introduced to study vertebrate development almost 30 years ago.[22] The teleosts are a class of bony fish that represent over 50% of all vertebrate species. There are several features of zebrafish that make it attractive for use in research. Zebrafish are small (2–4 centimeters long) and highly fecund (a single female can lay 200 eggs per spawning),[23] which are contributing factors to low husbandry costs. Furthermore, the rapid and external development, as well as transparency of zebrafish embryos, allows for visualization of embryogenesis and organogenesis, which is complete 2-4 days post fertilization.[24] Only in recent years with the sequencing and annotation of the zebrafish genome has the potential of zebrafish in proteomics been realized.[21, 25-29] The advent of the zebrafish genome sequencing project by the Wellcome Trust Sanger Institute in 2001[30] further promoted the use of zebrafish as a model system for vertebrate biology.

There is conservation of developmental pathways and genes between zebrafish and humans.[31] Aside from their prolific use as a model for understanding vertebrate development,[32-34] advances in forward and reverse genetics and genome sequencing have made zebrafish attractive as a model for studying human diseases, such as cardiovascular diseases,[35-39] renal diseases,[40-45] and neural diseases.[46-50]

Zebrafish embryos are amenable to high throughput screening and are gaining popularity in the drug discovery process.[51-53] Not only do zebrafish

embryos and larvae fit in individual wells of 96-well microtiter plates, but their small size translates to lower consumption of drugs, a cost-attractive feature. With emerging technologies, zebrafish will continue to play an important role as a vertebrate model.

1.2.2 Gill

The gill is a complex organ that performs numerous physiological functions including respiration, osmoregulation and ion regulation, acid-base regulation, and excretion of nitrogenous wastes.[54] The gill structure is supported by arches which branch out into multiple filaments and are further divided into thousands of lamellae. The large surface area of the filaments and lamellae are covered by an epithelium that provides a permeable barrier between the internal milieu of the fish and the external environment, and the diversity of the epithelial cell types specialize the gill to perform its numerous functions. The gill epithelium comprises over 50% of the total surface area of the fish[55] and provides a direct route for the uptake of contaminants,[56] making fish ideal for studying aquatic toxicology.

1.3 Metabolomics

Metabolomics is the comprehensive analysis of the complete set of metabolites in a cell, tissue, or organism. Although the term "metabolomics" is relatively new, the study of metabolites dates back to the 1970s.[57-59]

Metabolites are important regulators of cell processes and metabolic profiles provide valuable information in understanding cellular function.

Metabolites comprise a portion of the cellular products within a cell; transcripts and proteins comprise the other products. Correlation of transcript and protein expression is low[60-62] and that between transcription and metabolites even lower. Therefore, combining transcriptomics, proteomics, and metabolomics can provide a complete snapshot of the physiological status of an organism than any one of the "omics."

Metabolomics utilizes the analytical techniques of nuclear magnetic resonance (NMR) spectroscopy,[63-65] gas chromatography mass spectrometry (GC/MS),[66, 67] and liquid chromatography mass spectrometry (LC/MS),[68-70] and these techniques can be combined to provide complementary information.[71]

Metabolomics applications in human health include the identification of disease biomarkers, such as for liver cancer,[72] diabetes,[73] and inborn errors of metabolism.[74]

1.3.1 The Quadrupole Ion Trap Mass Analyzer

The quadrupole ion trap (QIT) mass analyzer is composed of two end-cap electrodes and one ring electrode, arranged to create a space where ions can be stored. A schematic of a quadrupole ion trap mass analyzer is shown in Figure 1.5. Applying DC and RF voltages to the end-cap and ring electrodes, respectively, confine ions in this space. A bath gas, typically helium, is introduced into the trap to provide collisional cooling of the analyte ions and to efficiently confine the ions to the center of the trap. In the absence of a bath gas, ions are distributed throughout the volume of the trap.

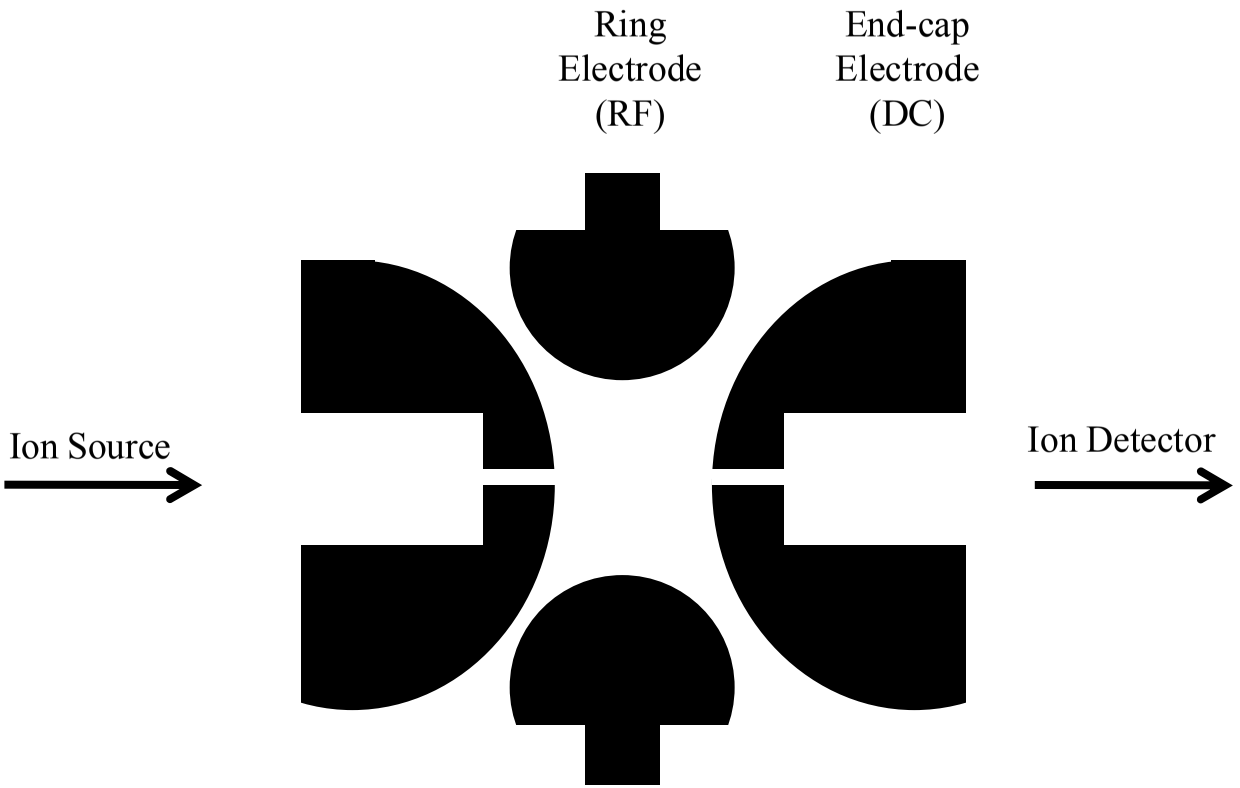

Figure 1.5. Quadrupole ion trap mass analyzer.

QITs can be operated in the mass-selective stability mode or the mass-selective instability mode. In the former, application of DC and RF voltages confine ions with a stable trajectory and the application of a DC voltage to the ring electrode ejects the ions. In the mass-selective instability mode, ions of specific mass-to-charge ratios are ejected from the trap selectively by increasing the RF voltage linearly on the ring electrode and leaving the DC voltage grounded. As the RF voltage increases, ions of different mass-to-charge ratios will move towards unstable trajectories and be ejected from the trap, resulting in a mass spectrum.

Tandem mass spectrometry can also be performed in the QIT. A selected ion can be isolated in the trap by applying a broad RF frequency containing all but the RF frequency of the ion of interest, thereby ejecting all but the precursor ion. An RF frequency matching that of the precursor ion is then applied, increasing the kinetic energy of the ions. The kinetic energy translates to internal energy, resulting in ion fragmentation. Operation in the mass-selective instability mode results in an MS/MS spectrum.

A Bruker esquire3000plus quadrupole ion trap mass spectrometer was used for the MS analysis described in Chapter 5.

1.4 Scope of the Thesis

Chapters 2 to 4 will describe our progress in using the zebrafish as a model in aquatic toxicology studies. We first established the proteome of the gill in Chapter 2 using 2D LC-ESI MS/MS. In Chapters 3 and 4, we evaluate the

effect of zebrafish exposure to naphthenic acids and oil sands process water, respectively, on the proteins of the gill using quantitative techniques in an attempt to understand the exposure effects on biochemical pathways. Chapter 5 focuses on developing a method to account for urinary volume variation. The thesis will end with a Conclusions and Future Work section described in Chapter 6.

1.5 Literature Cited

1. Edman, P., Method for determination of the amino acid sequence in peptides. *Acta Chem. Scand.* **1950,** 4, 283-293.

2. Barlow, C. K.; O'Hair, R. A. J., Gas-phase peptide fragmentation: how understanding the fundamentals provides a springboard to developing new chemistry and novel proteomic tools. *J. Mass Spectrom.* **2008,** 43, (10), 1301-1319.

3. Faull, K. F.; Dooley, A. N.; Halgand, F.; Shoemaker, L. D.; Norris, A. J.; Ryan, C. M.; Laganowsky, A.; Johnson, J. V.; Katz, J. E., An introduction to the basic principles and concepts of mass spectrometry. *Compr. Anal. Chem.* **2009,** 52, 1-46.

4. Volmer, D. A.; Sleno, L., Mass analyzers: An overview of several designs and their applications, Part II. *Spectroscopy (Duluth, MN, U. S.)* **2005,** 20, (12), 90-95.

5. Volmer, D. A.; Sleno, L., Mass analyzers: An overview of several designs and their applications, Part I. *Spectroscopy (Duluth, MN, U. S.)* **2005,** 20, (11), 20-26.

6. Kapp, E.; Schütz, F., Overview of tandem mass spectrometry (MS/MS) database search algorithms. *Curr Protoc Protein Sci* **2007,** 49, 25.2.1-25.2.19.

7. Iribarne, J. V.; Thomson, B. A., On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **1976,** 64, (6), 2287-2294.

8. Thomson, B. A.; Iribarne, J. V., Field-induced ion evaporation from liquid surfaces at atmospheric pressure. *J. Chem. Phys.* **1979,** 71, (11), 4451-4463.

9. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B., Molecular beams of macroions. *J. Chem. Phys.* **1968,** 49, (5), 2240-2249.

10. Jennings, K. R., Collision-induced decompositions of aromatic molecular ions. *Int. J. Mass Spectrom. Ion Phys.* **1968,** 1, (3), 227-35.

11. Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** 101, (26), 9528-9533.

12. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W., Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **1998,** 120, (13), 3265-3266.

13. Craig, R.; Beavis, R. C., TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **2004,** 20, (9), 1466-1467.

14. Eng, J. K.; McCormack, A. L.; Yates, J. R., III, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994,** 5, (11), 976-89.

15. Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X.; Shi, W.; Bryant, S. H., Open mass spectrometry search algorithm. *J. Proteome Res.* **2004,** 3, (5), 958-964.

16. Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999,** 20, (18), 3551-3567.

17. Pappin, D. J. C.; Hojrup, P.; Bleasby, A. J., Rapid indentification of proteins by peptide-mass fingerprinting. *Curr. Biol.* **1993,** 3, (6), 327-332.

18. McKim, J. M., Evaluation of tests with early life stages of fish for predicting long-term toxicity. *J. Fish. Res. Board Can.* **1977,** 34, (8), 1148-1154.

19. Kavlock, R. J.; Ankley, G. T., A perspective on the risk assessment process for endocrine-disruptive effects on wildlife and human health. *Risk Anal* **1996,** 16, (6), 731-739.

20. Larkin, P.; Sabo-Attwood, T.; Kelso, J.; Denslow, N. D., Analysis of Gene Expression Profiles in Largemouth Bass Exposed to 17-β-Estradiol and to Anthropogenic Contaminants that Behave as Estrogens. *Ecotoxicology* **2003,** 12, (6), 463-468.

21. Shrader, E. A.; Henry, T. R.; Greeley, M. S.; Bradley, B. P., Proteomics in zebrafish exposed to endocrine disrupting chemicals. *Ecotoxicology* **2003,** 12, (6), 485-488.

22. Streisinger, G.; Walker, C.; Dower, N.; Knauber, D.; Singer, F., Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature* **1981,** 291, (5813), 293-296.

23. Dahm, R.; Geisler, R., Learning from Small Fry: The zebrafish as a genetic model organism for aquaculture fish species. *Mar. Biotechnol.* **2006,** 8, (4), 329-345.

24. Kimmel, C. B.; Ballard, W. W.; Kimmel, S. R.; Ullmann, B.; Schilling, T. F., Stages of embryonic development of the zebrafish. *Dev Dyn* **1995,** 203, (3), 253-310.

25. Link, V.; Shevchenko, A.; Heisenberg, C.-P., Proteomics of early zebrafish embryos. *BMC Dev. Biol.* **2006,** 6:1, doi:10.1186/1471-213X-6-1.

26. Lucitt, M. B.; Price, T. S.; Pizarro, A.; Wu, W.; Yocum, A. K.; Seiler, C.; Pack, M. A.; Blair, I. A.; FitzGerald, G. A.; Grosser, T., Analysis of the zebrafish proteome during embryonic development. *Mol. Cell. Proteomics* **2008,** 7, (5), 981-994.

27. Posner, M.; Hawke, M.; LaCava, C.; Prince, C. J.; Bellanco, N. R.; Corbin, R. W., A proteome map of the zebrafish (Danio rerio) lens reveals similarities between zebrafish and mammalian crystallin expression. *Mol. Vision* **2008,** 14, 806-814.

28. Tay, T. L.; Lin, Q.; Seow, T. K.; Tan, K. H.; Hew, C. L.; Gong, Z., Proteomic analysis of protein profiles during early development of the zebrafish, Danio rerio. *Proteomics* **2006,** 6, (10), 3176-3188.

29. Wang, N.; MacKenzie, L.; De Souza, A. G.; Zhong, H.; Goss, G.; Li, L., Proteome profile of cytosolic component of zebrafish liver generated by LC-ESI MS/MS combined with trypsin digestion and microwave-assisted acid hydrolysis. *J. Proteome Res.* **2007,** 6, (1), 263-272.

30. Jekosch, K., The zebrafish genome project: sequence analysis and annotation. *Methods Cell Biol.* **2004,** 77, (Zebrafish: Genetics, Genomics, and Informatics), 225-239.

31. Barbazuk, W. B.; Korf, I.; Kadavi, C.; Heyen, J.; Tate, S.; Wun, E.; Bedell, J. A.; McPherson, J. D.; Johnson, S. L., The syntenic relationship of the zebrafish and human genomes. *Genome Res.* **2000,** 10, (9), 1351-1358.

32. Driever, W.; Stemple, D.; Schier, A.; Solnica-Krezel, L., Zebrafish: genetic tools for studying vertebrate development. *Trends Genet* **1994,** 10, (5), 152-159.

33. Strahle, U.; Blader, P., Early neurogenesis in the zebrafish embryo. *Faseb J* **1994,** 8, (10), 692-698.

34. Ungos, J.; Weinstein, B. M., Vascular development in the zebrafish. *Adv. Dev. Biol. (Amsterdam, Neth.)* **2008,** 18, (Cardiovascular Development), 301-332.

35. Baker, K.; Warren, K. S.; Yellen, G.; Fishman, M. C., Defective "pacemaker" current (Ih) in a zebrafish mutant with a slow heart rate. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** 94, (9), 4554-4559.

36. Poss, K. D., Getting to the heart of regeneration in zebrafish. *Semin. Cell Dev. Biol.* **2007,** 18, (1), 36-45.

37. Warren, K. S.; Baker, K.; Fishman, M. C., The slow mo mutation reduces pacemaker current and heart rate in adult zebrafish. *Am. J. Physiol.* **2001,** 281, (4, Pt. 2), H1711-H1719.

38. Warren, K. S.; Wu, J. C.; Pinet, F.; Fishman, M. C., The genetic basis of cardiac function: dissection by zebrafish (Danio rerio) screens. *Philos Trans R Soc Lond B Biol Sci* **2000,** 355, (1399), 939-944.

39. Xu, X.; Meiler, S. E.; Zhong, T. P.; Mohideen, M.; Crossley, D. A.; Burggren, W. W.; Fishman, M. C., Cardiomyopathy in zebrafish due to mutation in an alternatively spliced exon of titin. *Nat. Genet.* **2002,** 30, (2), 205-209.

40. Bingham, C.; Bulman, M. P.; Ellard, S.; Allen, L. I. S.; Lipkin, G. W.; van't Hoff, W. G.; Woolf, A. S.; Rizzoni, G.; Novelli, G.; Nicholls, A. J.; Hattersley, A. T., Mutations in the hepatocyte nuclear factor-1β gene are associated with familial hypoplastic glomerulocystic kidney disease. *Am. J. Hum. Genet.* **2001,** 68, (1), 219-224.

41. Drummond, I. A.; Majumdar, A.; Hentschel, H.; Elger, M.; Solnica-Krezel, L.; Schier, A. F.; Neuhauss, S. C. F.; Stemple, D. L.; Zwartkruis, F.; Rangini, Z.; Driever, W.; Fishman, M. C., Early development of the zebrafish pronephros and analysis of mutations affecting pronephric function. *Development (Cambridge, U. K.)* **1998,** 125, (23), 4655-4667.

42. Gabow, P. A., Autosomal dominant polycystic kidney disease. *N Engl J Med* **1993,** 329, (5), 332-342.

43. Hentschel, D. M.; Park, K. M.; Cilenti, L.; Zervos, A. S.; Drummond, I.; Bonventre, J. V., Acute renal failure in zebrafish: A novel system to study a complex disease. *Am. J. Physiol.* **2005,** 288, (5, Pt. 2), F923-F929.

44. Serluca, F. C.; Fishman, M. C., Pre-pattern in the pronephric kidney field of zebrafish. *Development (Cambridge, U. K.)* **2001,** 128, (12), 2233-2241.

45. Sun, Z.; Hopkins, N., vhnf1, the MODY5 and familial GCKD-associated gene, regulates regional specification of the zebrafish gut, pronephros, and hindbrain. *Genes Dev.* **2001,** 15, (23), 3217-3229.

46. Ardouin, O.; Legouis, R.; Fasano, L.; David-Watine, B.; Korn, H.; Hardelin, J.-P.; Petit, C., Characterization of the two zebrafish orthologues of the KAL-1 gene underlying X chromosome-linked Kallmann syndrome. *Mech. Dev.* **2000,** 90, (1), 89-94.

47. Guo, S., Linking genes to brain, behavior and neurological diseases: what can we learn from zebrafish? *Genes, Brain Behav.* **2004,** 3, (2), 63-74.

48. Karlovich, C. A.; John, R. M.; Ramirez, L.; Stainier, D. Y. R.; Myers, R. M., Characterization of the Huntington's disease (HD) gene homolog in the zebrafish Danio rerio. *Gene* **1998,** 217, (1-2), 117-125.

49. Leimer, U.; Lun, K.; Romig, H.; Walter, J.; Gruenberg, J.; Brand, M.; Haass, C., Zebrafish (Danio rerio) presenilin promotes aberrant amyloid β-peptide production and requires a critical aspartate residue for its function in amyloidogenesis. *Biochemistry* **1999,** 38, (41), 13602-13609.

50. Wulf, P.; Bernhardt, R. R.; Suter, U., Characterization of peripheral myelin protein 22 in zebrafish (zPMP22) suggests an early role in the development of the peripheral nervous system. *J. Neurosci. Res.* **1999,** 57, (4), 467-478.

51. Love, D. R.; Pichler, F. B.; Dodd, A.; Copp, B. R.; Greenwood, D. R., Technology for high-throughput screens: the present and future using zebrafish. *Curr. Opin. Biotechnol.* **2004,** 15, (6), 564-571.

52. Ma, C.; Parng, C.; Seng, W. L.; Zhang, C.; Willett, C.; McGrath, P., Zebrafish - an in vivo model for drug screening. *Innovations Pharm. Technol.* **2003,** 03, (12), 38, 40, 42, 44-45.

53. Zon, L. I.; Peterson, R. T., In vivo drug discovery in the zebrafish. *Nat. Rev. Drug Discovery* **2005,** 4, (1), 35-44.

54. Laurent, P.; Dunel, S., Morphology of gill epithelia in fish. *Am J Physiol* **1980,** 238, (3), R147-159.

55. Rombough, P. J.; Moroz, B. M., The scaling and potential importance of cutaneous and branchial surfaces in respiratory gas exchange in larval and juvenile walleye Stizostedion vitreum. *J. Exp. Biol.* **1997,** 200, (18), 2459-2468.

56. Wood, C. M.; Kelly, S. P.; Zhou, B.; Fletcher, M.; O'Donnell, M.; Eletti, B.; Part, P., Cultured gill epithelia as models for the freshwater fish gill. *Biochim. Biophys. Acta, Biomembr.* **2002,** 1566, (1-2), 72-83.

57. Krstulovic, A. M.; Matzura, C. T.; Bertani-Dziedzic, L.; Cerqueira, S.; Gitlow, S. E., Endogenous levels of free and conjugated urinary 3-methoxy-4-hydroxyphenylethyleneglycol in control subjects and patients with pheochromocytoma determined by reversed-phase liquid chromatography with electrochemical detection. *Clin. Chim. Acta* **1980,** 103, (1), 109-116.

58. Maggs, J. L.; Grimmer, S. F. M.; Orme, M. L.; Breckenridge, A. M.; Park, B. K.; Gilmore, I. T., The biliary and urinary metabolites of [3H]17α-ethynylestradiol in women. *Xenobiotica* **1983,** 13, (7), 421-431.

59. Williams, M. C.; Helton, E. D.; Goldzieher, J. W., Urinary metabolites of 17alpha-ethynylestradiol-9alpha,11xi-3H in women. Chromatographic profiling and identification of ethynyl and nonethynyl compounds. *Steroids* **1975,** 25, (2), 229-246.

60. Futcher, B.; Latter, G. I.; Monardo, P.; McLaughlin, C. S.; Garrels, J. I., A sampling of the yeast proteome. *Mol. Cell. Biol.* **1999,** 19, (11), 7357-7368.

61. Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R., Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **1999,** 19, (3), 1720-1730.

62. Matsumoto, M.; Hatakeyama, S.; Oyamada, K.; Oda, Y.; Nishimura, T.; Nakayama, K. I., Large-scale analysis of the human ubiquitin-related proteome. *Proteomics* **2005,** 5, (16), 4145-4151.

63. Griffin, J. L., Metabonomics: NMR spectroscopy and pattern recognition analysis of body fluids and tissues for characterisation of xenobiotic toxicity and disease diagnosis. *Curr. Opin. Chem. Biol.* **2003,** 7, (5), 648-654.

64. Lindon, J. C.; Nicholson, J. K.; Holmes, E.; Everett, J. R., Metabonomics: metabolic processes studied by NMR spectroscopy of biofluids. *Concepts Magn. Reson.* **2000,** 12, (5), 289-320.

65. Nicholson, J. K.; Lindon, J. C.; Holmes, E., "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological

NMR spectroscopic data. *Xenobiotica* **1999,** 29, (11), 1181-1189.

66. Fiehn, O.; Kopka, J.; Dormann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L., Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **2000,** 18, (11), 1157-1161.

67. Jonsson, P.; Gullberg, J.; Nordstroem, A.; Kusano, M.; Kowalczyk, M.; Sjoestroem, M.; Moritz, T., A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS. *Anal. Chem.* **2004,** 76, (6), 1738-1745.

68. Idborg-Bjoerkman, H.; Edlund, P.-O.; Kvalheim, O. M.; Schuppe-Koistinen, I.; Jacobsson, S. P., Screening of biomarkers in rat urine using LC/electrospray ionization-MS and two-way data analysis. *Anal. Chem.* **2003,** 75, (18), 4784-4792.

69. Plumb, R. S.; Stumpf, C. L.; Gorenstein, M. V.; Castro-Perez, J. M.; Dear, G. J.; Anthony, M.; Sweatman, B. C.; Connor, S. C.; Haselden, J. N., Metabonomics: the use of electrospray mass spectrometry coupled to reversed-phase liquid chromatography shows potential for the screening of rat urine in drug development. *Rapid Commun. Mass Spectrom.* **2002,** 16, (20), 1991-1996.

70. Plumb, R. S.; Stumpf, C. L.; Granger, J. H.; Castro-perez, J.; Haselden, J. N.; Dear, G. J., Use of liquid chromatography/time-of-flight mass spectrometry and multivariate statistical analysis shows promise for the detection of drug metabolites in biological fluids. *Rapid Commun. Mass Spectrom.* **2003,** 17, (23), 2632-2638.

71. Van, Q. N.; Klose, J. R.; Lucas, D. A.; Prieto, D. A.; Luke, B.; Collins, J.; Burt, S. K.; Chmurny, G. N.; Issaq, H. J.; Conrads, T. P.; Veenstra, T. D.; Keay, S. K., The use of urine proteomic and metabonomic patterns for the diagnosis of interstitial cystitis and bacterial cystitis. *Dis. Markers* **2003,** 19, (4,5), 169-183.

72. Yang, J.; Xu, G.; Zheng, Y.; Kong, H.; Pang, T.; Lu, S.; Yang, Q., Diagnosis of liver cancer using HPLC-based metabonomics avoiding false-positive result from hepatitis and hepatocirrhosis diseases. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2004,** 813, (1-2), 59-65.

73. Yang, J.; Xu, G.; Hong, Q.; Liebich, H. M.; Lutz, K.; Schmuelling, R. M.; Wahl, H. G., Discrimination of Type 2 diabetic patients from healthy controls by using metabonomics method based on their serum fatty acid profiles. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2004,** 813, (1-2), 53-58.

74. Constantinou, M. A.; Papakonstantinou, E.; Benaki, D.; Spraul, M.; Shulpis, K.; Koupparis, M. A.; Mikros, E., Application of nuclear magnetic resonance spectroscopy combined with principal component analysis in detecting inborn errors of metabolism using blood spots: a metabonomic approach. *Anal. Chim. Acta* **2004,** 511, (2), 303-312.

Chapter 2

Large-Scale Proteome Profile of the Zebrafish (*Danio rerio*) Gill for Physiological and Biomarker Discovery Studies*

2.1 Introduction

Establishing the complete proteome of an organism provides a strong base from which hypotheses concerning physiological adaptations or responses to stress can be designed. Proteomic profiling is particularly useful for identifying new biomarker proteins as it describes the biological characteristics underlying a particular phenotype.[1-4] It can also be applied to study the effects of emerging contaminants whose bioactive mechanisms may be complex and not easily predicted using standard toxicological methods.[5] Due to the complexity of the physiological stress response, single proteins are rarely specific or sensitive enough to be used reliably as biomarkers.[3] By studying entire biochemical pathways, proteomics can delineate characteristic patterns of change specific to a stimulus. In clinical settings, using a "panel" of biomarker proteins increases the specificity and accuracy of disease diagnosis relative to using an individual protein alone.[6, 7]

---

*A version of this chapter was published as: A. G. De Souza[¥], T. J. MacCormack[¥], N. Wang, L. Li, and G. G. Goss, **2009**, "Large scale proteome profile of the zebrafish (*Danio rerio*) gill for physiological and biomarker discovery studies", *Zebrafish*, **6**, 229-238. [¥]Authors contributed equally to this work.

While genomic studies provide valuable information, the transcriptome does not account for the post-transcriptional and post-translational regulation of protein expression. In many cases, there is poor correspondence between changes in transcript level and protein expression.[8, 9] The proteome accounts for these complex regulatory processes and potentially provides a much more accurate snapshot of an organism's physiological status. However, proteome profiling is much more challenging than transcriptome analysis and to date, technological restrictions have severely limited the number and type of proteins analyzed. Two-dimensional (2D) gel electrophoresis protein separation combined with tandem mass spectrometry (MS/MS) is the traditional method used to assess protein expression and proportional abundance changes between samples.[10, 11] The cumbersome nature of this method reduces the number of proteins addressed with most studies identifying less than 100 proteins.[2, 12, 13] In addition, membrane proteins tend to be under-represented in gel-based analyses because hydrophobic proteins may precipitate during isoelectric focusing and are not transferred to the subsequent SDS-PAGE separation.[14, 15]

Two-dimensional liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI MS/MS) shotgun methods typically identify more proteins than gel-based techniques and provide a more favourable separation environment for hydrophobic proteins.[15] Wang *et al.*[16] used LC-ESI MS/MS to identify over 1200 proteins from the cytosolic fraction of the zebrafish liver, increasing proteome coverage by more than an order of magnitude over previous studies of whole fish tissues.[17-21] Recently, close to 1400 proteins were identified

in zebrafish embryos using similar techniques.[22] These studies illustrate the power and sensitivity of LC-ESI MS/MS shotgun methods and show that a comprehensive proteome is within reach utilizing a combination of existing techniques.

The zebrafish (*Danio rerio*) is an excellent candidate for proteomic studies as it is a well-characterized vertebrate model widely used in a variety of disciplines including genetics, developmental biology and physiology.[23-25] The genome of the zebrafish is mostly sequenced and partially annotated, which facilitates proteomics analyses by allowing for the identification and characterization of proteins using existing databases. Fish are a valuable model system for investigating the impact of toxins on aquatic environments and the zebrafish is gaining popularity as a model for this discipline.[26-31] Fish require a large surface area for gas and ion exchange with their environment and these functions are largely accomplished by the gills.[32] The gills can make up greater than 50% of the total surface area of the animal[33] and provide a direct route for the uptake of contaminants.[34]

The goal of the current study was to develop and apply a 2D LC-ESI MS/MS shotgun technique for the analysis of the zebrafish gill and determine the proteome coverage achievable using current state-of-the-art technology. This work provides a baseline proteome of the zebrafish gill for use in physiological, developmental and toxicological studies. The analysis of the proteome focused on proteins relevant to the normal function of the gill and on biomarkers of toxicant and stress exposure. Using our method, we were able to identify and

partially characterize the most complete proteome of a fish tissue published to date. The data provided excellent coverage of predicted biochemical pathway components and identified numerous established and potential biomarkers. In this context, we also highlight a number of bioinformatics challenges that must be resolved in order to facilitate the widespread use of high-throughput methods for quantitative proteomic profiling.

2.2 Experimental

2.2.1 Chemicals and Reagents

All chemicals and reagents were purchased from Sigma-Aldrich (Oakville, ON, Canada) unless otherwise stated. Tricaine methanesulfonate (TMS) was purchased from Syndel Laboratories Ltd. (Vancouver, BC, Canada). The total protein extraction kit was purchased from Biochain Institute Inc. (Hayward, CA) and the bicinchoninic acid (BCA) assay kit was purchased from Pierce Biotechnology (Rockford, IL). Dithiothreitol (DTT), iodoacetamide (IAA), and sodium dodecyl sulphate (SDS) were purchased from BioRad (Mississauga, ON, Canada). Sequencing-grade modified trypsin and LC-MS grade acetonitrile, acetone, water, and methanol (MeOH) were purchased from Fisher Scientific Canada (Edmonton, AB, Canada).

2.2.2 Zebrafish Gill and Sample Preparation

A total of 24 zebrafish (12 male, 12 female, body mass 200-400 mg) were taken from the Zebrafish Breeding Facility at the University of Alberta

Biosciences Aquatics Facility. All fish were reared and maintained at 28 °C. Nacre strain (A/B background) zebrafish were used in this study.

To reduce the content of blood and plasma proteins in the gill tissue, the circulatory system was flushed with ice-cold phosphate buffered saline (PBS) containing 50 IU/mL sodium heparin. Perfusion cannulae consisted of a length of PE10 tubing with an 8 mm long tip fabricated from a 30G needle. The cutting edge of the needle was removed and the tip was polished. Fish were anaesthetized in a 1.0 g/L solution of TMS and positioned ventral-side up on a moistened sponge under a dissecting microscope. The body cavity was then opened and quickly rinsed with ice-cold PBS. The heart was exposed by blunt dissection and one ventricular wall was carefully pierced with a sharp 30G needle. The cannula was inserted and held in place by clamping the ventricle wall around it using blunt forceps. The heart was subsequently perfused at 300 μL/min with ice-cold PBS containing 50 IU/mL sodium heparin until gill tissue was visibly cleared of blood (1-5 min).

A full workflow is outlined in Figure 2.1. Following perfusion, whole gill baskets were excised and placed in a flat-bottomed microcentrifuge tube on dry ice and proteins extracted according to the protocol provided by the manufacturer (Biochain Institute Inc., Hayward, CA). A modification to the protocol involved the addition of 2 mM phenylmethylsulfonyl fluoride (dissolved in EtOH) to the extraction buffer immediately prior to solubilization. Each gill was homogenized on ice using a disposable pestle, rotated on an end-over-end rotator (4 °C, 20 min), centrifuged (20 800 g, 4 °C, 20 min), and the supernatant collected and

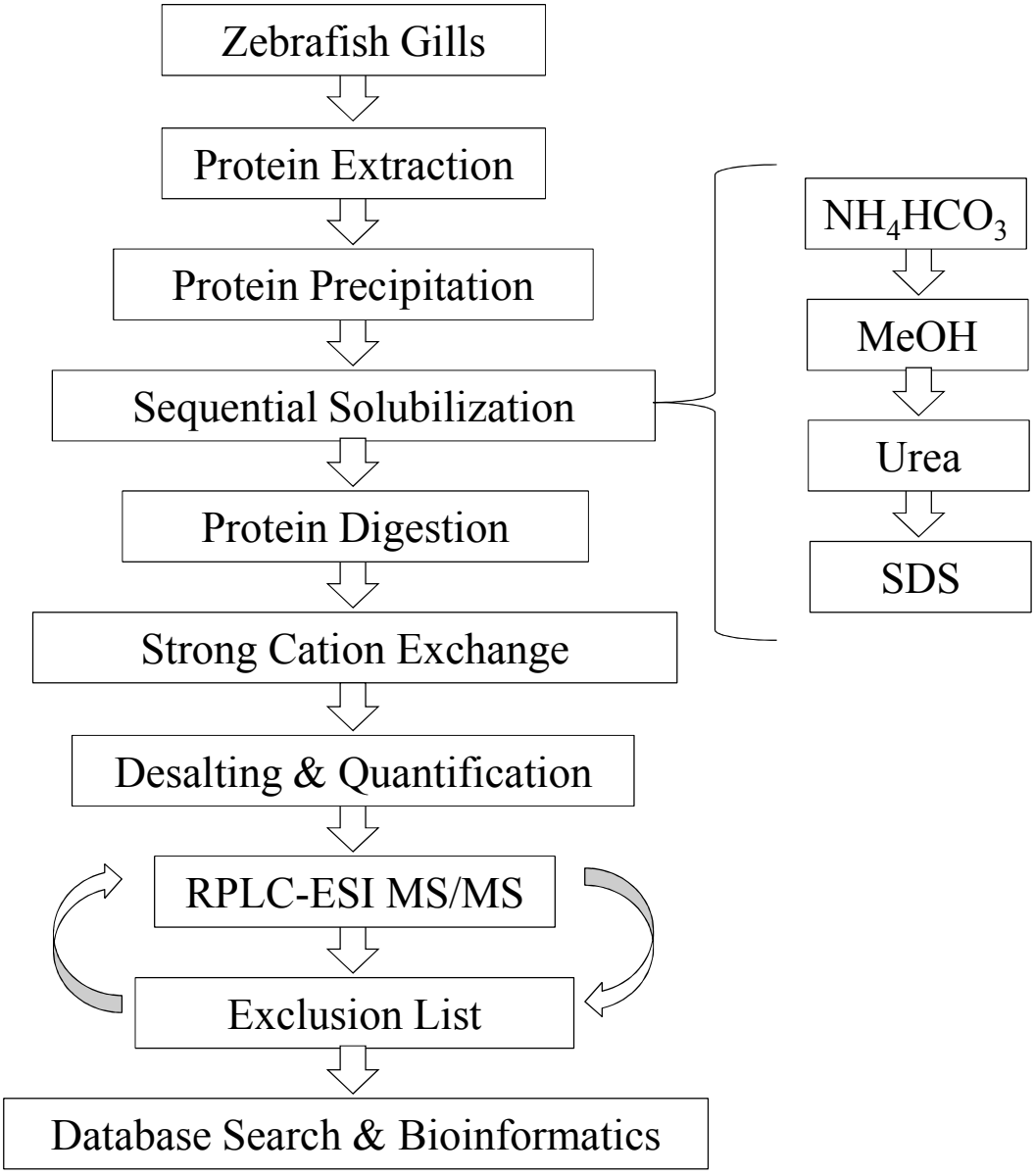

Figure 2.1. Workflow for proteome analysis of zebrafish gill.

stored at -80 °C. A BCA assay was conducted to determine the gill protein concentration using bovine serum albumin as the standard.

2.2.3 Acetone Precipitation and In-Solution Digestion

The individual gill protein extracts were pooled to generate one sample. Proteins were reduced with 900 mM DTT for one hour at 37 °C, cooled at room temperature, and alkylated with 2.2 molar equivalents of 450 mM IAA for 1 h at room temperature in the dark. The concentrations of DTT and IAA were chosen to reduce sample dilution upon reduction and alkylation. For every 5 µg of protein, 0.225 µmol of DTT and 0.500 µmol of IAA were used to reduce and alkylate the disulfide bonds, respectively. Proteins were then precipitated with four times the volume of acetone (pre-cooled at -80 °C), kept at -20 °C overnight, centrifuged (20 800 g, 4 °C, 10 min), and the supernatant removed and discarded.

The protein pellet was partially solubilized in 100 mM NH_4HCO_3, sonicated in an ice bath (1 min), vortexed (1 h, 4 °C), and centrifuged (20 800 g, 4 °C, 10 min). The supernatant was collected and diluted to 50 mM NH_4HCO_3. Solubilization of the protein in the remaining pellet proceeded using the following solvents sequentially: 60% MeOH, 8 M urea, and 1% SDS. A BCA assay was conducted to determine the protein concentration in each protein sample. The urea and SDS samples were diluted to 1 M and 0.05%, respectively, prior to digestion with trypsin. The trypsin-to-protein ratio was 1:25 (w/w), and digestion occurred at 37 °C overnight.

2.2.4 Strong Cation Exchange Chromatography

Each of the 4 peptide samples was further separated on a PolySULFOETHYL A column (PolyLC, Columbia, MD, 5 μm, 300 Å, 2.1 x 250 mm) by strong cation exchange (SCX) on an Agilent 1100 HPLC system (Palo Alto, CA). Mobile Phase A consisted of 10 mM KH_2PO_4 (pH 2.76) and Mobile Phase B consisted of 10 mM KH_2PO_4 and 0.5 M KCl (pH 2.76). The elution gradient was 0% B for 7 min, 0-6% B from 7 to 8 min, 6-28% B from 8 to 36 min, 28-40% B from 36 to 44 min, 40-60% B from 44 to 49 min, 60-100% B from 49 to 53 min, 100% B from 53 to 58 min, 100-0% B from 58 to 60 min, and held at 0% B for 10 min at a flow rate of 0.200 mL/min. Fractions were collected at 1 min intervals from 16 min to 70 min.

2.2.5 Peptide Desalting and Quantitation

Prior to LC-ESI MS/MS, the samples collected from the SCX were subjected to salt removal on a Polaris C18A column (Palo Alto, CA, 3 μm, 300 Å, 4.6 x 50 mm) on an Agilent 1100 HPLC system with a UV detector. In addition to salt removal, simultaneous peptide quantitation was performed based on the area of the peptide peak at 214 nm.[35] Mobile Phase A consisted of 0.1% trifluoroacetic acid (TFA) in water and Mobile Phase B consisted of 0.1% TFA in acetonitrile. The elution gradient was 2.5% B for 5.50 min, 2.5-85% B from 5.50 to 5.51 min, 85% B from 5.51 to 15.50 min, 85-2.5% B from 15.50 to 15.51 min, and held at 2.5% B from 15.51 to 30.00 min at a flow rate of 1.000 mL/min. Fraction collection occurred at 7.50 min for a total of 1.10 min. Adjacent

fractions from the SCX fractionation were pooled if the total amount of peptides in individual fractions was less than 1 μg.

2.2.6 LC-ESI QToF MS and MS/MS Analysis

The desalted SCX peptide samples were analyzed using a Q-Tof Premier™ mass spectrometer equipped with a nanoACQUITY® UPLC system (Waters, Milford, MA). Approximately 1 μg of sample was loaded onto an Atlantis dC18 column (Waters, 3 μm, 100 Å, 75 μm x 150 mm). Mobile Phase A consisted of 0.1% formic acid in water and Mobile Phase B consisted of 0.1% formic acid in acetonitrile. The elution gradient was 2-7% B from 0 to 2 min, 7-20% B from 2 to 85 min, 20-30% B from 85 to 110 min, 30-45% B from 110 to 115 min, 45-90% B from 115 to 120 min, held at 90% B for 5 min, and 90-2% B for 5 min at 250 nL/min. A precursor ion exclusion (PIE) strategy in LC MS/MS was used where ions whose acquired MS/MS spectra resulted in positive peptide identification were excluded from MS/MS acquisition in the subsequent run of the adjacent SCX sample.[36] This strategy improves the overall peptide identification efficiency, as well as enables the sequencing of low abundance peptides.

2.2.7 Protein Identification from MS/MS Data

Raw LC MS/MS data were lock-mass corrected, de-isotoped, and converted to peak list files by ProteinLynx Global Server 2.2.5 (Waters). Peptide sequences were identified by automated database searches of peak list files generated from each LC-ESI MS/MS run using the Mascot search program (Matrix Science, London, UK). Mascot searches were performed individually for the SCX fractions of the tryptic digest generated from each of the solvents used to

dissolve the proteome sample. These individual search results were then merged for each solubilization method. A final merge of all results from four solvents generated the final protein list.

For each LC-ESI MS/MS run, database searches were restricted to *Danio rerio* in the NCBInr database (searched in August 2007 with a database containing 4815286 sequences and 1665828716 residues). The following search parameters were selected for all database searches: enzyme, trypsin; maximum missed cleavages, 1; fixed modifications, carbamidomethyl (C); variable modifications, N-Acetyl (Protein), oxidation (M), pyro-glu (N-term E), pyro-glu (N-term Q); peptide tolerance, ± 30 ppm; fragment mass tolerance, ± 0.2 Da; peptide charge, 1+, 2+, 3+; instrument type, ESI-QUAD-TOF modified to include immonium ions and a-series ions as possible fragmentations. Two additional variable modifications were selected for the urea-solubilized samples: carbamyl (K), carbamyl (N-term). The search results, including protein names, accession IDs, molecular mass, unique peptide sequences, ion score, Mascot threshold score for identity, calculated molecular mass of the peptide, and the difference (error) between the experimental and calculated masses were extracted to Excel files using in-house software. The identified peptides with scores lower than the Mascot threshold score for identity at a confidence level of 95% were subsequently removed from the protein list. Redundant peptides identified for different proteins were deleted, and redundant proteins identified under the same gene name but different accession IDs were also deleted. The final unique protein or peptide list was generated by merging all the protein or peptide lists according

to the following rules: only unique proteins (under unique gene names) and peptides with the highest scores were retained; each peptide was associated to one unique protein; only the first hit within each identified protein group was kept as a representative protein. All redundant peptides with lower identification scores were deleted, in addition to redundant proteins with either lower scores or lower number of peptides.

The target-decoy search strategy was applied to assess the false positive peptide matching rate in our analysis by searching the MS/MS spectra against the forward and reversed zebrafish proteome sequences.[37, 38] This approach involved re-searching the matched spectra from the Mascot database search using the forward or correct proteome sequences against the reversed proteome sequence or decoy. The decoy peptide matches with scores above the threshold scores at the 95% confidence level were then compared to those in the forward sequence search. If the score of a MS/MS spectrum matched with a decoy peptide was equal to or higher than that of the same spectrum matched with a correct peptide, a false positive match was registered. The false positive matching rate was calculated using the equation: $\frac{\mathbf{2}(\boldsymbol{n_{reversed}})}{\left(\boldsymbol{n_{forward}}+\boldsymbol{n_{reversed}}\right)}$, where $n_{reversed}$ and $n_{forward}$ are the number of matches from the reversed (decoy) and forward (correct) sequences, respectively.

2.2.8 Protein Annotation

Figure 2.2 illustrates the general workflow used for the analysis of the final list of identified proteins generated by the methods described above. Proteins identified with only an alphanumerical code and the descriptors

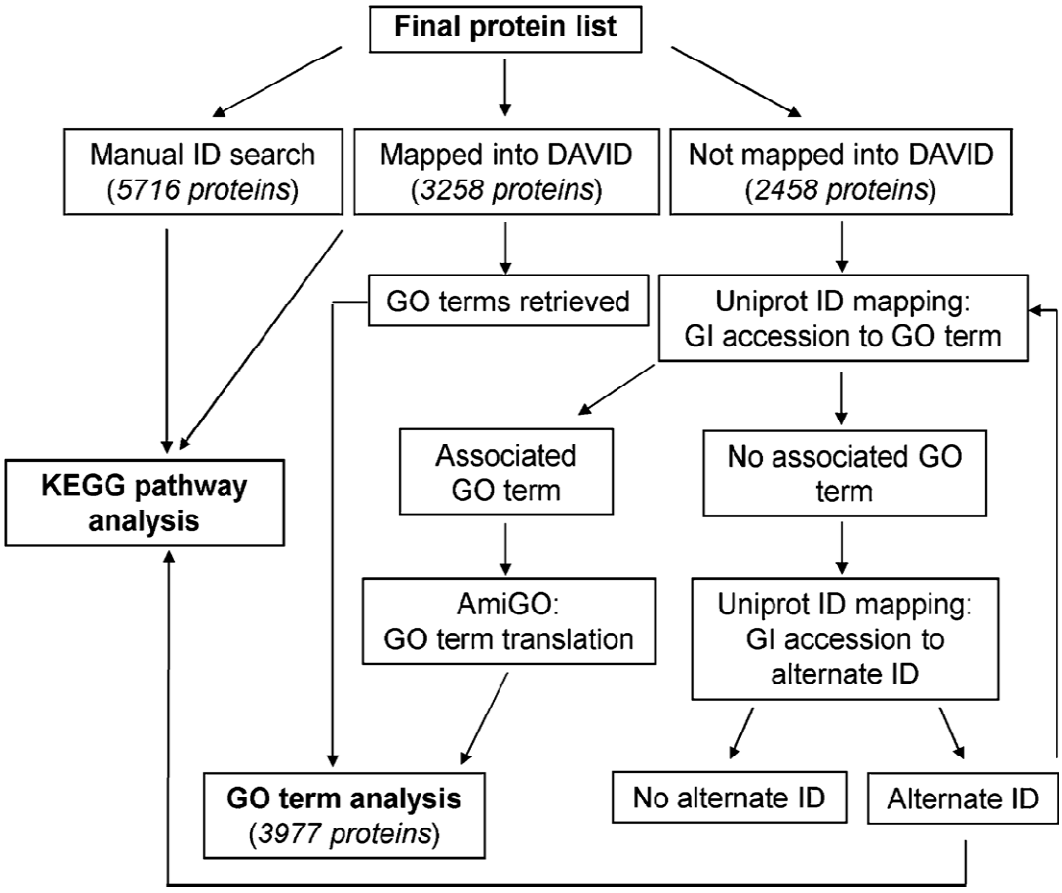

Figure 2.2. Workflow for bioinformatics analysis of zebrafish gill.

"hypothetical", "novel", and/or "predicted" were further analyzed using the Universal Protein Resource[39] ID mapping and retrieve tools to identify alternate, biologically relevant protein names for subsequent categorization and pathway analysis.

Gene Ontology (GO) terms[40] associated with the identified proteins were retrieved using The Database for Annotation, Visualization, and Integrated Discovery (DAVID) functional annotation tool.[41] For proteins that DAVID was unable to retrieve any associated GO terms, GI accession numbers were mapped to GO terms using the UniProt ID mapping tool. When necessary, GO codes were translated using the AmiGO tool on the Gene Ontology website.

In some cases, proteins were categorized according to their participation in specific biochemical pathways using the Kyoto Encyclopedia of Genes and Genomes (KEGG) PATHWAY database[42], the DAVID functional annotation tool, and manual analysis based on GO terms and primary literature searches. For pathway coverage data, the gill proteome was manually analyzed against KEGG pathways predicted for *Danio rerio*.

2.3 Results

Using the advanced LC-ESI-MS/MS method developed in our lab, we have identified 5716 proteins expressed in the zebrafish gill (see Supplementary Information 2.1). Raw data is available upon request from the authors and, to increase data accessibility and provide interactive data processing of the raw and metadata generated from one or more organs of the zebrafish, efforts are currently

underway to establish an integrated, publicly available database of the zebrafish proteome. The false positive discovery rate was estimated at 1.34% using a previously described method.[37, 38] Although 49% of the proteins identified were from single-peptide-matches, the confidence of protein identification based on single-peptide matches is still high, as judged by the low false positive matching rate determined from the dataset. It should be noted that, compared to a low resolution mass analyzer such as an ion trap for recording collision-induced dissociation (CID) mass spectra, the Q-Tof instrument generates MS/MS spectra with high resolution and high mass measurement accuracy which contributes to a low false positive matching rate.

To improve proteome coverage, we employed several novel approaches, including sequential solubilization of proteins in the sample for simplifying the proteome prior to 2D-LC MS/MS analysis. Tissue samples contain a complex mixture of proteins with varying solubilities. Exploiting a number of solvents with different properties after acetone precipitation can fractionate the proteome sample into sub-proteomes and allow a more thorough analysis by 2D-LC MS/MS. In a previous study from our lab,[16] similar methods were used to expand proteome coverage to over 1200 proteins in just the cytosolic fraction of the zebrafish liver. Figure 2.3 shows the distribution of the number of proteins identified in each of the four fractions generated by the sequential solubilization protocol. Eight hundred three (14%), 104 (2%), 1297 (23%), and 860 (15%) proteins were exclusively found in the NH_4HCO_3, MeOH, urea, and SDS fractions, respectively. This data clearly shows that when dealing with a complex

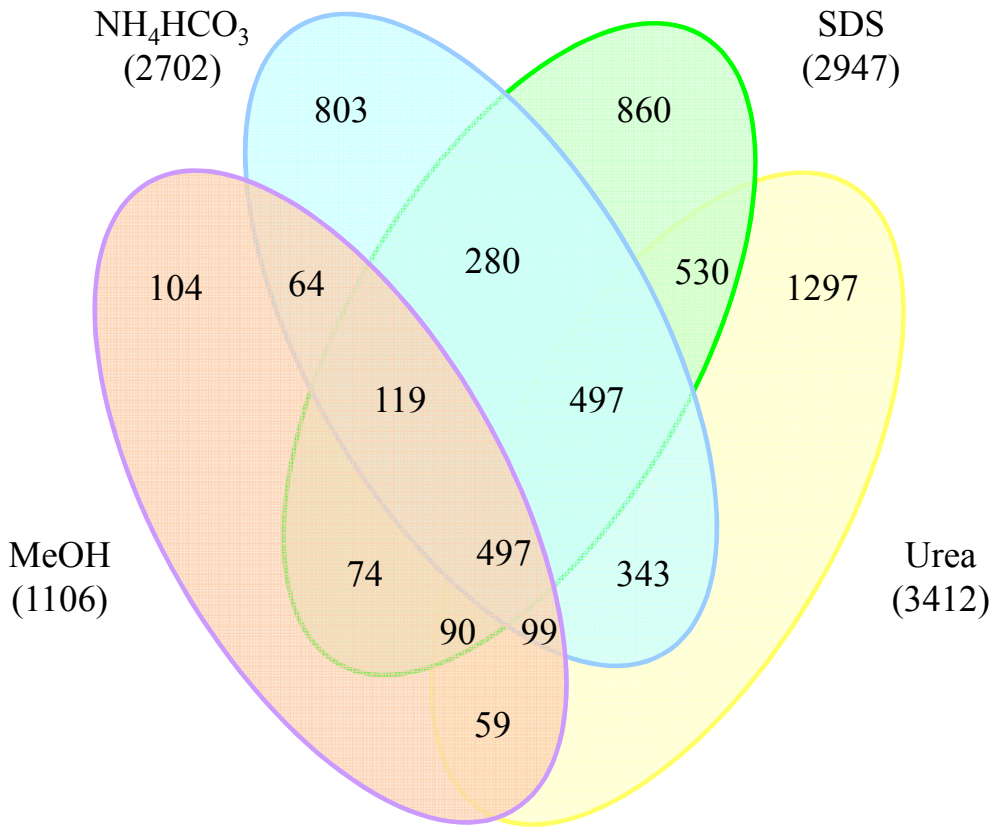

Figure 2.3. A comparison of the protein identification results from the four protein solubilization techniques. The numbers in brackets indicate the total number of proteins identified in each fraction.

sample, sequential solubilization significantly increases proteome coverage compared to a single solubilization protocol. Previously published proteomic analyses on fish tissue typically identified 10-100 proteins. Only two studies have identified more than 1000 proteins (1384 proteins[22] and 1204 proteins[16]) and both used 2D LC-ESI MS/MS. Our work clearly demonstrates the potential of this technique for more comprehensive coverage.

Sequential solubilization is a powerful technique to fractionate the protein sample, but steps must be taken during the subsequent analysis to ensure that protein identification is accurate. Urea and SDS are effective for solubilizing hydrophobic proteins,[43, 44] but these solutes can introduce problems in shotgun methods based on proteome digestion and LC MS/MS analysis.[45] Urea efficiently solubilizes membrane proteins but the formation of cyanic acid can be problematic because it reacts with primary amines. This side reaction can be accounted for in database searches by selecting the variable modifications of carbamylation at lysines and N-termini of proteins. SDS molecules can strongly bind to some peptides and suppress peptide signals in MS analysis. To reduce SDS interference, we used SCX prior to reversed phase separation and introduction of the sample into the mass spectrometer. Using these additional sample preparation techniques, we were able to increase proteome coverage.

Proteome coverage was also improved through the use of an optimal PIE technique developed in our lab.[36] Although the initial protein sample was fractionated *via* solubilization, SCX, and RP prior to analysis, it is still extremely complex and not all peptides are able to be analyzed by MS/MS. The PIE

method involves generating a list of positively identified peptides from one SCX fraction and then excluding these ions from MS/MS spectral acquisition in the subsequent SCX fractions. In other studies, the application of the PIE method improved the likelihood of scanning less abundant peptides and increased proteome coverage by 45% or more.[36]

Protein identification was based on the sequence match of one or more peptides. Protein modifications are difficult to detect, as sequence coverage by peptides is usually low (in this work, only a few protein isoforms were detected where different regions of peptide sequences were identified). The zebrafish genome contains approximately 17330 genes that can potentially code for proteins.[46] The identification of 5716 unique proteins therefore represents 36% of the total potential proteome expressed in the gill alone. Given that only a single tissue was investigated under steady-state conditions, we feel it is reasonable to suggest that this represents a significant proportion of the total number of proteins present in the gill at the time of sampling.

A significant issue encountered in the subsequent proteome analysis was that 44% of the proteins identified in the gill were described only as hypothetical, predicted, novel, or by an alphanumeric code (e.g., Zebrafish Genome Collection 'ZGC' proteins). The lack of descriptive names made it difficult to classify these proteins in a meaningful biological context, particularly in the absence of associated GO terms (see below). To resolve this issue and improve the biological analysis of the proteome, we manually searched the available databases to identify more relevant names for the 2139 "hypothetical" proteins. Alternate

names were retrieved for 1105 of these proteins with the majority of these described only as novel proteins or by ZGC numbers. Similar searches were performed for "novel" and ZGC protein names and a number of more biologically relevant descriptors were retrieved for use in the pathway analysis described below.

Some degree of annotation information was retrieved for approximately 3977 proteins. The DAVID functional annotation tool[41] was able to retrieve GO terms for roughly 57% of the identified gill proteins, with the remaining annotations found using the UniProt ID mapping tool. The paucity of information for the other approximately 1739 proteins highlights the urgent need for advancements in annotation and bioinformatics techniques to facilitate large-scale proteomic analyses of this type.

Despite this gap in knowledge, the proteome exhibited excellent coverage of a number of biochemical pathways relevant to the function of the gill (Table 2.1) as well as numerous established biomarkers (Table 2.2). We utilized the KEGG database to examine 12 pathways important to the physiological role of the zebrafish gill and coverage ranged from 16-88% (Table 2.1). In several cases, we identified multiple proteins matching a single component of a particular pathway. An example of the extent of coverage we were able to achieve is given for glycolysis (Figure 2.4) where we identified multiple proteins for many steps in the pathway.

Table 2.1. Coverage of selected biochemical pathways by proteins expressed in the zebrafish gill.

Pathway	# Predicted	# Identified	% Coverage
Citrate cycle (TCA cycle)	16	14	88
Proteasome	25	22	88
Ribosome (small subunit)	27	23	85
Ribosome (large subunit)	28	22	79
Pyruvate metabolism	16	12	75
Glycolysis	23	17	74
Metabolism of xenobiotics by cytochrome p450	11	7	64
Nitrogen metabolism: reduction and fixation	8	5	63
Fatty acid elongation in mitochondria	5	3	60
Starch and sucrose metabolism	10	6	60
Urea cycle and metabolism of amino groups	15	9	60
Fatty acid metabolism	14	8	57
Oxidative phosphorylation	79	37	47
Ubiquitin-mediated proteolysis	76	12	16

Pathway components were predicted by the Kyoto Encyclopedia of Genes and Genomes database.
TCA = tricarboxylic acid

Table 2.2. Biomarker proteins identified in the zebrafish gill proteome

Protein Name	Protein Function
Acyl-coenzyme A dehydrogenase family, member 8	2-Methylnaphthalene degradation
Alcohol dehydrogenase 5	2-Methylnaphthalene degradation
	Methane metabolism
	Metabolism of xenobiotics by cytochrome p450
Glutathione s-transferase M	Metabolism of xenobiotics by cytochrome p450
Glutathione s-transferase PI	Metabolism of xenobiotics by cytochrome p450
Cytochrome p450 2AD3	Metabolism of xenobiotics by cytochrome p450
Cytochrome p450 family 2, subfamily J, polypeptide 24	Metabolism of xenobiotics by cytochrome p450
Cytochrome p450 family 3, subfamily C, polypeptide 1	Metabolism of xenobiotics by cytochrome p450
Novel protein similar to cytochrome p450 family 2, subfamily j	Metabolism of xenobiotics by cytochrome p450
NADPH-cytochrome p450 oxidoreductase	Metabolism of xenobiotics by cytochrome p450
ZGC:66393 (UDP glucuronosyltransferase 1 family A, B)	Metabolism of xenobiotics by cytochrome p450
Predicted: epoxide hydrolase 1, microsomal, partial	Metabolism of xenobiotics by cytochrome p450
Hydroxysteroid (17-beta) dehydrogenase 4	Caprolactam degradation
Sirtuin 2 (silent mating type information regulation 2, homolog)	Caprolactam degradation

NADPH, nicotinamide adenine dinucleotide phosphate; ZGC, Zebrafish Genome Collection; UDP, uridine diphosphate

Table 2.2. Biomarker proteins identified in the zebrafish gill proteome

Protein Name	Protein Function
Enoyl coenzyme A hydratase, short chain, 1, mitochondrial	Caprolactam degradation
	Benzoate degradation via coA ligation
Acetyl-coA acetyltransferase 2	Benzoate degradation via coA ligation
Manganese superoxide dismutase	Response to oxidative stress
Predicted: similar to extracellular superoxide dismutase	Response to oxidative stress
Superoxide dismutase 1, soluble	Response to oxidative stress
Homogentisate 1,2-dioxygenase	Styrene degradation
Adenosine deaminase, RNA-specific	Atrazine degradation
Metallothionein 2	Response to methylmercury
Arsenate resistance protein 2	Response to arsenic
Predicted: similar to heat shock cognate 70 kDa protein	Response to cadmium ion,
	Response to xenobiotic stimulus
	Response to heat
Vitellogenin 1	Response to endocrine disrupting compound
Vitellogenin 2	Response to endocrine disrupting compound
Vitellogenin 3 precursor	Response to endocrine disrupting compound

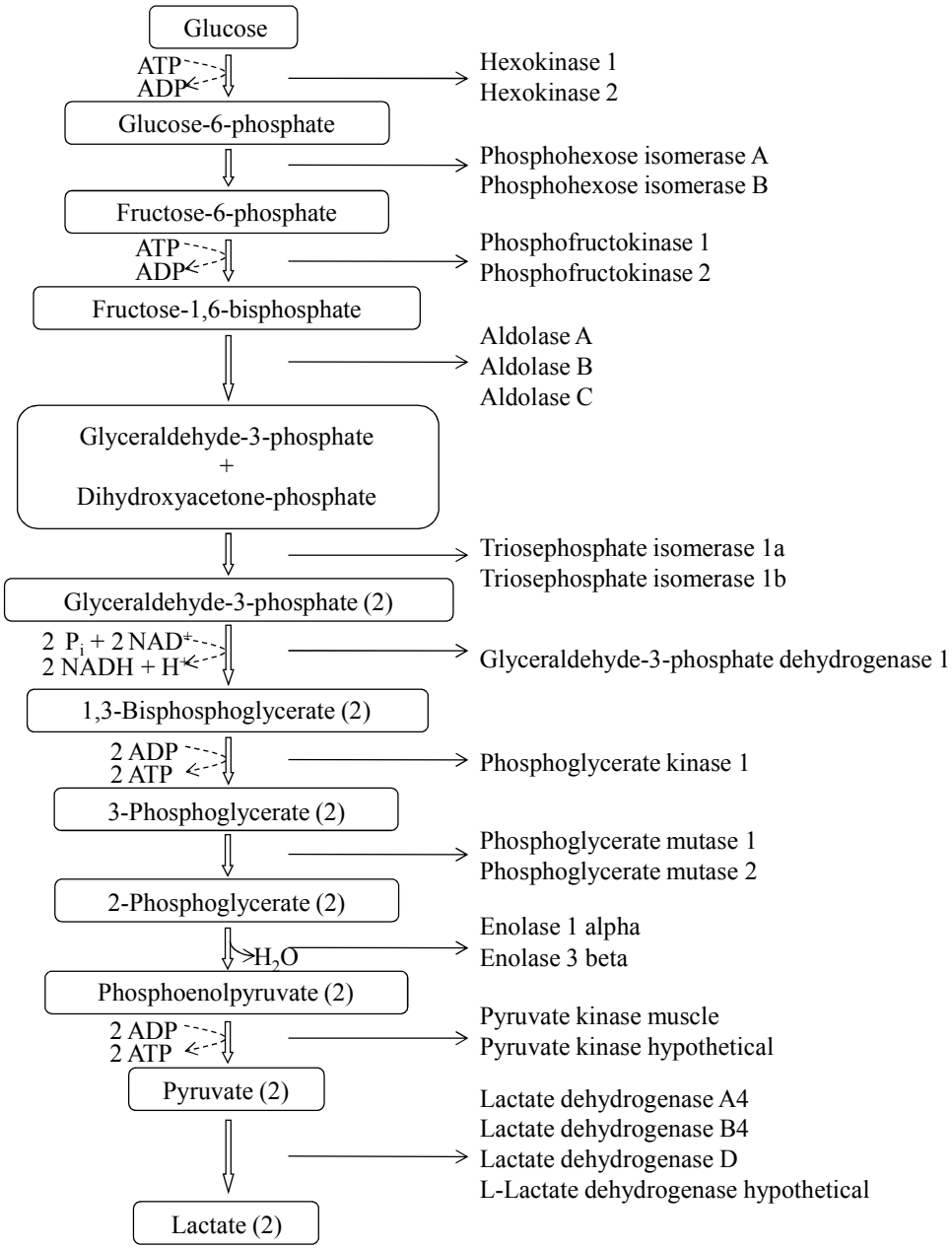

Figure 2.4. Zebrafish gill proteome coverage of the enzymes involved in the metabolism of glucose to lactate via glycolysis and fermentation.

2.4 Discussion

A major drawback of transcriptomic studies is the potentially weak correspondence between transcript levels and protein expression.[9] The expression of some proteins can vary more than 20-fold in the absence of changes in mRNA levels and similar changes in transcript levels can occur with no effect on protein expression.[47] The expression and function of proteins determines the biology of an organism so developing a method for analyzing a complete proteome is particularly valuable for physiological studies.

The extensive level of proteome coverage we have established in this study should greatly facilitate subsequent physiological, toxicological, and comparative studies of the fish gill. Unlike genomic studies, we clearly identify which proteins are expressed in the zebrafish gill under control conditions and provide important baseline information. With this database, we can now design specific hypotheses to investigate the function of single proteins, protein-protein interactions, or whole biochemical pathways. To emphasize the utility and relevance of proteomic profiling to these types of studies, the proteome was analyzed based upon established physiological roles of the fish gill and the biochemical pathways supporting these functions.

2.4.1 Respiration and Ion and Acid-Base Homeostasis

The gill is the predominant site of oxygen uptake in fish and it contains a population of cells specialized for carrying out this function.[32] Approximately 14 proteins with GO terms related to oxygen binding, oxygen transport, or reactive oxygen species metabolism were identified in the zebrafish gill proteome (see

Supplementary Information 2.2). These consisted of various alpha and beta hemoglobin subunits and myoglobin. Although all gills were fully flushed prior to analysis, it is likely that the high sensitivity of our methods identified proteins originating from trace amounts of blood in the tissue.

Freshwater fish maintain a high internal osmolarity relative to their environment and the gill has a substantial architecture of ion regulatory machinery to sustain that gradient.[48] Zebrafish naturally inhabit ion-poor environments but they are tolerant to a range of salinities from essentially ion-free to brackish waters.[49] Greater than 60 ion transport related proteins were identified in the gill proteome (see Supplementary Information 2.2), encompassing the majority of proteins hypothesized to be responsible for ion regulation and acid-base homeostasis in freshwater fish.[50, 51] These include numerous isoforms of Na^+/K^+-ATPase and V-type H^+-ATPases, sodium bicarbonate co-transporters, anion exchangers, and several carbonic anhydrases. However, several proteins of interest such as the Na^+/H^+ exchanger and epithelial Na^+ channel were not identified in our analysis. A number of other Na^+, K^+, and Ca^{2+} channels typically found in muscle and nervous tissue were also identified. This is expected, given that gill tissue contains both vascular smooth muscle cells, neuroepithelial cells and mitochondrion-rich cells.[50, 52, 53]

2.4.2 Protein Turnover

The gill plays a pivotal role in maintaining physiological homeostasis and, as such, it is expected that gill tissue would exhibit significant plasticity in the face of environmental changes.[54] The gills respond to changes in the internal and

external environment by changing their complement of membrane transport proteins,[55] the levels of enzymes involved in energy metabolism,[56] and even the physical structure of the lamellae.[57] Protein synthesis rates in the fish gill are consistently high relative to those measured in other tissues.[54, 58, 59] It is not surprising, therefore, that we identified a large number of proteins related to protein synthesis and degradation (see Supplementary Information 2.3). The ribosome is the primary cellular component involved in protein synthesis and our analysis identified 23 of 27 and 22 of 28 of the expected small and large ribosomal subunit proteins, respectively. The proteasome and lysosome are responsible for protein degradation[60] with short-lived proteins generally degraded by the proteasome and long-lived proteins degraded by the lysosome.[60, 61] The proteasome also plays an important role in a variety of basic cellular processes, such as the regulation of cell cycle, division, development, and differentiation, and modulation of immune and inflammatory responses.[62, 63] We identified 22 out of 25 of the protein subunits of the 26S proteasome. Information was not available for expected lysosomal proteins for *Danio rerio*. Only 12 of the expected 76 proteins involved in ubiquitin-mediated proteolysis were present in the identified zebrafish gill proteome (see Supplementary Information 2.3). It is possible that a number of proteins in this pathway were not identified due to relatively low protein abundance in the sample.

2.4.3 Energy Metabolism

High protein turnover rates and ion and acid-base regulation are energetically expensive processes and necessitate a relatively high organ-specific

metabolic rate. In rainbow trout, gill NaCl uptake alone can account for 37% of tissue oxygen demand and up to 4% of the animal's total energy expenditure.[64] The gill is largely an aerobic tissue that exhibits a strong dependence upon glucose and lactate for ATP production.[65, 66] The zebrafish gill proteome contained the full complement of glycolytic enzymes with several isoforms present for most proteins in the pathway (Figure 2.4). A number of proteins critical to the regulation of carbohydrate metabolism were also found, including several subunits of AMP-activated protein kinase, an insulin-like growth factor 2 receptor, a glucagon-like peptide-2 receptor, pyruvate dehydrogenase phosphatase 2, as well as a predicted protein similar to glucose transporter X. In total, 222 proteins involved in the metabolism of carbohydrates, amino acids, and lipids were identified in the gill (see Supplementary Information 2.4).

Coverage is slightly less comprehensive for biochemical pathways such as oxidative phosphorylation (47%), whose constituents are largely localized to membranes. We have far greater coverage for pathways such as the citrate cycle (88%) or glycolysis (74%) that contain few membrane-associated proteins. Greater than 240 integral membrane proteins were identified from the zebrafish gill proteome. Although this number represents only 4% of the proteins identified, it is still far more than the numbers identified from fish tissue in any other study to date (~45 membrane proteins;[22] 78 membrane proteins;[67] see Supplementary Information 2.5). Eight, 6, 17, and 19% of the integral membrane proteins were identified in the NH_4HCO_3, MeOH, urea, and SDS fractions, respectively. The remaining 53% were identified in more than one solvent, but

always included urea or SDS as one of the solvents, further illustrating the importance of these two solvents in the solubilization of membrane proteins. The number of integral membrane proteins we identified is still likely a substantial underestimate of the true number present considering that annotation information is unavailable for >30% of the proteome.

2.4.4 Biomarkers

Comprehensive proteomic profiling is a powerful technique for biomarker discovery and it will be an important tool for defining the bioactivity of new compounds. A number of well-characterized biomarkers of environmental stress and toxicant exposure were expressed in the zebrafish gill (Table 2.2). Only 5 proteins were predicted to be involved in the metabolism of xenobiotics by cytochrome P450 in zebrafish but we identified approximately 10 unique proteins associated with this GO term in the gill proteome. A number of proteins related to the degradation of specific toxins like 1- and 2-methylnaphthalene, caprolactam, and atrazine were also expressed. Several proteins significant to metal toxicity were identified, including metallothionein 2.

In addition to the specific biomarkers referred to above, greater than 40 heat shock proteins (HSPs), chaperonins, and related proteins were expressed in the gill (see Supplementary Information 2.6). These proteins generally interact with unfolded or denatured proteins and facilitate refolding, repair, or degradation processes.[68] Heat shock proteins are ubiquitously present in all organisms and their expression levels are sensitive to a variety of stresses including heat, hypoxia, pathogens, and pH and osmotic disturbances. This well-conserved

response has lead to considerable interest in the use of HSPs as biomarkers of stress in animals.[68] The complexity and species-to-species variability of the heat shock response, however, has prohibited the use of HSPs as specific stress indicators in fish.[69] The sensitivity of proteomic profiling methods may help to clarify the relationship between HSP expression patterns and other biomarkers in response to a particular stimulus. Exploiting the well-characterized zebrafish model system can also minimize the uncertainties of working with lesser studied fish species.

2.5 Conclusions

We present the first large scale proteome profile of a teleost fish tissue. Using advanced sample preparation and data dependent MS/MS analysis techniques developed in our lab, we have identified 5716 proteins expressed in the zebrafish gill. The data provide critical baseline information on tissue-specific protein expression for future studies on gill physiology and aquatic toxicology. Future work will proceed on two fronts; one focused on refining the current method for increased proteome coverage (e.g., identifying more membrane proteins by enriching membrane proteome fractions from tissue extracts) and the second on the development and application of quantitative profiling techniques. We are currently optimizing a 2MEGA isotope labelling method for relative quantification of proteins from differently treated samples.[70-72] We envision that the combination of this labelling method and the current protocol should provide a means of carrying out large scale quantitative proteome profiling studies.

Analyzing such large-scale quantitative datasets will require advanced bioinformatics tools which are currently unavailable. With the progression of genomic sequencing of non-model systems and improved bioinformatics analyses, this technique should be applicable for physiological studies on any organism.

2.6 Literature Cited

1. Bosworth, C. A. I. V.; Chou, C.-W.; Cole, R. B.; Rees, B. B., Protein expression patterns in zebrafish skeletal muscle: Initial characterization and the effects of hypoxic exposure. *Proteomics* **2005,** 5, (5), 1362-1371.

2. Link, V.; Shevchenko, A.; Heisenberg, C. P., Proteomics of early zebrafish embryos. *BMC Dev. Biol.* **2006,** 6, 9.

3. Polanski, M.; Anderson, N. L., A list of candidate cancer biomarkers for targeted proteomics. *Biomarker Insights* **2006**, 1-48.

4. Wei, Y.; Chan, L. L.; Wang, D.; Zhang, H.; Wang, J.; Dai, J., Proteomic analysis of hepatic protein profiles in rare minnow (Gobiocypris rarus) exposed to perfluorooctanoic acid. *J. Proteome Res.* **2008,** 7, (4), 1729-1739.

5. MacCormack, T. J.; Goss, G. G., Identifying and predicting biological risks associated with manufactured nanoparticles in aquatic ecosystems. *J. Ind. Ecol.* **2008,** 12, (3), 286-296.

6. Mor, G.; Visintin, I.; Lai, Y.; Zhao, H.; Schwartz, P.; Rutherford, T.; Yue, L.; Bray-Ward, P.; Ward, D. C., Serum protein markers for early detection of ovarian cancer. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** 102, (21), 7677-7682.

7. Xiao, T.; Ying, W.; Li, L.; Hu, Z.; Ma, Y.; Jiao, L.; Ma, J.; Cai, Y.; Lin, D.; Guo, S.; Han, N.; Di, X.; Li, M.; Zhang, D.; Su, K.; Yuan, J.; Zheng, H.; Gao, M.; He, J.; Shi, S.; Li, W.; Xu, N.; Zhang, H.; Liu, Y.; Zhang, K.; Gao, Y.; Qian, X.; Cheng, S., An approach to studying lung cancer-related proteins in human blood. *Mol. Cell. Proteomics* **2005,** 4, (10), 1480-1486.

8. Marchant, G. E., Toxicogenomics and toxic torts. *Trends Biotechnol.* **2002,** 20, (8), 329-332.

9. Pandey, A.; Mann, M., Proteomics to study genes and genomes. *Nature (London)* **2000,** 405, (6788), 837-846.

10. Chanson, A.; Sayd, T.; Rock, E.; Chambon, C.; Sante-Lhoutellier, V.; Potier de Courcy, G.; Brachet, P., Proteomic analysis reveals changes in the liver protein pattern of rats exposed to dietary folate deficiency. *J. Nutr.* **2005,** 135, (11), 2524-2529.

11. Strey, C. W.; Winters, M. S.; Markiewski, M. M.; Lambris, J. D., Partial hepatectomy induced liver proteome changes in mice. *Proteomics* **2005,** 5, (1), 318-325.

12. Melle, C.; Ernst, G.; Schimmel, B.; Bleul, A.; Koscielny, S.; Wiesner, A.; Bogumil, R.; Moeller, U.; Osterloh, D.; Halbhuber, K.-J.; von Eggeling, F., Biomarker discovery and identification in laser microdissected head and neck squamous cell carcinoma with ProteinChip technology, two-dimensional gel electrophoresis, tandem mass spectrometry, and immunohistochemistry. *Mol. Cell. Proteomics* **2003,** 2, (7), 443-452.

13. Tay, T. L.; Lin, Q.; Seow, T. K.; Tan, K. H.; Hew, C. L.; Gong, Z., Proteomic analysis of protein profiles during early development of the zebrafish, Danio rerio. *Proteomics* **2006,** 6, (10), 3176-3188.

14. Macher, B. A.; Yen, T.-Y., Proteins at membrane surfaces - a review of approaches. *Mol. BioSyst.* **2007,** 3, (10), 705-713.

15. Speers, A. E.; Wu, C. C., Proteomics of Integral Membrane Proteins - Theory and Application. *Chem. Rev. (Washington, DC, U. S.)* **2007,** 107, (8), 3687-3714.

16. Wang, N.; MacKenzie, L.; De Souza, A. G.; Zhong, H.; Goss, G.; Li, L., Proteome profile of cytosolic component of zebrafish liver generated by LC-ESI MS/MS combined with trypsin digestion and microwave-assisted acid hydrolysis. *J. Proteome Res.* **2007,** 6, (1), 263-272.

17. Brunt, J.; Hansen, R.; Jamieson, D. J.; Austin, B., Proteomic analysis of rainbow trout (Oncorhynchus mykiss, Walbaum) serum after administration of probiotics in diets. *Vet. Immunol. Immunopathol.* **2008,** 121, (3-4), 199-205.

18. Ky, C. L.; de Lorgeril, J.; Hirtz, C.; Sommerer, N.; Rossignol, M.; Bonhomme, F., The effect of environmental salinity on the proteome of the sea bass (Dicentrarchus labrax L.). *Anim. Genet.* **2007,** 38, (6), 601-608.

19. Martin, S. A. M.; Mohanty, B. P.; Cash, P.; Houlihan, D. F.; Secombes, C. J., Proteome analysis of the Atlantic salmon (Salmo salar) cell line SHK-1 following recombinant IFN-γ stimulation. *Proteomics* **2007,** 7, (13), 2275-2286.

20. Mezhoud, K.; Praseuth, D.; Puiseux-Dao, S.; Francois, J.-C.; Bernard, C.; Edery, M., Global quantitative analysis of protein expression and phosphorylation status in the liver of the medaka fish (Oryzias latipes) exposed to microcystin-LR I. Balneation study. *Aquat. Toxicol.* **2008,** 86, (2), 166-175.

21. Smith, R. W.; Wang, J.; Bucking, C. P.; Mothersill, C. E.; Seymour, C. B., Evidence for a protective response by the gill proteome of rainbow trout exposed to X-ray induced bystander signals. *Proteomics* **2007,** 7, (22), 4171-4180.

22. Lucitt, M. B.; Price, T. S.; Pizarro, A.; Wu, W.; Yocum, A. K.; Seiler, C.; Pack, M. A.; Blair, I. A.; FitzGerald, G. A.; Grosser, T., Analysis of the zebrafish proteome during embryonic development. *Mol. Cell. Proteomics* **2008,** 7, (5), 981-994.

23. Dahm, R.; Geisler, R., Learning from small fry: The zebrafish as a genetic model organism for aquaculture fish species. *Mar. Biotechnol.* **2006,** 8, (4), 329-345.

24. Jonz, M. G.; Nurse, C. A., Epithelial mitochondria-rich cells and associated innervation in adult and developing zebrafish. *J. Comp. Neurol.* **2006,** 497, (5), 817-832.

25. Phelps, H. A.; Neely, M. N., Evolution of the zebrafish model: from development to immunity and infectious disease. *Zebrafish* **2005,** 2, (2), 87-103.

26. Henry, T. R.; Spitsbergen, J. M.; Hornung, M. W.; Abnet, C. C.; Peterson, R. E., Early life stage toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in zebrafish (Danio rerio). *Toxicol. Appl. Pharmacol.* **1997,** 142, (1), 56-68.

27. Jorgensen, A.; Andersen, O.; Bjerregaard, P.; Rasmussen, L. J., Identification and characterization of an androgen receptor from zebrafish Danio rerio. *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2007,** 146C, (4), 561-568.

28. Kausch, U.; Alberti, M.; Haindl, S.; Budczies, J.; Hock, B., Biomarkers for exposure to estrogenic compounds: gene expression analysis in zebrafish (Danio rerio). *Environ. Toxicol.* **2008,** 23, (1), 15-24.

29. Ramírez, O. A. B.; Garcia, F. P., Genotoxic damage in zebra fish (Danio rerio) by arsenic in waters from Zimapan, Hidalgo, Mexico. *Mutagenesis* **2005,** 20, (4), 291-295.

30. Schnurstein, A.; Braunbeck, T., Tail moment versus tail length--application of an in vitro version of the comet assay in biomonitoring for genotoxicity in native surface waters using primary hepatocytes and gill cells from zebrafish (Danio rerio). *Ecotoxicol. Environ. Saf.* **2001,** 49, (2), 187-196.

31. Wu, S. M.; Zheng, Y. D.; Kuo, C.-H., Expression of mt2 and smt-B upon cadmium exposure and cold shock in zebrafish (Danio rerio). *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2008,** 148C, (2), 184-193.

32. Perry, S. F., The chloride cell: structure and function in the gills of freshwater fishes. *Annu. Rev. Physiol.* **1997,** 59, 325-347.

33. Rombough, P. J.; Moroz, B. M., The scaling and potential importance of cutaneous and branchial surfaces in respiratory gas exchange in larval and juvenile walleye Stizostedion vitreum. *J. Exp. Biol.* **1997,** 200, (18), 2459-2468.

34. Wood, C. M.; Kelly, S. P.; Zhou, B.; Fletcher, M.; O'Donnell, M.; Eletti, B.; Part, P., Cultured gill epithelia as models for the freshwater fish gill. *Biochim. Biophys. Acta, Biomembr.* **2002,** 1566, (1-2), 72-83.

35. Wang, N.; Xie, C.; Young, J. B.; Li, L., Off-line two-dimensional liquid chromatography with maximized sample loading to reversed-phase liquid chromatography-electrospray ionization tandem mass spectrometry for shotgun proteome analysis. *Anal. Chem. (Washington, DC, U. S.)* **2009,** 81, (3), 1049-1060.

36. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Anal. Chem. (Washington, DC, U. S.)* **2008,** 80, (12), 4696-4710.

37. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **2007,** 4, (3), 207-214.

38. Gong, Y.; Wang, N.; Wu, F.; Cass, C. E.; Damaraju, S.; Mackey, J. R.; Li, L., Proteome profile of human breast cancer tissue generated by LC-ESI-MS/MS combined with sequential protein precipitation and solubilization. *J. Proteome Res.* **2008,** 7, (8), 3583-3590.

39. The UniProt Consortium. The universal protein resource (UniProt). *Nucleic Acids Res.* **2008,** 36, (Database Iss), D190-D195.

40. Ashburner, M.; Ball, C. A.; Blake, J. A.; Botstein, D.; Butler, H.; Cherry, J. M.; Davis, A. P.; Dolinski, K.; Dwight, S. S.; Eppig, J. T.; Harris, M. A.; Hill, D. P.; Issel-Tarver, L.; Kasarskis, A.; Lewis, S.; Matese, J. C.; Richardson, J. E.; Ringwald, M.; Rubin, G. M.; Sherlock, G., Gene ontology: tool for the unification of biology. *Nat. Genet.* **2000,** 25, (1), 25-29.

41. Dennis, G.; Sherman, B. T.; Hosack, D. A.; Yang, J.; Gao, W.; Lane, H. C.; Lempicki, R. A., DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol.* **2003,** 4, (9), 11.

42. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; Yamanishi, Y., KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **2008,** 36, (Database Iss), D480-D484.

43. Aivaliotis, M.; Corvey, C.; Tsirogianni, I.; Karas, M.; Tsiotis, G., Membrane proteome analysis of the green-sulfur bacterium Chlorobium tepidum. *Electrophoresis* **2004,** 25, (20), 3468-3474.

44. Santoni, V.; Molloy, M.; Rabilloud, T., Membrane proteins and proteomics: un amour impossible? *Electrophoresis* **2000,** 21, (6), 1054-1070.

45. Weiner, J. H.; Li, L., Proteome of the Escherichia coli envelope and technological challenges in membrane proteome analysis. *Biochim. Biophys. Acta, Biomembr.* **2008,** 1778, (9), 1698-1713.

46. Flicek, P.; Aken, B. L.; Beal, K.; Ballester, B.; Caccamo, M.; Chen, Y.; Clarke, L.; Coates, G.; Cunningham, F.; Cutts, T.; Down, T.; Dyer, S. C.; Eyre, T.; Fitzgerald, S.; Fernandez-Banet, J.; Graef, S.; Haider, S.; Hammond, M.; Holland, R.; Howe, K. L.; Howe, K.; Johnson, N.; Jenkinson, A.; Kahari, A.; Keefe, D.; Kokocinski, F.; Kulesha, E.; Lawson, D.; Longden, I.; Megy, K.; Meidl, P.; Overduin, B.; Parker, A.; Pritchard, B.; Prlic, A.; Rice, S.; Rios, D.; Schuster, M.; Sealy, I.; Slater, G.; Smedley, D.; Spudich, G.; Trevanion, S.; Vilella, A. J.; Vogel, J.; White, S.; Wood, M.; Birney, E.; Cox, T.; Curwen, V.; Durbin, R.; Fernandez-Suarez, X. M.; Herrero, J.; Hubbard, T. J. P.; Kasprzyk, A.; Proctor, G.; Smith, J.; Ureta-Vidal, A.; Searle, S., Ensembl 2008. *Nucleic Acids Res.* **2008,** 36, (Database Iss), D707-D714.

47. Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R., Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **1999,** 19, (3), 1720-1730.

48. Goss, G. G.; Perry, S. F.; Fryer, J. N.; Laurent, P., Gill morphology and acid-base regulation in freshwater fishes. *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.* **1998,** 119A, (1), 107-115.

49. Engeszer Raymond, E.; Patterson Larissa, B.; Rao Andrew, A.; Parichy David, M., Zebrafish in the wild: a review of natural history and new notes from the field. *Zebrafish* **2007,** 4, (1), 21-40.

50. Parks, S. K.; Tresguerres, M.; Goss, G. G., Interactions between Na+ channels and Na+-HCO-3 cotransporters in the freshwater fish gill MR cell: a model for transepithelial Na+ uptake. *Am. J. Physiol.* **2007,** 292, (2, Pt. 1), C935-C944.

51. Parks, S. K.; Tresguerres, M.; Goss, G. G., Theoretical considerations underlying Na+ uptake mechanisms in freshwater fishes. *Comp. Biochem. Physiol. C-Toxicol. Pharmacol.* **2008,** 148, (4), 411-418.

52. Jonz, M. G.; Nurse, C. A., New developments on gill innervation: insights from a model vertebrate. *J. Exp. Biol.* **2008,** 211, (15), 2371-2378.

53. Taylor, E. W.; Jordan, D.; Coote, J. H., Central control of the cardiovascular and respiratory systems and their interactions in vertebrates. *Physiol. Rev.* **1999,** 79, (3), 855-916.

54. Lyndon, A. R.; Houlihan, D. F., Gill protein turnover: costs of adaptation. *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.* **1998,** 119A, (1), 27-34.

55. Tresguerres, M.; Parks, S. K.; Katoh, F.; Goss, G. G., Microtubule-dependent relocation of branchial V-H+-ATPase to the basolateral membrane in the Pacific spiny dogfish (Squalus acanthias): a role in base secretion. *J. Exp. Biol.* **2006,** 209, (4), 599-609.

56. Driedzic, W. R.; Gesser, H., Energy metabolism and contractility in ectothermic vertebrate hearts: hypoxia, acidosis, and low temperature. *Physiol. Rev.* **1994,** 74, (1), 221-258.

57. Sollid, J.; De Angelis, P.; Gundersen, K.; Nilsson, G. E., Hypoxia induces adaptive and reversible gross morphological changes in crucian carp gills. *J. Exp. Biol.* **2003,** 206, (20), 3667-3673.

58. Lewis, J. M.; Costa, I.; Val, A. L.; Almeida-Val, V. M. F.; Gamperl, A. K.; Driedzic, W. R., Responses to hypoxia and recovery: repayment of oxygen debt is not associated with compensatory protein synthesis in the

Amazonian cichlid, Astronotus ocellatus. *J. Exp. Biol.* **2007,** 210, (11), 1935-1943.

59. Lewis, J. M.; Driedzic, W. R., Tissue-specific changes in protein synthesis associated with seasonal metabolic depression and recovery in the north temperate labrid, Tautogolabrus adspersus. *Am. J. Physiol.* **2007,** 293, (1, Pt. 2), R474-R481.

60. Demarchi, F.; Brancolini, C., Altering protein turnover in tumor cells: New opportunities for anti-cancer therapies. *Drug Resist. Updates* **2005,** 8, (6), 359-368.

61. Ohsumi, Y., Protein turnover. *IUBMB Life* **2006,** 58, (5-6), 363-369.

62. Glickman, M. H.; Ciechanover, A., The ubiquitin-proteasome proteolytic pathway: Destruction for the sake of construction. *Physiol. Rev.* **2002,** 82, (2), 373-428.

63. Richardson, P.; Hideshima, T.; Mitsiades, C.; Anderson, K., Proteasome inhibition in hematologic malignancies. *Ann. Med. (Basingstoke, U. K.)* **2004,** 36, (4), 304-314.

64. Morgan, J. D.; Iwama, G. K., Energy cost of NaCl transport in isolated gills of cutthroat trout. *Am. J. Physiol.* **1999,** 277, (3, Pt. 2), R631-R639.

65. Mommsen, T. P., Biochemical characterization of the rainbow trout gill. *Journal of Comparative Physiology* **1984,** 154, (2), 191-198.

66. Soengas, J. L.; Barciela, P.; Aldegunde, M.; Andrés, D., Gill carbohydrate metabolism of rainbow trout is modified during gradual adaptation to sea water. *J. Fish Biol.* **1995,** 46, (5), 845-56.

67. Nazarian, J.; Hathout, Y.; Vertes, A.; Hoffman, E. P., The proteome survey of an electricity-generating organ (Torpedo californica electric organ). *Proteomics* **2007,** 7, (4), 617-627.

68. Basu, N.; Todgham, A. E.; Ackerman, P. A.; Bibeau, M. R.; Nakano, K.; Schulte, P. M.; Iwama George, K., Heat shock protein genes and their functional significance in fish. *Gene* **2002,** 295, (2), 173-183.

69. Iwama, G. K.; Afonso, L. O. B.; Todgham, A.; Ackerman, P.; Nakano, K., Are hsps suitable for indicating stressed states in fish? *J. Exp. Biol.* **2004,** 207, (1), 15-19.

70. Ji, C.; Guo, N.; Li, L., Differential dimethyl labeling of N-termini of peptides after guanidination for proteome analysis. *J. Proteome Res.* **2005,** 4, (6), 2099-2108.

71. Ji, C.; Lo, A.; Marcus, S.; Li, L., Effect of 2MEGA labeling on membrane proteome analysis using LC-ESI QTOF MS. *J. Proteome Res.* **2006,** 5, (10), 2567-2576.

72. Ji, C.; Zhang, N.; Damaraju, S.; Damaraju, V. L.; Carpenter, P.; Cass, C. E.; Li, L., A study of reproducibility of guanidination-dimethylation labeling and liquid chromatography matrix-assisted laser desorption ionization mass spectrometry for relative proteome quantification. *Anal. Chim. Acta* **2007,** 585, (2), 219-226.

Chapter 3

Using Stable Isotope Labelling to Quantitate the Proteome Effects on Gills from Zebrafish (*Danio rerio*) Exposed to Naphthenic Acids*

3.1 Introduction

The aquatic environment is particularly susceptible to pollution through the intentional and accidental release of chemicals into rivers, lakes, seas, and oceans. These chemicals can be classified as metals, nanoparticles, persistent organic pollutants, pesticides, petroleum hydrocarbons, and pharmaceuticals. The question remains as to what effects these chemicals have on aquatic wildlife. Endocrine disrupting compounds (EDCs) have been a large focus of aquatic research, due to threat to drinking water and concern about possible estrogenic and other detrimental effects on wildlife and humans. Similarly, oil sands process water (OSPW), particularly the naphthenic acids (NAs) component, has emerged as a hot topic among environmentalists and researchers.

*A version of this chapter is in preparation for publication as: A. G. De Souza, T. J. MacCormack, L. Li, and G. G. Goss, "A quantitative proteomic analysis of the biological effects on the gills of zebrafish (*Danio rerio*) exposed to commercial naphthenic acids and oil sands process water from the Athabasca oil sands in Alberta, Canada"

NAs are saturated aliphatic and alicyclic carboxylic acids whose structures are represented by the general chemical formula $C_nH_{2n+Z}O_2$, where n indicates the carbon number and Z is zero or a negative, even integer that specifies the hydrogen deficiency (Figure 3.1). NAs are natural constituents in petroleum, oil sands bitumen, and crude oil. In the Athabasca oil sands in northeastern Alberta, Canada, bitumen is separated from oil sands using the Clark hot water extraction method,[1] resulting in the extraction of NAs into the resulting wastewater. This wastewater, known as OSPW, is stored in tailings ponds on site in accordance with existing policy.[1] OSPW not only contains NAs at concentrations ranging from 40-120 mg/L[2] but also residual bitumen, polycyclic aromatic hydrocarbons (PAHs), and heavy metals.[3-6] OSPW has been shown to be toxic to aquatic organisms[7, 8] and NAs have been implicated in OSPW toxicity.[2, 3, 9-19] A study by Dokholyan and Magomedov (1983) investigated the toxicity of commercial sodium naphthenates and found 96 h LC_{50} values of 25 mg/L in chum salmon (*Oncorhynchus keta*), 50 mg/L in kutum (*Rutilus frisii kutum*), young roach (*Rutilus rutilus caspicus*), and sturgeon (*Acipenser gueldenstaedti*), and 75 mg/L in adult roach.[20] On the basis of haematological and biochemical parameters, they determined the maximum permissible concentration of sodium naphthenate in sea water to be 0.15 mg/L. Peters *et al.* (2007) studied the effects of OSPW and commercial NAs on yellow perch (*Perca flavescens*) and Japanese medaka (*Orizias latipes*), and found that exposure to both solutions resulted in similar types and severities of deformities between the two species, but at different concentrations.[21]

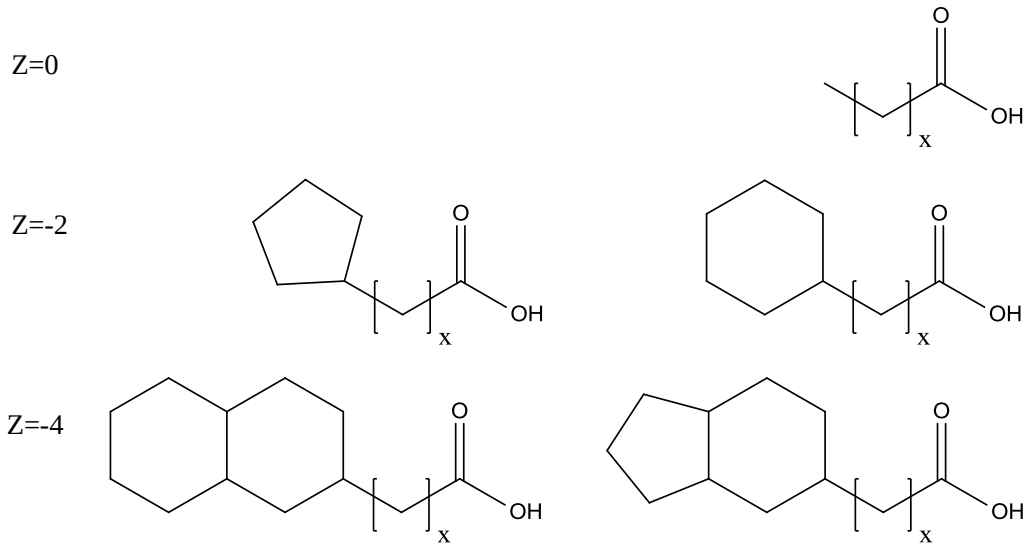

Figure 3.1. Sample structures of NAs of various hydrogen deficiencies, Z.

The objective of this study was to determine the effects of exposure to NAs on the proteome of zebrafish gill. In Chapter 2, we established a proteome profile of zebrafish gill with excellent coverage of important biochemical pathways relevant to the function of the gill in respiration, ion and acid-base homeostasis, and energy metabolism.[22] Numerous established and potential biomarkers of stress, diseases, and environmental contamination were also expressed in the gill. In this chapter, zebrafish were exposed for seven days to commercial naphthenic acids at two concentrations: 0.3 mg/L and 3 mg/L. We utilized a stable isotope labelling method (2MEGA)[23] to quantitate protein expression levels in NAs-exposed zebrafish and controls to evaluate the effects of NAs exposure on the biochemical pathways of zebrafish gill.

3.2 Experimental

3.2.1 Chemicals and Reagents

Tricaine methanesulfonate (TMS) was purchased from Syndel Laboratories Ltd. (Vancouver, BC, Canada). The total protein extraction kit was purchased from Biochain Institute Inc. (Hayward, CA, USA) and the bicinchoninic acid (BCA) assay kit was purchased from Pierce Biotechnology (Rockford, IL, USA). Dithiothreitol (DTT) and iodoacetamide (IAA) were purchased from BioRad (Mississauga, ON, Canada). Sequencing-grade modified trypsin was purchased from Promega (Madison, WI, USA). LC-MS grade acetonitrile (ACN), methanol (MeOH) and water were purchased from Fisher Scientific Canada (Edmonton, AB, Canada). Isotopically-enriched formaldehyde

($^{13}CD_2O$) was purchased from Cambridge Isotope Laboratories (Andover, MA, USA). Naphthenic acids from Merichem Chemicals and Refinery Services LLC (Houston, Texas) were obtained from the laboratory of Dr. Philip Fedorak in the Department of Biological Sciences at the University of Alberta. All other chemicals and reagents were purchased from Sigma-Aldrich (Oakville, ON, Canada).

3.2.2 Zebrafish Exposure to Commercial Naphthenic Acids

Nacre strain (A/B background) zebrafish were taken from the Zebrafish Breeding Facility at the University of Alberta Biosciences Aquatics Facility and distributed among three tanks, where the fish acclimated for a period of two days. At the end of the acclimation period, one tank was assigned as a control, and the other two tanks were dosed to a total concentration of 0.3 mg/L (low dose) or 3 mg/L (high dose) NAs. All tanks, including controls, were adjusted with 0.1 M NaOH to pH 8 to facilitate the dissolution of NAs. The experiment was performed in triplicate and the start date staggered over three days to maintain consistency in gill harvesting. Fifty percent of the water in the exposure tanks and the control tank were exchanged every other day for a period of 7 days with either NAs water or Zebrafish Breeding Facility water. At the end of the experiment, fish were sacrificed on three consecutive days.

3.2.3 Gill Tissue Collection and Protein Extraction

Fish were anaesthetized in a 1.0 g/L solution of tricaine methanesulfonate. To reduce the content of blood proteins in the gill tissue, the circulatory system was flushed with ice-cold phosphate-buffered saline (PBS) containing 50 IU/mL

sodium heparin. Perfusion cannulae consisted of a length of PE10 tubing with an 8 mm long tip fabricated from a 30G needle. The cutting edge of the needle was removed and the tip was polished. Fish were positioned ventral-side up on a moistened sponge under a dissecting microscope. The body cavity was then opened and quickly rinsed with ice-cold PBS. The heart was exposed by blunt dissection and one ventricular wall was carefully pierced with a sharp 30G needle. The cannula was inserted and held in place by clamping the ventricle wall around it using blunt forceps. The heart was subsequently perfused at 300 µL/min with ice-cold PBS containing 50 IU/mL sodium heparin until gill tissue was visibly cleared of blood (1-5 min).

Following perfusion, whole gill baskets were excised and placed in a flat-bottomed microcentrifuge tube on dry ice, and proteins extracted according to the protocol provided by the manufacturer (Biochain Institute Inc., Hayward, CA) with slight modifications. One mL of 100 mM phenylmethylsulfonyl fluoride (dissolved in ethanol) per mL of 50X protease inhibitor (PI) was added to the extraction buffer, and 10 mL of 1X PI added per gram of tissue. A BCA assay was conducted to determine the gill protein concentration using bovine serum albumin as the standard.

The individual gill protein extracts from each tank were pooled to generate three sets of samples (Biological Replicates A, B, and C), each consisting of a low dose exposure, a high dose exposure, and a control. A full workflow is outlined in Figure 3.2.

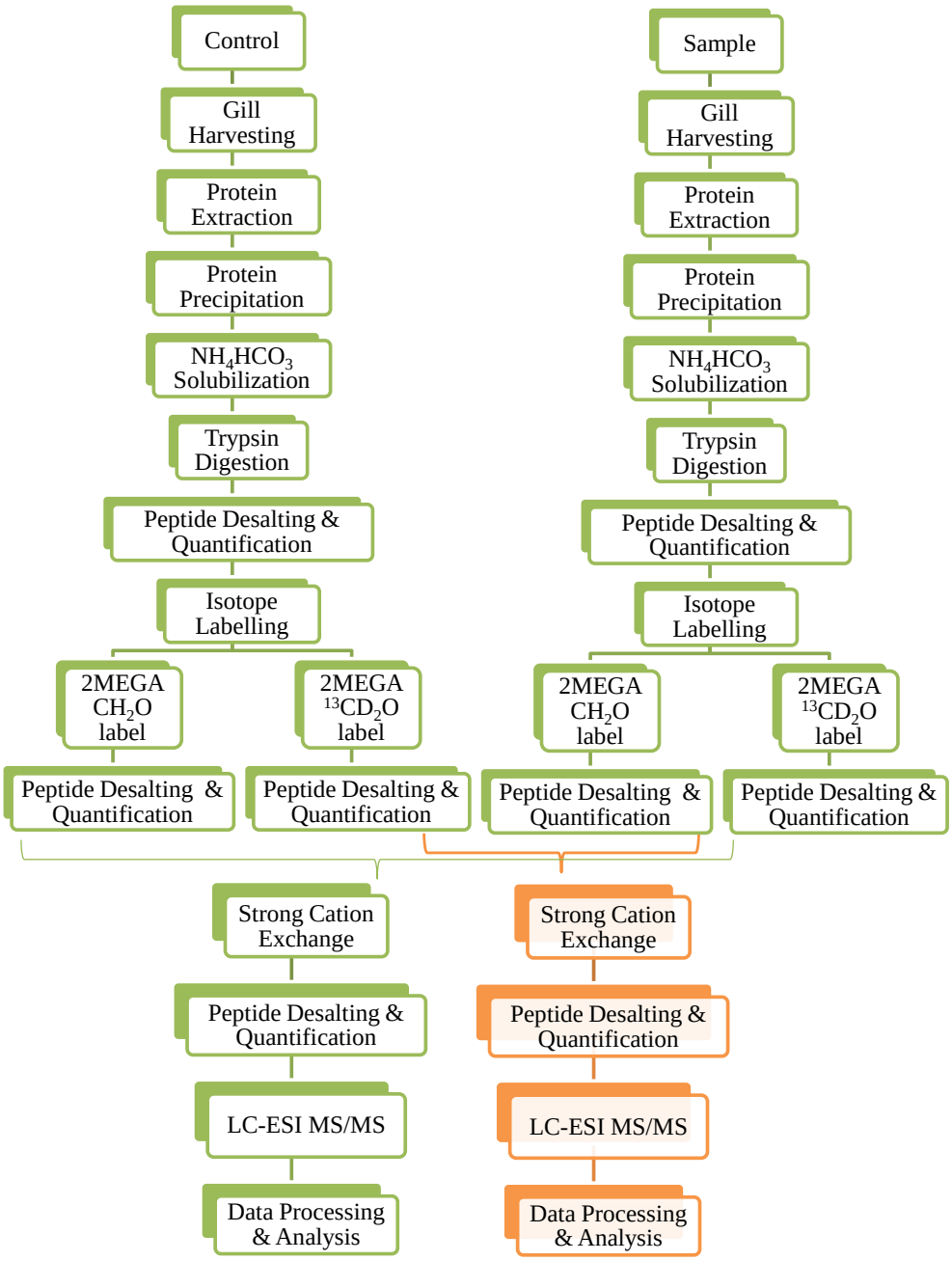

Figure 3.2. Workflow for the experimental design.

3.2.4 Protein Precipitation and In-Solution Digestion

Proteins were reduced and alkylated with DTT and IAA, respectively, and then precipitated with four times the volume of acetone (pre-cooled at -80 °C), kept at -20 °C overnight, centrifuged (20 800 g, 4 °C, 10 min), and the supernatant removed and discarded.

The protein pellet was partially solubilized in 100 mM NH_4HCO_3, sonicated in an ice bath for 5 s followed by a 30 s rest on ice (repeated a total of four times), shaken on a vortex-shaker (2 h, 4 °C), and centrifuged (20 800 g, 4 °C, 10 min). The supernatant was collected and diluted to 50 mM NH_4HCO_3. This solubilization process was repeated on the protein pellet. A BCA assay was conducted to determine the protein concentration for each solubilization. Negligible protein was detected in the second NH_4HCO_3 solubilization sample; therefore, protein digestion was conducted only on the first NH_4HCO_3 solubilization sample. The trypsin-to-protein ratio was 1:25 (w/w), and digestion occurred at 37 °C overnight.

The sequential solubilization scheme in Chapter 2 identified more proteins in the urea solubilized fraction versus the NH_4HCO_3 solubilized fraction. There was concern about carbamylation at lysines and N-termini of peptides, which would affect 2MEGA labelling, therefore we did not use urea as a solubilizing agent. Results from a study conducted after the NAs exposure study showed that urea induces carbamylation on the lysine of approximately 1% of peptides.[24]

3.2.5 Peptide Desalting and Labelling

Simultaneous salt removal and peptide quantitification was performed on the digests using a Polaris C18A column (Palo Alto, CA, USA, 3 µm, 300 Å, 4.6 x 50 mm) on an Agilent 1100 HPLC system with a UV detector.[25] Mobile Phase A consisted of 0.1% trifluoroacetic acid (TFA) in water and Mobile Phase B consisted of 0.1% TFA in acetonitrile. The elution gradient was 2.5% B for 5.50 min, 2.5-85% B from 5.50 to 5.51 min, 85% B from 5.51 to 15.50 min, 85-2.5% B from 15.50 to 15.51 min, and held at 2.5% B from 15.51 to 25.00 min at a flow rate of 1.000 mL/min. Fraction collection occurred at 7.25 min for a total of 1.00 min.

The desalted peptides were isotopically-labelled using the 2MEGA strategy, where either CH_2O or $^{13}CD_2O$ was used to dimethylate the N-termini of peptides after guanidinylation of lysines.[23] The reaction scheme is illustrated in Figure 3.3. The peptides were divided into two sets for light and heavy labelling by CH_2O and $^{13}CD_2O$, respectively. A Gilson 215 automated liquid handler was used for the labelling reaction, requiring some modification of the protocol. The procedure was performed at 37 ^{o}C and all liquids were dispensed by the liquid handler. In brief, 50 µL of 2 M O-methylisourea (prepared in 1:1 (v/v) 2 M NaOH and 100 mM $NaHCO_3$) was added to 100 µL of 0.5 µg/µL peptide solution. The pH was adjusted with 12 µL of 2 M NaOH and the guanidinylation reaction was allowed to proceed for 60 minutes. At the end of the 60 minutes, the liquid handler dispensed 30 µL of 6 M HCl and 24 µL of 1 M acetate buffer to adjust the pH. Then, 20 µL of 1 M picoline borane (prepared in MeOH) and 10

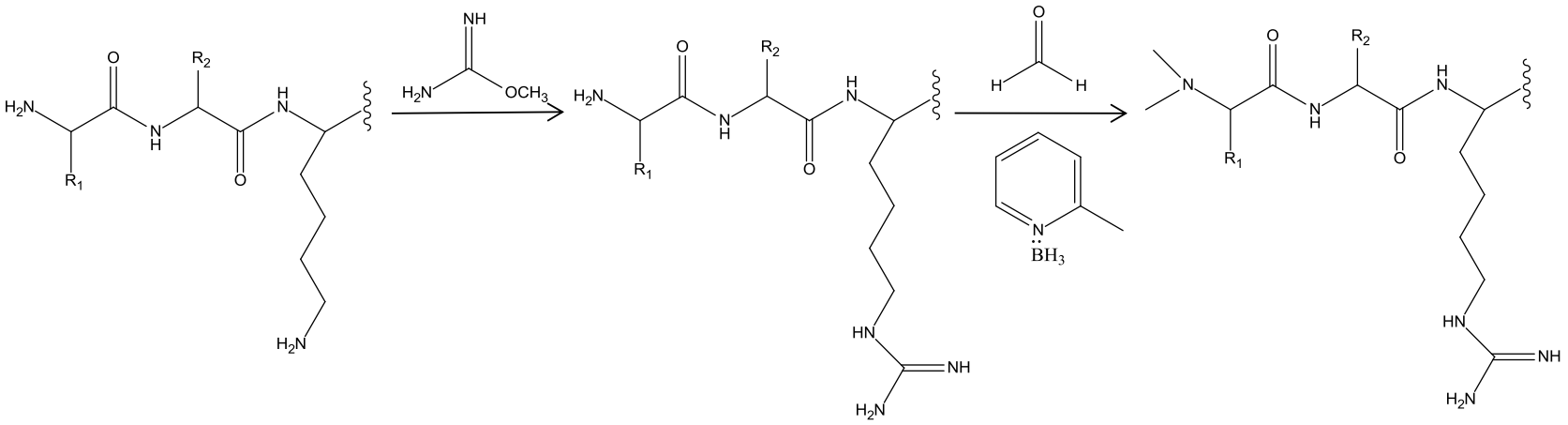

Figure 3.3. Reaction scheme of 2MEGA illustrating guanidinylation of lysine followed by reductive dimethylation of the N-terminus.

µL of 1% (v/v) CH_2O (for light labelling) or $^{13}CD_2O$ (for heavy labelling) were dispensed and the reaction allowed to proceed for 30 minutes. The reaction was quenched with the addition of 16 µL of 1 M NH_4HCO_3. The labelled digests were mixed such that a light-labelled control and a heavy-labelled sample constituted the forward case, and a heavy-labelled control and a light-labelled sample constituted the reverse case.

3.2.6 Strong Cation Exchange Chromatography

The forward- and reverse-labelled mixtures were separated into individual fractions on a PolySULFOETHYL A column (PolyLC, Columbia, MD, 5 µm, 300 Å, 2.1 x 250 mm) by strong cation exchange (SCX) on an Agilent 1100 HPLC system (Palo Alto, CA). Mobile Phase A consisted of 10 mM KH_2PO_4 in 20% ACN (pH 2.76) and Mobile Phase B consisted of 10 mM KH_2PO_4 and 0.5 M KCl in 20% ACN (pH 2.76). The elution gradient was 0% B for 7 min, 0-6% B from 7 to 8 min, 6-28% B from 8 to 36 min, 28-40% B from 36 to 44 min, 40-60% B from 44 to 49 min, 60-100% B from 49 to 53 min, 100% B from 53 to 58 min, 100-0% B from 58 to 60 min, and held at 0% B for 10 min at a flow rate of 0.200 mL/min. Fractions were collected at 1 min intervals.

3.2.7 LC-ESI QToF MS and MS/MS Analysis

Prior to LC-ESI MS/MS, the collected fractions from the SCX were subjected to salt removal and peptide quantitation as previously described in Section 3.2.5. A modification to the desalting and quantification was the use of a smaller Polaris C18A column (Palo Alto, CA, USA, 3 µm, 300 Å, 2.0 x 50 mm) with an elution gradient of 2.5% B for 5.50 min, 2.5-85% B from 5.50 to 5.51

min, 85% B from 5.51 to 10.00 min, 85-2.5% B from 10.00 to 10.01 min, and held at 2.5% B from 10.01 to 15.00 min at a flow rate of 0.200 mL/min. Fraction collection occurred at 7.00 min for a total of 1.00 min. A smaller sample for desalting facilitated the use of a smaller column. Adjacent fractions from the SCX fractionation were pooled if the total amount of peptides in individual fractions was less than 1 µg.

The desalted SCX peptide samples were analyzed using a Q-Tof Premier™ mass spectrometer equipped with a nanoACQUITY® UPLC system (Waters, Milford, MA). Approximately 1 µg of sample was loaded onto an Atlantis dC18 column (Waters, 3 µm, 100 Å, 75 µm x 150 mm). Mobile Phase A consisted of 0.1% formic acid in water and Mobile Phase B consisted of 0.1% formic acid in acetonitrile. The elution gradient was 2-7% B from 0 to 2 min, 7-20% B from 2 to 85 min, 20-30% B from 85 to 110 min, 30-45% B from 110 to 115 min, 45-90% B from 115 to 120 min, held at 90% B for 5 min, and 90-2% B for 5 min at 250 nL/min. A precursor ion exclusion (PIE) strategy in LC MS/MS was used where ions whose acquired MS/MS spectra resulted in positive peptide identification were excluded from MS/MS acquisition in the subsequent run of the adjacent SCX sample.[26]

3.2.8 Peptide Quantitation from MS/MS Data

Mascot Distiller 2.3.2.0 (Matrix Science, London, UK) was used to generate quantitation results for each LC-ESI MS/MS run. Raw LC MS/MS data were lock-mass corrected and de-isotoped, and searched using the Mascot search program (Matrix Science, London, UK) within Mascot Distiller.

For each LC-ESI MS/MS run, database searches were restricted to *Danio rerio* in the NCBInr database (containing 7836153 sequences and 2700770747 residues). The following search parameters were selected for all database searches: enzyme, trypsin; maximum missed cleavages, 1; fixed modifications, carbamidomethyl (C), guandinyl (K); variable modifications, dimethylation0 (N-term), dimethylation6 (N-term); peptide tolerance, ± 30 ppm; fragment mass tolerance, ± 0.2 Da; peptide charge, 1+, 2+, 3+; quantitation method, 2MEGA_QuantTest. Mascot Distiller generated a list of quantitated peptide ratios.

A final list of quantitated peptide ratios was generated for each of the forward and reverse cases for each of Biological Replicates A, B, and C using the following criteria: For redundant peptides in each list, the peptide ratio was calculated as a geometric average of the observed Mascot Distiller ratios. Peptides were considered quantitated only if peptide quantitation ratios were obtained in both the forward and reverse cases and the relative difference was equal to or less than 0.5, as previously rationalized.[27] Relative difference was calculated using the following equation: $\frac{\mathbf{2}(\boldsymbol{ratio_f - ratio_r^{-1}})}{(\boldsymbol{ratio_f + ratio_r^{-1}})}$, where $ratio_f$ and $ratio_r$ refer to the peptide quantitation ratios in the forward and reverse cases, respectively. A conservative list of peptides was generated by calculating coefficients of variation (CVs) for each peptide and discarding peptides with CVs greater than 50%. The remaining peptides were manually checked and discarded if the Mascot Distiller peak fitting was incorrect.

The peptides were then organized by proteins according to their accession identification number with their associated quantitation ratio from each biological replicate. Geometric averages were calculated for each protein using the peptide quantitation ratios from the Biological Replicates. Proteins were considered differentially-expressed if the protein quantitation ratios were greater than 1.5 or less than 0.67. The threshold for differential abundance represents two standard deviations and 95% confidence determined in a previous study.[27]

3.3 Results and Discussion

We investigated NAs exposure on zebrafish at two concentrations: 3 mg/L and 0.3 mg/L. These concentrations were chosen after consultation with literature[10, 28] and Dr. Phillip Fedorak from the Department of Biological Sciences at the University of Alberta to observe sublethal effects of NAs exposure. Although zebrafish gill protein extracts were obtained from both the 3 mg/L and 0.3 mg/L NAs exposures, we only present results from the 3 mg/L NAs exposure, the results of which indicated that analyzing the 0.3 mg/L NAs exposure would not be beneficial.

Our strategy to analyze the biological effects of zebrafish exposure to NAs involved the 2MEGA stable isotope labelling method. In this study, we used both forward and reverse labelling, as well as pooled biological replicates, to corroborate our quantitation results. Forward and reverse labelling served as atypical technical replicates. Biological replicates were performed in triplicate with each replicate consisting of pooled zebrafish gill. The distribution of fish is

shown in Table 3.1. Pooling gill sample addressed two issues: 1) limited tissue availability and 2) reducing biological variation. We recognize that pooling does not eliminate the need for technical replicates to determine inter-pool variability. Utilizing both technical and biological replicates lead to greater confidence in the final list of differentially-expressed proteins not being due to random artefacts.

For each biological replicate, peptides were considered quantitated if quantitation ratios were observed by Mascot Distiller in both the forward and reverse labelling cases. Furthermore, peptides were not considered for quantitation if their Mascot score was lower than the Mascot threshold score for identity at 95% confidence level. Based on these criteria, 1053, 1223, and 915 peptides were quantitated in Biological Replicates A, B, and C, respectively (Figure 3.4). Figure 3.5 shows the MS spectra of a tryptic peptide for forward- and reverse-labelled samples. Of these quantitated peptides, 15 were differentially-expressed in at least two of the three biological replicates, most of which were downregulated in the NAs-exposed sample compared to the control. This data is listed in Table 3.2.

While we were initially surprised by the low number of differentially-expressed proteins, we also realize that the concentration of NAs used in our study was lower than those used in toxicology studies of NAs with fish (12 – 100 mg/L).[20] Our purpose was not to determine LC_{50} but to monitor a sublethal response upon NAs exposure. It is not unreasonable to expect that exposure to higher concentrations of NAs might result in a greater number of differentially-expressed proteins. Consequently, we chose not to analyze the 0.3 mg/L NAs-

Table 3.1. Distribution of male and female fish in controls and samples of Biological Replicates A, B, and C.

Sex	Biological Replicate					
	A		B		C	
	Control	Sample	Control	Sample	Control	Sample
Male	6	5	5	6	4	5
Female	4	5	5	3	6	4

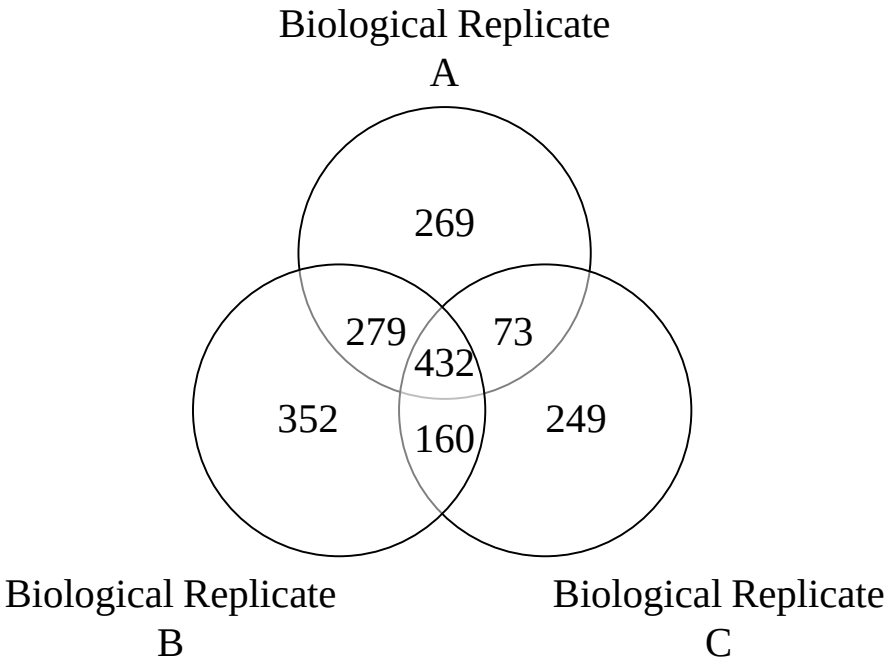

Figure 3.4. Quantitated peptides in Biological Replicates A, B, and C.

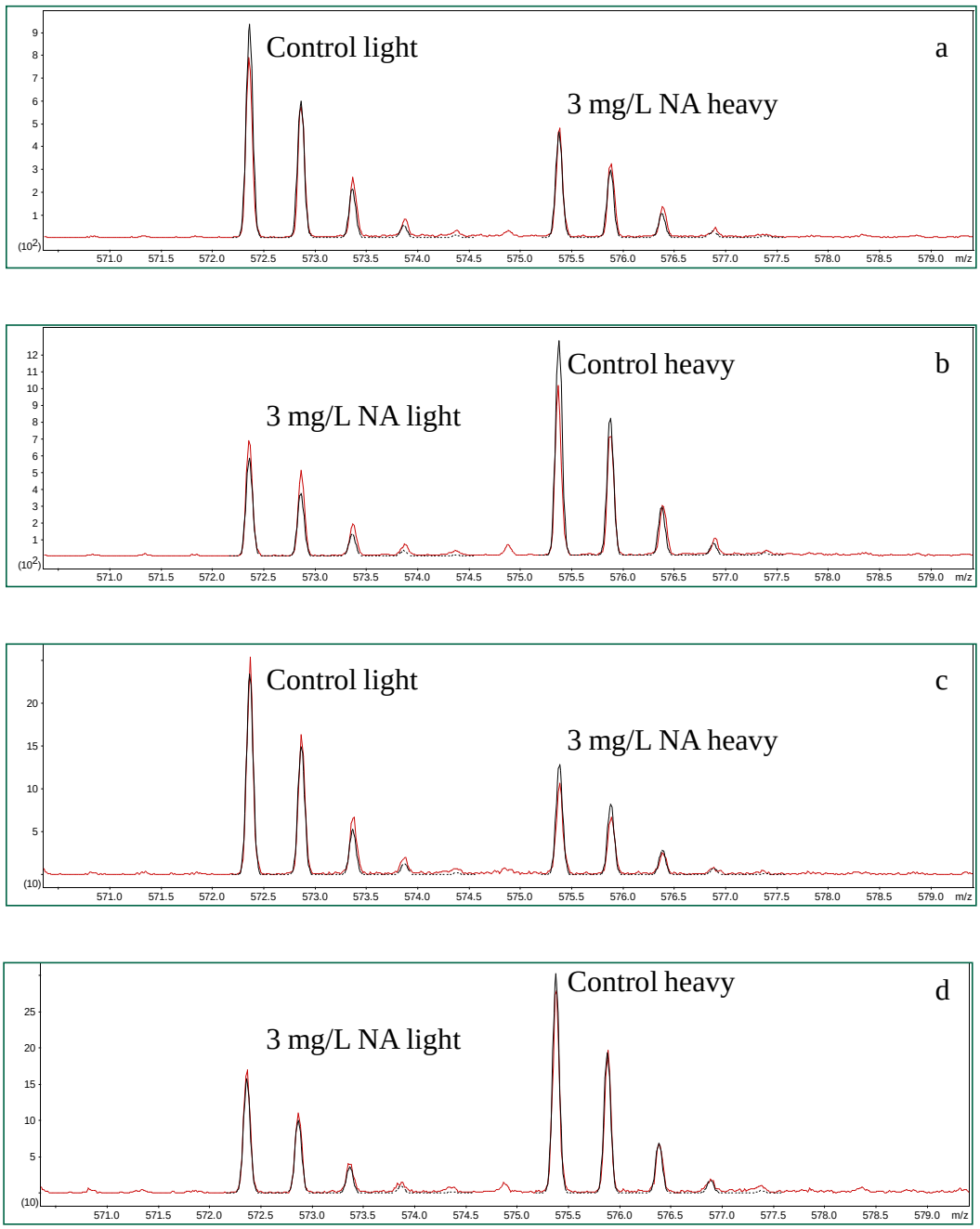

Figure 3.5. Peak fitting results for a representative peptide FVQLVQLLR from protein gi|94733733, a novel protein similar to vitellogenin 1 (vg1), for the forward- and reverse-labelled samples of Biological Replicates B and C at 3 mg/L NA exposure. The trace in red is the observed spectrum. The trace in black is the theoretical peak fit determined by Mascot Distiller for quantitation. The legend is as follows: a) Biological Replicate B forward labelling, b) Biological Replicate B reverse labelling, c) Biological Replicate C forward labelling, and d) Biological Replicate C reverse labelling.

Table 3.2. Differentially-expressed proteins quantified in at least two of the three biological replicates.

#	gi|	Protein Name	Peptides			GA[1]	SD_{geo}[2]
			A	B	C		
1	577534	major histocompatibility complex class I UAA	1	1	0	0.59	1.26
2	6636423	major histocompatibility complex class I antigen	1	1	0	0.60	1.30
3	18859281	proteasome activator subunit 2	2	1	0	0.65	1.32
4	34785174	splicing factor, arginine/serine-rich I	2	1	0	1.66	1.24
5	41152046	chaperonin containing TCP1, subunit 6A (zeta 1)	1	1	0	0.57	1.23
6	47086021	aldolase a, fructose-bisphosphate, b	1	0	1	0.66	1.08
7	94732723	novel protein similar to vitellogenin 1 (vg1)	1	0	1	0.43	1.19
8	94733730	vitellogenin 1	3	2	0	0.55	1.14
9	94733733	novel protein similar to vitellogenin 1 (vg1)	1	1	2	0.46	1.20
10	94733734	novel protein similar to vitellogenin 1 (vg1)	1	0	2	0.35	1.37
11	113678458	vitellogenin 2 isoform 1	6	0	2	0.41	1.10
12	125811426	PREDICTED: hypothetical protein	2	5	1	0.52	1.21
13	166795887	vitellogenin 1	1	0	2	0.40	1.26
14	167621538	hypothetical protein LOC100136867	2	1	0	0.65	1.24
15	189528835	PREDICTED: sb:cb37	1	1	1	0.62	1.06

[1]GA = geometric average of peptide ratios across biological replicates
[2]SD_{geo} = geometric standard deviation

exposed zebrafish.

3.3.1 Naphthenic Acids and Reproductive Response

Among the differentially-expressed proteins were six vitellogenin proteins (gi|94733723 novel protein similar to vitellogenin 1, gi|94733730 vitellogenin 1, gi|94733733 novel protein similar to vitellogenin 1, gi|94733734 novel protein similar to vitellogenin 1, gi|166795887 vitellogenin 1, and gi|113678458 vitellogenin 2 isoform 1). Vitellogenins are the major precursor of egg yolk proteins, which provide energy reserves for embryonic development in oviparous organisms. They are typically found in the liver and blood of adult female oviparous vertebrates, but have been detected in the gill of fathead minnows.[29] In mature females, endogenous estrogens trigger the synthesis of vitellogenins in the liver, which are then released into the bloodstream and stored in developing oocytes. Males have a silent vitellogenin gene, although the gene can be activated by EDCs, such as xenobtioic estrogens. EDCs are capable of altering endocrine function through a variety of mechanisms, such as binding to hormone receptors or influencing signalling pathways. Consequently, the presence and upregulation of vitellogenin in males and females, respectively, has been used an indicator of exposure to estrogenic compounds.[30-32]

In this study, we observed a downregulation of vitellogenin in the NAs-exposed group compared to controls; vitellogenin levels were decreased upon NA exposure. Our finding supports inhibition of vitellogenin by NAs and it would appear that NAs might interfere with reproductive development. Decreasing levels of vitellogenins for the oocytes may have significant developmental

repercussions. Our study contained pools of both sexes of fish, therefore it would be advantageous to perform a study measuring vitellogenin levels in each sex to provide further insight into vitellogenin protein expression upon NAs exposure. It is also possible that, as a result of NAs exposure, energy resources are required elsewhere than reproduction.

3.3.2 Naphthenic Acids and Immune Response

Although vitellogenin is classified as a reproductive protein, there is evidence suggesting that it may play a role in immune and stress response.[32-35] Shi *et al.* (2006) showed that vitellogenin purified from rosy barb (*Puntius conhonius*), a freshwater fish, possessed antibacterial properties against *Escherichia coli* both *in vitro* and *in vivo*.[34] Bacterial colony growth was inhibited at a minimum vitellogenin concentration of 40 µg/mL and rosy barb injected with *Escherichia coli* showed increased levels of vitellogenin compared to controls, further suggesting an immune-based response and novel vitellogenin function. In another study, Tong *et al.* (2009) demonstrated that the injection of foreign molecules into male zebrafish upregulated transcriptional and translational sera levels of vitellogenin.[36] In our study, the downregulation of vitellogenin compared to controls suggests that NAs may have an immunosuppressive effect. This postulation is not only supported by the downregulation of vitellogenins, but also by the downregulation of two major histocompatibility complex (MHC) class I proteins (gi|577534 MHC class I UAA and gi|6636423 MHC class I antigen).

MHC class I proteins are expressed on the surface of most nucleated cells and are involved in immune response. MHC class I proteins bind to peptide

fragments of degraded molecules in the cell cytosol and carry them to the cell surface for recognition by T-lymphocytes. In this way, MHC class I proteins help the immune system recognize foreign substances.

The observed downregulation of MHC class I proteins and vitellogenin proteins suggests that NAs have an inhibitory effect on immune response. It is possible that cell resources are not being utilized for maximum benefit in the immune system because their resources are required elsewhere. Of course, this postulate needs further investigation as the immune response is complicated, involving a diverse class of molecules and pathways.

3.4 Conclusion

Zebrafish were exposed for seven days to commercially-purchased NAs at 3 mg/L and their NH_4HCO_3 soluble gill proteins analyzed using quantitative mass spectrometry techniques. Using the 2MEGA peptide labelling strategy, we simultaneously identified and quantitated 15 differentially-expressed proteins in at least two of the three Biological Replicates. Among the differentially-expressed proteins were two classes of downregulated proteins: vitellogenin and major histocompatibility complex class I. The downregulation of these specific proteins suggests that NAs exposure triggers an immunosuppressive response in zebrafish. Furthermore, the downregulation of vitellogenin proteins implies a reproductive system effect from NAs exposure. Further studies on individual sexes are required to establish NAs effect on reproductive-related proteins.

It would be beneficial to study the protein expression levels in other soluble gill fractions, as was done in Chapter 2. We would expect to detect more hydrophobic proteins, such as membrane proteins, in urea soluble fractions and, therefore, observe NAs effects on these proteins.

We are interested in studying what effects OSPW, at similar concentrations of NAs, will have on zebrafish gill, as there may be components other than NAs that contribute to the overall toxicity of OSPW. A study by Lister *et al.* found that OSPW exposure effects on goldfish (*Carassius auratus*) were not reproducible with NAs exposure,[4] thereby giving credence to the argument that NAs are not solely responsible for OSPW toxicity.

3.5 Literature Cited

1. FTFC (Fine Tailings Fundamentals Consortium), *Advances in oil sands tailings research*. Alberta Department of Energy: Edmonton, 1995; Vol. II.

2. Holowenko, F. M.; MacKinnon, M. D.; Fedorak, P. M., Characterization of naphthenic acids in oil sands wastewaters by gas chromatography-mass spectrometry. *Water Res.* **2002,** 36, (11), 2843-2855.

3. Frank, R. A.; Fischer, K.; Kavanagh, R.; Burnison, B. K.; Arsenault, G.; Headley, J. V.; Peru, K. M.; Van Der Kraak, G.; Solomon, K. R., Effect of carboxylic acid content on the acute toxicity of oil sands naphthenic acids. *Environ. Sci. Technol.* **2009,** 43, (2), 266-271.

4. Lister, A.; Nero, V.; Farwell, A.; Dixon, D. G.; Van Der Kraak, G., Reproductive and stress hormone levels in goldfish (Carassius auratus) exposed to oil sands process-affected water. *Aquat. Toxicol.* **2008,** 87, (3), 170-177.

5. MacKinnon, M. D., Development of the tailings pond at Syncrude's oil sands plant: 1978-1987. *AOSTRA J. Res.* **1989,** 5, (2), 109-133.

6. Madill, R. E. A.; Orzechowski, M. T.; Chen, G.; Brownlee, B. G.; Bunce, N. J., Preliminary risk assessment of the wet landscape option for

reclamation of oil sands mine tailings: bioassays with mature fine tailings pore water. *Environ. Toxicol.* **2001,** 16, (3), 197-208.

7. Farrell, A. P.; Kennedy, C. J.; Kolok, A., Effects of wastewater from an oil-sands-refining operation on survival, hematology, gill histology and swimming of fathead minnows. *Canadian Journal of Zoology* **2004,** 82, 1519-1527.

8. van den Heuvel, M. R.; Power, M.; Richards, J.; MacKinnon, M.; Dixon, D. G., Disease and gill lesions in yellow perch (Perca flavescens) exposed to oil sands mining-associated waters. *Ecotoxicol. Environ. Saf.* **2000,** 46, (3), 334-341.

9. Armstrong, S. A.; Headley, J. V.; Peru, K. M.; Germida, J. J., Differences in phytotoxicity and dissipation between ionized and nonionized oil sands naphthenic acids in wetland plants. *Environ. Toxicol. Chem.* **2009,** 28, (10), 2167-2174.

10. Clemente, J. S.; MacKinnon, M. D.; Fedorak, P. M., Aerobic biodegradation of two commercial naphthenic acids preparations. *Environ. Sci. Technol.* **2004,** 38, (4), 1009-1016.

11. Colavecchia, M. V.; Backus, S. M.; Hodson, P. V.; Parrott, J. L., Toxicity of oil sands to early life stages of fathead minnows (Pimephales promelas). *Environ. Toxicol. Chem.* **2004,** 23, (7), 1709-1718.

12. Crowe, A. U.; Han, B.; Kermode, A. R.; Bendell-Young, L. I.; Plant, A. L., Effects of oil sands effluent on cattail and clover: photosynthesis and the level of stress proteins. *Environ. Pollut. (Oxford, U. K.)* **2001,** 113, (3), 311-322.

13. Farwell, A.; Nero, V.; Croft, M.; Bal, P.; Dixon, D. G., Modified Japanese medaka embryo-larval bioassay for rapid determination of developmental abnormalities. *Arch. Environ. Contam. Toxicol.* **2006,** 51, (4), 600-607.

14. Frank, R. A.; Kavanagh, R.; Burnison, B. K.; Arsenault, G.; Headley, J. V.; Peru, K. M.; Van Der Kraak, G.; Solomon, K. R., Toxicity assessment of collected fractions from an extracted naphthenic acid mixture. *Chemosphere* **2008,** 72, (9), 1309-1314.

15. Gentes, M.-L.; Waldner, C.; Papp, Z.; Smits, J. E. G., Effects of oil sands tailings compounds and harsh weather on mortality rates, growth and detoxification efforts in nestling tree swallows (Tachycineta bicolor). *Environ. Pollut. (Amsterdam, Neth.)* **2006,** 142, (1), 24-33.

16. Harms, N. J.; Fairhurst Graham, D.; Bortolotti Gary, R.; Smits Judit, E. G., Variation in immune function, body condition, and feather corticosterone in nestling tree swallows (Tachycineta bicolor) on reclaimed wetlands in the Athabasca oil sands, Alberta, Canada. *Environ Pollut* 158, (3), 841-848.

17. Leung, S. S.; MacKinnon, M. D.; Smith, R. E. H., The ecological effects of naphthenic acids and salts on phytoplankton from the Athabasca oil sands region. *Aquat. Toxicol.* **2003,** 62, (1), 11-26.

18. Pollet, I.; Bendell-Young, L. I., Amphibians as indicators of wetland quality in wetlands formed from oil sands effluent. *Environ. Toxicol. Chem.* **2000,** 19, (10), 2589-2597.

19. Tetreault, G. R.; McMaster, M. E.; Dixon, D. G.; Parrott, J. L., Using reproductive endpoints in small forage fish species to evaluate the effects of athabasca oil sands activities. *Environ. Toxicol. Chem.* **2003,** 22, (11), 2775-2782.

20. Dokholyan, V. K.; Magomedov, A. K., Effects of sodium naphthenate on survival and some physiological-biochemical parameters of some fishes. *J. Ichthyol.* **1983,** 23, (6), 125-132.

21. Peters, L. E.; MacKinnon, M.; Van Meer, T.; van den Heuvel, M. R.; Dixon, D. G., Effects of oil sands process-affected waters and naphthenic acids on yellow perch (Perca flavescens) and Japanese medaka (Oryzias latipes) embryonic development. *Chemosphere* **2007,** 67, (11), 2177-2183.

22. De Souza, A. G.; MacCormack, T. J.; Wang, N.; Li, L.; Goss, G. G., Large-scale proteome profile of the zebrafish (Danio rerio) gill for physiological and biomarker discovery studies. *Zebrafish* **2009,** 6, (3), 229-238.

23. Ji, C.; Guo, N.; Li, L., Differential dimethyl labeling of N-termini of peptides after guanidination for proteome analysis. *J. Proteome Res.* **2005,** 4, (6), 2099-2108.

24. Lo, A. Quantitative proteomics by stable isotope labelling liquid chromatography mass spectrometry. Chapter 4: Automation of 2MEGA labelling chemistry for high throughput proteomics applications. 2011.

25. Wang, N.; Xie, C.; Young, J. B.; Li, L., Off-line two-dimensional liquid chromatography with maximized sample loading to reversed-phase liquid chromatography-electrospray ionization tandem mass spectrometry for shotgun proteome analysis. *Anal. Chem. (Washington, DC, U. S.)* **2009,** 81, (3), 1049-1060.

26. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Anal. Chem. (Washington, DC, U. S.)* **2008,** 80, (12), 4696-4710.

27. Wang, P.; Lo, A.; Young, J.B.; Song, J.; Lai, R.; Kneteman, N.; Hao, C.; Li, L., Targeted quantitative mass spectrometric identification of differentially expressed proteins between Bax-expressing and deficient colorectal carcinoma cells. *J. Proteome Res.* **2009,** 8, (7), 3403-3414.

28. Lo, C. C.; Brownlee, B. G.; Bunce, N. J., Mass spectrometric and toxicological assays of Athabasca oil sands naphthenic acids. *Water Res.* **2006,** 40, (4), 655-664.

29. Lattier, D. L.; Reddy, T. V.; Gordon, D. A.; Lazorchak, J. M.; Smith, M. E.; Williams, D. E.; Wiechman, B.; Flick, R. W.; Miracle, A. L.; Totii, G. P., 17β-ethynylestradiol-induced vitellogenin gene transcription quantified in livers of adult males, larvae, and gills of fathead minnows (Pimephales promelas). *Environ. Toxicol. Chem.* **2002,** 21, (11), 2385-2393.

30. Emmersen, B. K.; Petersen, I. M., Natural occurrence, and experimental induction by estradiol-17-β, of a lipophosphoprotein (vitellogenin) in flounder (Platichtys flesus, L.). *Comp. Biochem. Physiol. B* **1976,** 54, (4B), 443-446.

31. Mundall, E. C.; Tobe, S. S.; Stay, B., Induction of vitellogenin and growth of implanted oocytes in male cockroaches. *Nature (London)* **1979,** 282, (5734), 97-98.

32. Purdom, C. E.; Hardiman, P. A.; Bye, V. J.; Eno, N. C.; Tyler, C. R.; Sumpter, J. P., Estrogenic effects of effluents from sewage treatment works. *Chem. Ecol.* **1994,** 8, (4), 275-285.

33. Ding, J. L.; Lim, E. H.; Lam, T. J., Cortisol-induced hepatic vitellogenin mRNA in Oreochromis aureus (Steindachner). *Gen. Comp. Endocrinol.* **1994,** 96, (2), 276-287.

34. Shi, X.; Zhang, S.; Pang, Q., Vitellogenin is a novel player in defense reactions. *Fish Shellfish Immunol.* **2006,** 20, (5), 769-772.

35. Zhang, S.; Sun, Y.; Pang, Q.; Shi, X., Hemagglutinating and antibacterial activities of vitellogenin. *Fish Shellfish Immunol.* **2005,** 19, (1), 93-95.

36. Tong, Z.; Li, L.; Pawar, R.; Zhang, S., Vitellogenin is an acute phase protein with bacterial-binding and inhibiting activities. *Immunobiology* 215, (11), 898-902.

Chapter 4

Quantitative Proteomic Analysis of Gills from Zebrafish (*Danio rerio*) Exposed to Oil Sands Process Water from the Athabasca Oil Sands in Alberta, Canada*

4.1 Introduction

Oil sands consist of a mixture of sand, clay, water, and crude bitumen, a tar-like form of petroleum. The Athabasca oil sands are one of the three major oil sands deposits in northeastern Alberta, Canada and are the largest global deposit of crude bitumen. Bitumen is typically separated from oil sands using the Clark hot water extraction method resulting in wastewater, known as oil sands process water (OSPW), which is stored in tailings ponds on site in accordance with existing policy.[1] The OSPW contains water, sand, clay, silt, residual bitumen, and dissolved inorganics and organics, including naphthenic acids (NAs), polycyclic aromatic hydrocarbons (PAHs), and heavy metals.[2-5] OSPW is toxic and, because the tailing ponds will eventually be reclaimed, there is interest in understanding the toxicity of OSPW.[5-12] There are numerous studies investigating the toxicity and exposure effects of OSPWs and NAs in bacteria and phytoplankton,[7, 9, 13, 14] plants,[6, 15] frogs,[16, 17] birds,[18, 19] rats,[12] and fish.[3, 20-23]

*A version of this chapter is in preparation for publication as: A. G. De Souza, T. J. MacCormack, L. Li, and G. G. Goss, "A quantitative proteomic analysis of the biological effects on the gills of zebrafish (*Danio rerio*) exposed to commercial naphthenic acids and oil sands process water from the Athabasca oil sands in Alberta, Canada"

Although studies have been conducted using fish, to the best of the authors' knowledge, results from studies involving zebrafish (*Danio rerio*) exposure to OSPW have not been previously reported. The zebrafish is commonly used as a model organism to study the effect of aquatic contaminants.[24-35] Fish require a large surface area for gas and ion exchange with their environment and these functions are largely accomplished by the gills.[36] The gills can make up greater than 50% of the total surface area of the animal[37] and provide a direct route for the uptake of contaminants.[38] In Chapter 2, we established a comprehensive proteome profile of the zebrafish gill and identified over 5700 unique proteins with significant coverage of important biochemical pathways relevant to gill function.[39] Numerous established and potential biomarkers of stress, diseases, and environmental contamination were also expressed in the gill.

The objective of this study was to identify potential environmental markers and to determine mechanisms of toxicity and response for OSPW exposure in zebrafish gill based on differential protein expression analysis. This study utilized a stable isotope labelling strategy involving dimethylation of N-termini of peptides after guanidinylation of lysines (2MEGA),[40] followed by two dimensional liquid chromatography electrospray ionization tandem mass spectrometry (2D LC-ESI MS/MS) analysis. In Chapter 3, an initial study involving exposure of zebrafish to commercially available NAs surprisingly found only a small number of differentially expressed proteins. We hypothesized

that other components in OSPW, in addition to NAs, must play an important role in its toxicity.

4.2 Experimental

4.2.1 Chemicals and Reagents

Tricaine methanesulfonate (TMS) was purchased from Syndel Laboratories Ltd. (Vancouver, BC, Canada). The total protein extraction kit was purchased from Biochain Institute Inc. (Hayward, CA, USA) and the bicinchoninic acid (BCA) assay kit was purchased from Pierce Biotechnology (Rockford, IL, USA). Dithiothreitol (DTT) and iodoacetamide (IAA) were purchased from BioRad (Mississauga, ON, Canada). Sequencing-grade modified trypsin was purchased from Promega (Madison, WI, USA). LC-MS grade acetonitrile (ACN) and water were purchased from Fisher Scientific Canada (Edmonton, AB, Canada). Isotopically-enriched formaldehyde ($^{13}CD_2O$) was purchased from Cambridge Isotope Laboratories (Andover, MA, USA). Oil sands process water from the Athabasca Oil Sands was obtained from the laboratory of Dr. Mohamed Gamal El-Din in the Faculty of Engineering at the University of Alberta. All other chemicals and reagents were purchased from Sigma-Aldrich (Oakville, ON, Canada).

4.2.2 Zebrafish Exposure to Oil Sands Process Water

A total of 54 Nacre strain (A/B background) zebrafish were taken from the Zebrafish Breeding Facility at the University of Alberta Biosciences Aquatics

Facility. In order to maintain consistency in gill harvesting, three replicates of 12 fish from exposed zebrafish and from unexposed controls were staggered over three days. Fish were distributed among 6 L-tanks (4 L water volume) to maintain similar ratios of male and female across the tanks. Fish acclimated for two days in 4 L of Zebrafish Breeding Facility water and, at the end of the acclimation period, one of the two tanks was assigned for exposure and the other as the control. The exposure tank consisted of OSPW diluted 10 times in Zebrafish Breeding Facility water to give a final concentration of 3 mg/L NAs. Fifty percent of the water in the exposed zebrafish tank and the unexposed control tank were exchanged every other day for a period of 7 days with diluted OSPW and Zebrafish Breeding Facility water, respectively. The water temperature in the tanks during the acclimation period and experiment was maintained at 25 ^{o}C. At the end of the experiment, fish were sacrificed on three consecutive days.

4.2.3 Gill Tissue Collection and Protein Extraction

Fish were anaesthetized in a 1.0 g/L solution of tricaine methanesulfonate. To reduce the content of blood proteins in the gill tissue, the circulatory system was flushed with ice-cold phosphate-buffered saline (PBS) containing 50 IU/mL sodium heparin. Perfusion cannulae consisted of a length of PE10 tubing with an 8 mm long tip fabricated from a 30G needle. The cutting edge of the needle was removed and the tip was polished. Fish were positioned ventral-side up on a moistened sponge under a dissecting microscope. The body cavity was then opened and quickly rinsed with ice-cold PBS. The heart was exposed by blunt dissection and one ventricular wall was pierced using #5 forceps. The cannula

was inserted through the ventricle, into the bulbus arteriosus, and held in place by clamping the ventricle wall around it using blunt forceps. The heart was subsequently perfused at 300 µL/min with ice-cold PBS containing 50 IU/mL sodium heparin until gill tissue was visibly cleared of blood (1-5 min).

Following perfusion, whole gill baskets were excised and placed in a flat-bottomed microcentrifuge tube on dry ice, and proteins extracted according to the protocol provided by the manufacturer (Biochain Institute Inc., Hayward, CA) with slight modifications. One mL of 2 mM phenylmethylsulfonly fluoride (dissolved in ethanol) per mL of 50X protease inhibitor (PI) was added to the extraction buffer, and 10 mL of 1X PI added per gram of tissue. Each gill was homogenized on ice using a disposable pestle, and kept on dry ice prior to sonication in an ice-bath, followed by shaking on a vortex-shaker at 4 ^{o}C for 2 h. Post-vortexing, the sample was centrifuged (18 000 g, 4 ^{o}C, 20 min) and the supernatant collected and stored at -80 ^{o}C. A BCA assay was conducted to determine the gill protein concentration using bovine serum albumin as the standard.

The individual gill protein extracts from each tank were pooled to generate three sets of samples (Sets A, B, and C), each consisting of OSPW-exposed sample and a control.

4.2.4 Protein Precipitation and In-Solution Digestion

Proteins were reduced with 900 mM DTT for one hour at 37 °C, cooled at room temperature, and alkylated with 2.2 molar equivalents of 450 mM IAA for 1 h at room temperature in the dark. The concentrations of DTT and IAA were

chosen to reduce sample dilution upon reduction and alkylation. For every 5 µg of protein, 0.225 µmol of DTT and 0.500 µmol of IAA were used to reduce and alkylate the disulfide bonds, respectively. Proteins were then precipitated with four times the volume of acetone (pre-cooled at -80 °C), kept at -20 °C overnight, centrifuged (20 800 g, 4 °C, 10 min), and the supernatant removed and discarded.

The protein pellet was partially solubilized in 100 mM NH_4HCO_3, sonicated in an ice bath for 5 s followed by a 30 s rest on ice (repeated a total of four times), shaken on a vortex-shaker (2 h, 4 °C), and centrifuged (20 800 g, 4 °C, 10 min). The supernatant was collected and diluted to 50 mM NH_4HCO_3. This solubilization process was repeated on the protein pellet. A BCA assay was conducted to determine the protein concentration for each solubilization. Negligible protein was detected in the second NH_4HCO_3 solubilization step, but both the first and second NH_4HCO_3 solubilization samples were pooled and digested. The trypsin-to-protein ratio was 1:25 (w/w), and digestion occurred at 37 °C overnight.

The sequential solubilization scheme in Chapter 2 identified more proteins in the urea solubilized fraction versus the NH_4HCO_3 solubilized fraction. There was concern about carbamylation at lysines and N-termini of peptides, which would affect 2MEGA labelling, therefore we did not use urea as a solubilizing agent. Since then, results from a study have shown that urea induces carbamylation on the lysine of approximately 1% of peptides.[41]

4.2.5 Peptide Desalting and Labelling

Simultaneous salt removal and peptide quantitification was performed on the digests using a Polaris C18A column (Palo Alto, CA, USA, 3 µm, 300 Å, 4.6 x 50 mm) on an Agilent 1100 HPLC system with a UV detector.[42] Mobile Phase A consisted of 0.1% trifluoroacetic acid (TFA) in water and Mobile Phase B consisted of 0.1% TFA in acetonitrile. The elution gradient was 2.5% B for 5.50 min, 2.5-85% B from 5.50 to 5.51 min, 85% B from 5.51 to 15.50 min, 85-2.5% B from 15.50 to 15.51 min, and held at 2.5% B from 15.51 to 25.00 min at a flow rate of 1.000 mL/min. Fraction collection occurred at 7.25 min for a total of 1.00 min.

The desalted peptides were isotopically-labelled using the 2MEGA strategy, where either CH_2O or $^{13}CD_2O$ was used to dimethylate the N-termini of peptides after guanidinylation of lysines.[40] In brief, 100 µL of 2 M O-methylisourea (prepared in 1:1 (v/v) 2 M NaOH and 100 mM $NaHCO_3$) was added to 100 µL of 0.5 µg/µL peptide solution. The pH was adjusted to 11.5 with 2 M NaOH, if necessary, and incubated at 60 ^oC for 20 minutes with intermittent shaking. At the end of the 20 minutes, the samples were removed from heat and pH adjusted to 7 using 10% trifluoroacetic acid (TFA) followed by the addition of 50 µL of 1 M acetate buffer. The dimethylation reaction proceeded with the addition of 4 µL of 1 M sodium cyanoborohydride and 1 µL of 4% (w/w) CH_2O or $^{13}CD_2O$, followed by incubation at 37 ^oC for 20 minutes. At the end of the incubation period, 10% TFA was used to adjust the solution to pH 2. A note of

caution: sodium cyanoborohydride releases hydrogen cyanide gas upon exposure to strong acid.

4.2.6 Strong Cation Exchange Chromatography

The labelled digests were mixed such that a light-labelled control and a heavy-labelled sample constituted the forward case and vice versa for the reverse case for each of Sets A, B, and C. The forward- and reverse-labelled mixtures were separated into individual fractions on a PolySULFOETHYL A column (PolyLC, Columbia, MD, 5 µm, 300 Å, 2.1 x 250 mm) by strong cation exchange (SCX) on an Agilent 1100 HPLC system (Palo Alto, CA). Mobile Phase A consisted of 10 mM KH_2PO_4 in 20% ACN (pH 2.76) and Mobile Phase B consisted of 10 mM KH_2PO_4 and 0.5 M KCl in 20% ACN (pH 2.76). The elution gradient was 0% B for 7 min, 0-6% B from 7 to 8 min, 6-28% B from 8 to 36 min, 28-40% B from 36 to 44 min, 40-60% B from 44 to 49 min, 60-100% B from 49 to 53 min, 100% B from 53 to 58 min, 100-0% B from 58 to 60 min, and held at 0% B for 10 min at a flow rate of 0.200 mL/min. Fractions were collected at 1 min intervals.

4.2.7 LC-ESI QToF MS and MS/MS Analysis

Prior to LC-ESI MS/MS, the collected fractions from the SCX were subjected to salt removal and peptide quantification, as previously described. A modification to the desalting and quantification was the use of a smaller Polaris C18A column (Palo Alto, CA, USA, 3 µm, 300 Å, 2.0 x 50 mm) with an elution gradient of 2.5% B for 5.50 min, 2.5-85% B from 5.50 to 5.51 min, 85% B from 5.51 to 10.00 min, 85-2.5% B from 10.00 to 10.01 min, and held at 2.5% B from

10.01 to 15.00 min at a flow rate of 0.200 ml/min. Fraction collection occurred at 7.00 min for a total of 1.00 min. A smaller sample for desalting facilitated the use of a smaller column. Adjacent fractions from the SCX fractionation were pooled if the total amount of peptides in individual fractions was less than 1 µg.

The desalted SCX peptide samples were analyzed using a Q-Tof Premier™ mass spectrometer equipped with a nanoACQUITY® UPLC system (Waters, Milford, MA). Approximately 1 µg of sample was loaded onto an Atlantis dC18 column (Waters, 3 µm, 100 Å, 75 µm x 150 mm). Mobile Phase A consisted of 0.1% formic acid in water and Mobile Phase B consisted of 0.1% formic acid in acetonitrile. The elution gradient was 2-7% B from 0 to 2 min, 7-20% B from 2 to 85 min, 20-30% B from 85 to 110 min, 30-45% B from 110 to 115 min, 45-90% B from 115 to 120 min, held at 90% B for 5 min, and 90-2% B for 5 min at 250 nL/min. A precursor ion exclusion (PIE) strategy in LC MS/MS was used where ions whose acquired MS/MS spectra resulted in positive peptide identification were excluded from MS/MS acquisition in the subsequent run of the adjacent SCX sample.[43]

4.2.8 Peptide Quantification from MS/MS Data

Mascot Distiller 2.3.2.0 (Matrix Science, London, UK) was used to generate quantification results for each LC-ESI MS/MS run. Raw LC MS/MS data were lock-mass corrected and de-isotoped, and searched using the Mascot search program (Matrix Science, London, UK) within Mascot Distiller.

For each LC-ESI MS/MS run, database searches were restricted to *Danio rerio* in the NCBInr database (containing 9069431 sequences and 3106408609

residues). The following search parameters were selected for all database searches: enzyme, trypsin; maximum missed cleavages, 2; fixed modifications, carbamidomethyl (C), guandinyl (K); variable modifications, dimethylation0 (N-term), dimethylation6 (N-term); peptide tolerance, ± 30 ppm; fragment mass tolerance, ± 0.2 Da; peptide charge, 1+, 2+, 3+; quantitation method, 2MEGA_QuantTest. Mascot Distiller generated a list of quantitated peptide ratios.

A final list of quantitated peptide ratios was generated for each of the forward and reverse cases for each of Biological Replicates A, B, and C using the following criteria: For redundant peptides in each list, the peptide ratio was calculated as a geometric average of the observed Mascot Distiller ratios. Peptides were considered quantitated only if peptide quantitation ratios were obtained in both the forward and reverse cases and the relative difference was equal to or less than 0.5, as previously rationalized.[44] Relative difference was calculated using the following equation: $\frac{\mathbf{2}(\boldsymbol{ratio_f - ratio_r^{-1}})}{(\boldsymbol{ratio_f + ratio_r^{-1}})}$, where $ratio_f$ and $ratio_r$ refer to the peptide quantitation ratios in the forward and reverse cases, respectively. A conservative list of peptides was generated by calculating coefficients of variation (CVs) for each peptide and discarding peptides with CVs greater than 50%. The remaining peptides were manually checked and discarded if the Mascot Distiller peak fitting was incorrect.

The peptides were then organized by proteins according to their accession identification number with their associated quantitation ratio from each biological replicate. Geometric averages were calculated for each protein using the peptide

quantitation ratios from the biological replicates. Proteins were considered differentially-expressed if the protein quantitation ratios were greater than 1.5 or less than 0.67. The threshold for differential abundance represents two standard deviations and 95% confidence determined in a previous study.[44] A final list of quantitated peptide ratios was generated for each of the forward and reverse cases for each of Biological Replicates A, B, and C.

4.3 Results and Discussion

Based on the results of our previous study (Chapter 3) on 3 mg/L NAs exposure on zebrafish, where we found 15 differentially-expressed proteins, we decided to examine the exposure effects of OSPW on zebrafish gill. OSPW contains components, other than NAs, that may contribute to toxicity, such as PAHs and heavy metals.

The NAs concentration in the OSPW used in this study was determined to be 28 mg/L by the laboratory of Dr. Jonathan Martin (Faculties of Medicine and Dentistry and School of Public Health) at the University of Alberta.[45] Based on the NAs concentration, dilutions of OSPW were prepared at 20-, 10-, and 6-fold to visually observe the effects of seven day exposures on zebrafish. The zebrafish in the 6-fold dilution OSPW exposure exhibited slower motility after two days of exposure, while those in the 20- and 10-fold dilutions OSPW exposures did not exhibit mortality or visible changes in motility after a seven day exposure. Consequently, we decided to use a 9.3-fold OSPW dilution exposure for the zebrafish gill protein study for two reasons: 1) low likelihood of death due to

toxicity, and 2) NAs concentration in OSPW corresponded to the commercial NAs concentration (3 mg/L) used in an earlier study (Chapter 3).

The results of this study found 99 proteins that were differentially-expressed upon OSPW exposure, 18 of which were expressed in all three biological replicates (Table 4.1) and 81 of which were expressed in at least two of the three biological replicates (Table 4.2). Most of the differentially-expressed proteins were upregulated in OSPW exposures compared to controls.

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.7 was used to categorize the differentially-expressed proteins by their function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Table 4.3 lists select KEGG pathways and the number of differentially-expressed proteins observed for each pathway.

Protein levels change in response to stimuli and include those proteins involved in repairing tissue damage, protecting against pathogens, and restoring the homeostatic state prior to the stimuli. Through the use of quantitative MS techniques, we observed changes in protein expression upon OSPW exposure.

4.3.1 Tissue Damage and Repair

Metals are among the many constituents of OSPW and studies have shown that exposure to elevated concentrations of metals (e.g. silver, copper, cadmium, zinc) are detrimental to fish.[46-48] The OSPW was sent for metal content analysis by inductively coupled plasma (ICP) MS. Table 4.4 shows the concentration of metals in OSPW compared to the Canadian Water Quality Guidelines for the Protection of Aquatic Life set by the Canadian Council of Ministers of the

Table 4.1. Differentially-expressed proteins quantitated in all three biological replicates.

#	gi|	Protein Name	Peptides			GA[1]	SD_{geo}[2]
			A	B	C		
1	18858197	glutathione S-transferase pi	2	3	1	1.8	1.18
2	18858335	β actin1	1	1	4	1.9	1.22
3	18858427	muscle creatine kinase a	2	2	1	2.5	1.15
4	18859505	alpha-tropomyosin	1	1	1	1.6	1.05
5	21426835	glucose phosphate isomerase a	1	1	2	2.1	1.12
6	28194094	parvalbumin isoform 4a	1	2	2	1.8	1.06
7	28278942	eukaryotic translation elongation factor 2, like	1	1	1	2.3	1.21
8	31419240	ckmb protein	1	1	3	3.0	1.05
9	38707983	aconitate hydratase, mitochondrial	2	2	1	2.3	1.22
10	40363541	adenosylhomocysteinase	1	1	1	3.1	1.06
11	47085883	malate dehydrogenase, mitochondrial	3	1	2	1.8	1.17
12	68435369		1	1	1	2.0	1.34
13	112180551	hpx protein	1	1	1	2.0	1.01
14	169403947	glyceraldehyde-3-phosphate dehydrogenase	1	1	2	2.2	1.14
15	189527793	PREDICTED: AHNAK nucleoprotein	4	2	1	0.56	1.02
16	189529246		1	1	3	2.6	1.17
17	189535920	PREDICTED: filamin A, alpha	1	1	1	1.9	1.05
18	190339564	zgc:194249 protein	1	1	3	1.9	1.11

[1]GA = geometric average of peptide ratios across biological replicates
[2]SD_{geo} = geometric standard deviation

Table 4.2. Differentially-expressed proteins quantitated in two biological replicates.

#	gi|	Protein Name	Peptides			GA[1]	SD_{geo}[2]
			A	B	C		
1	2495341	RecName: Full=Heat shock cognate 71 kDa protein; AltName: Full=Heat shock 70 kDa protein 8	1	0	1	2.1	1.18
2	6630992	NCC receptor protein 1	1	0	1	2.8	1.22
3	6644113	nucleoside diphosphate kinase-Z2	3	2	0	1.7	1.01
4	18858281	apolipoprotein A-I	0	1	2	2.6	1.20
5	18859279	proteasome activator subunit 1	1	1	0	1.6	1.02
6	18859307	ras-related nuclear protein	3	1	0	1.5	1.50
7	22671688	aldolase A	0	2	2	2.1	1.12
8	27464846	transferrin	0	1	2	2.2	1.18
9	27545193	brain creatine kinase b	2	4	0	1.9	1.13
10	27881963	sb:cb825 protein	1	1	0	1.7	1.11
11	28194102	parvalbumin isoform 1c	0	1	1	0.58	1.21
12	29124460	rplp0 protein	1	0	1	1.9	1.11
13	29477118	aldolase b, fructose-bisphosphate	2	1	0	2.0	1.49
14	29612581	zgc:66168 protein	0	1	1	1.6	1.02
15	29648593	glutathione peroxidase	2	1	0	3.0	1.28
16	31795559	troponin T3b, skeletal, fast isoform 2	1	1	0	0.52	1.41
17	34740143	CDC48	1	1	0	2.0	1.25
18	34849746	calr protein	1	1	0	2.0	1.33

[1]GA = geometric average of peptide ratios across biological replicates
[2]SD_{geo} = geometric standard deviation

Table 4.2. Differentially-expressed proteins quantitated in two biological replicates.

#	gi|	Protein Name	Peptides			GA[1]	SD_{geo}[2]
			A	B	C		
19	37590349	enolase 1, (alpha)	1	2	0	2.5	1.02
20	38488700	bisphosphoglycerate mutase 1a	0	1	1	1.6	1.00
21	41053309	hydroxyacylglutathione hydrolase	0	1	1	1.9	1.33
22	41053315	vacuolar protein sorting 29	2	1	0	1.6	1.46
23	41053339	hypothetical protein LOC336637	2	1	0	1.8	1.10
24	41054227	bridging integrator 2	0	1	1	0.63	1.14
25	41054435	tropomyosin 4	1	2	0	3.1	1.00
26	41054651	isocitrate dehydrogenase 2 (NADP+), mitochondrial	3	0	1	1.9	1.17
27	41054980	tubulin cofactor a	1	1	0	0.58	1.02
28	41055439	cell division cycle 42 isoform 1	1	1	0	2.0	1.03
29	41055742	selenium binding protein 1	2	1	0	1.9	1.05
30	41056147	dynein light chain 2	1	1	0	5.0	1.00
31	41152342	sarcomeric mitochondrial creatine kinase	0	1	1	1.9	1.25
32	41393141	tropomyosin 3	2	0	3	2.0	1.04
33	42542740	vat1 protein	0	2	1	2.6	1.03
34	45387521	actin related protein 2/3 complex subunit 4	1	0	1	1.6	1.33
35	45387531	NAD(P)H dehydrogenase, quinone 1	0	2	2	3.3	1.06
36	45829832	nsfl1c protein	1	1	0	1.7	1.02

[1]GA = geometric average of peptide ratios across biological replicates
[2]SD_{geo} = geometric standard deviation

Table 4.2. Differentially-expressed proteins quantitated in two biological replicates.

#	gi|	Protein Name	Peptides			GA[1]	SD_{geo}[2]
			A	B	C		
37	47085823	medium-chain acyl-CoA dehydrogenase isoform 1	1	1	0	1.7	1.14
38	47085887	MARCKS-like 1	0	2	1	2.3	1.40
39	47086573	methyl CpG binding protein 2	0	1	1	1.6	1.05
40	47086793	alpha-2-HS-glycoprotein	0	1	2	1.9	1.12
41	47086875	glutamate dehydrogenase 1a	1	1	0	1.8	1.10
42	47087353	heterogeneous nuclear ribonucleoprotein D-like	1	1	0	1.5	1.18
43	47174751	non-muscle cofilin 1	1	0	1	2.1	1.24
44	47271422	triosephosphate isomerase 1b	1	2	0	2.3	1.25
45	47550733	SERPINE1 mRNA binding protein 1	1	0	1	1.8	1.20
46	47551327	hypothetical protein LOC407981	0	1	1	0.49	1.27
47	49618989	nucleolin	0	1	2	3.2	1.12
48	50344790	hypothetical protein LOC415158	2	0	1	0.57	1.09
49	50344894	tropomyosin 2 (beta)	1	0	1	2.2	1.17
50	50539990	thioredoxin	1	0	1	2.7	1.02
51	50539996	peroxiredoxin 2	0	2	1	2.1	1.02
52	51011067	pyruvate kinase, muscle, b	4	1	0	1.5	1.11
53	51468069	proteasome (prosome, macropain) subunit, beta type, 1	0	1	1	3.6	1.18
54	55925442	3-oxoacid CoA transferase 1a	1	1	0	1.8	1.18

[1]GA = geometric average of peptide ratios across biological replicates
[2]SD_{geo} = geometric standard deviation

Table 4.2. Differentially-expressed proteins quantitated in two biological replicates.

#	gi|	Protein Name	Peptides			GA[1]	SD_{geo}[2]
			A	B	C		
55	56790262	superoxide dismutase 1, soluble	0	1	1	3.2	1.16
56	57525607	GPN-loop GTPase 1	0	1	1	3.0	1.04
57	58801528	SET translocation (myeloid leukemia-associated) A	0	1	1	1.6	1.05
58	62955567	cold inducible RNA binding protein	0	1	1	1.7	1.12
59	66472396	hypothetical protein LOC553706	1	0	1	0.64	1.05
60	66472640	glutathione reductase	1	2	0	1.8	1.15
61	66472750	dihydropyrimidinase-like 3	0	1	1	1.8	1.27
62	80751159	hypothetical protein LOC641325	1	3	0	0.58	1.52
63	113680776	EF-hand domain family, member D2	1	0	2	2.5	1.33
64	113865949	zgc:161979	1	2	0	0.61	1.04
65	115529393	hypothetical protein LOC767787	0	1	1	6.3	1.27
66	116004521	coagulation factor XIII, A1 polypeptide	0	3	1	3.3	1.27
67	116325975	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1, cardiac muscle	2	0	1	3.3	1.11
68	118150552	FK506 binding protein 3, 25kDa isoform 2	0	1	1	0.55	1.05
69	123702889	hypothetical protein LOC791196 3	1	1	0	1.5	1.29
70	125846520	PREDICTED: RETRotransposon-like family member (retr-1)-like 3	2	0	1	1.6	1.25
71	134024899	zgc:162944 protein	0	2	2	1.6	1.30
72	160333126	aldehyde dehydrogenase 5 family, member A1 (succinate-semialdehyde dehydrogenase)	0	1	2	1.9	1.07

[1]GA = geometric average of peptide ratios across biological replicates
[2]SD_{geo} = geometric standard deviation

Table 4.2. Differentially-expressed proteins quantitated in two biological replicates.

#	gi|	Protein Name	Peptides			GA[1]	SD_{geo}[2]
			A	B	C		
73	167621556	polymerase I and transcript release factor isoform 2	2	0	1	0.56	1.04
74	168823552	hypothetical protein LOC322909	0	2	2	3.4	1.03
75	187608218	phosphoglucomutase 5	2	0	1	1.6	1.04
76	189521576	PREDICTED: similar to Non-muscle caldesmon (CDM) (L-caldesmon)	0	1	1	0.53	1.25
77	189524946	PREDICTED: similar to Rho family guanine-nucleotide exchange factor	1	2	0	2.1	1.42
78	189526614	PREDICTED: similar to toxin-1	1	1	0	0.39	1.48
79	189530381	PREDICTED: hypothetical protein LOC100000597	1	0	1	3.2	1.11
80	189530625	PREDICTED: spectrin repeat containing, nuclear envelope 1-like 3	0	1	3	5.1	1.34
81	193788703	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), beta polypeptide	0	1	1	0.56	1.09

[1]GA = geometric average of peptide ratios across biological replicates
[2]SD_{geo} = geometric standard deviation

Table 4.3. Select KEGG pathways.

KEGG Pathway	# Proteins Quantified & Differentially-Expressed
Glycolysis / Gluconeogenesis	8
Arginine and proline metabolism	4
Cardiac muscle contraction	4
Glutathione metabolism	4
Regulation of actin cytoskeleton	4
MAPK signaling pathway	3
Pentose phosphate pathway	3
Pyruvate metabolism	3
Tight junction	3
Citrate cycle (TCA cycle)	3
Adherens junction	2
Alanine, aspartate and glutamate metabolism	2
Butanoate metabolism	2
Endocytosis	2
Focal adhesion	2
Fructose and mannose metabolism	2
PPAR signaling pathway	2
Proteasome	2
Purine metabolism	2
Ribosome	2
Valine, leucine and isoleucine degradation	2
Glyoxylate and dicarboxylate metabolism	2
Amino sugar and nucleotide sugar metabolism	1

MAPK, mitogen-activated protein kinase; TCA, tricarboxylic acid; PPAR, peroxisome proliferator-activated receptor; GnRH, gonadotropin-releasing hormone; VEGF, vascular endothelial growth factor

Table 4.3. Select KEGG pathways

KEGG Pathway	# Proteins Quantified & Differentially-Expressed
Apoptosis	1
Arachidonic acid metabolism	1
beta-Alanine metabolism	1
Cysteine and methionine metabolism	1
D-Glutamine and D-glutamate metabolism	1
Drug metabolism	1
Fatty acid metabolism	1
Fructose and mannose metabolism	1
GnRH signaling pathway	1
Inositol phosphate metabolism	1
Metabolism of xenobiotics by cytochrome P450	1
Nitrogen metabolism	1
Oxidative phosphorylation	1
p53 signaling pathway	1
Propanoate metabolism	1
Pyrimidine metabolism	1
RNA degradation	1
Selenoamino acid metabolism,	1
Spliceosome	1
Starch and sucrose metabolism	1
Synthesis and degradation of ketone bodies	1
Vascular smooth muscle contraction	1
VEGF signaling pathway	1

MAPK, mitogen-activated protein kinase; TCA, tricarboxylic acid; PPAR, peroxisome proliferator-activated receptor; GnRH, gonadotropin-releasing hormone; VEGF, vascular endothelial growth factor

Environment (CCME).[49] As can be seen in Table 4.4, the OSPW concentration is above the guidelines for a number of metals (e.g. cadmium, copper, iron, molybdenum, and zinc). Therefore, one might expect to observe metal effects on the zebrafish gills.

The gill has an important role in gas and ion exchange between the internal and external environments of fish and in maintaining physiological homeostasis. Metals bind to ion transport sites in the gill epithelium, thereby inhibiting ion transport across the epithelium, resulting in damage to the epithelium, and disrupting ionoregulation and acid-base regulation. Our data showed evidence of damage to the structural integrity of the gill through the upregulation of proteins involved in pathways related to maintaining tissue architecture. These pathways include the adherens junction, focal adhesion, regulation of actin cytoskeleton, and tight junction pathways (Table 4.5). The cell attempts to repair damage and an upregulation of proteins in these pathways could indicate repair to tissue damage.

Once the structural integrity of the gill is compromised, so is the function of the gill. Tissue damage may result in leaky epithelial cells, thereby affecting the gill's ability to maintain ion and pH balance. Consequently, metabolism increases to compensate for ion loss, as observed through the upregulation of numerous proteins involved in metabolism pathways, as listed in Table 4.6.

Tissue restructuring is an energetically-demanding process and we found eight upregulated proteins involved in the glycolysis/gluconeogenesis pathway

Table 4.4 Metal content concentrations in OSPW compared to Canadian Water Quality Guidelines for the Protection of Aquatic Life set by the CCME.

Metal	OSPW Concentration (µg/L)	CCME Guideline (µg/L)
Arsenic	5.04	5
Cadmium	0.31	0.01
Chromium[1]	21.76	1 9
Copper	9.55	4
Iron	391	300
Molybdenum	153	73
Silver	0.04	0.1
Thallium	0.8	0.8
Uranium	7.94	15
Zinc	16.9	30

CCME guideline of 1 and 9 µg/L for chromium(VI) and chromium(III), respectively.

Table 4.5. Differentially-expressed proteins involved in structure remodelling.

KEGG Pathway	gi|	Protein Name	Differential Expression
Adherens junction	18858335	β actin1	1.9
Adherens junction	41055439	cell division cycle 42 isoform 1	2.0
Regulation of actin cytoskeleton	41055439	cell division cycle 42 isoform 1	2.0
Regulation of actin cytoskeleton	45387521	actin related protein 2/3 complex subunit 4	1.6
Regulation of actin cytoskeleton	18858335	β actin1	1.9
Regulation of actin cytoskeleton	47174751	non-muscle cofilin 1	2.1
Tight junction	18858335	β actin1	1.9
Tight junction	41055439	cell division cycle 42 isoform 1	2.0
Focal adhesion	18858335	β actin1	1.9
Focal adhesion	41055439	cell division cycle 42 isoform 1	2.0

Table 4.6. Differentially-expressed proteins involved in metabolism.

KEGG Pathway	gi|	Protein Name	Differential Expression
Alanine, aspartate and glutamate metabolism	160333126	aldehyde dehydrogenase 5 family, member A1 (succinate-semialdehyde dehydrogenase)	1.9
Alanine, aspartate and glutamate metabolism	47086875	glutamate dehydrogenase 1a	1.8
Amino sugar and nucleotide sugar metabolism	21426835	glucose phosphate isomerase a	2.1
Arachidonic acid metabolism	29648593	glutathione peroxidase	3.0
Arginine and proline metabolism	47086875	glutamate dehydrogenase 1a	1.8
Arginine and proline metabolism	27545193	brain creatine kinase b	1.9
Arginine and proline metabolism	18858427	muscle creatine kinase a	2.5
Arginine and proline metabolism	41152342	sarcomeric mitochondrial creatine kinase	1.9
β-Alanine metabolism	47085823	medium-chain acyl-CoA dehydrogenase isoform 1	1.7
Butanoate metabolism	55925442	3-oxoacid CoA transferase 1a	1.8
Butanoate metabolism	160333126	aldehyde dehydrogenase 5 family, member A1 (succinate-semialdehyde dehydrogenase)	1.9
Cysteine and methionine metabolism	40363541	adenosylhomocysteinase	3.1
D-Glutamine and D-glutamate metabolism	47086875	glutamate dehydrogenase 1a	1.8
Drug metabolism	18858197	glutathione S-transferase pi	1.8
Fatty acid metabolism	47085823	medium-chain acyl-CoA dehydrogenase isoform 1	1.7
Fructose and mannose metabolism	47271422	triosephosphate isomerase 1b	2.3

Table 4.6. Differentially-expressed proteins involved in metabolism.

KEGG Pathway	gi|	Protein Name	Differential Expression
Fructose and mannose metabolism	22671688	aldolase A	2.1
Fructose and mannose metabolism	29477118	aldolase b, fructose-bisphosphate	2.0
Glutathione metabolism	18858197	glutathione S-transferase pi	1.8
Glutathione metabolism	29648593	glutathione peroxidase	3.0
Glutathione metabolism	41054651	isocitrate dehydrogenase 2 (NADP+), mitochondrial	1.9
Glutathione metabolism	66472640	glutathione reductase	1.8
Glyoxylate and dicarboxylate metabolism	47085883	malate dehydrogenase, mitochondrial	1.8
Glyoxylate and dicarboxylate metabolism	38707983	aconitate hydratase, mitochondrial	2.3
Nitrogen metabolism	47086875	glutamate dehydrogenase 1a	1.8
Propanoate metabolism	47085823	medium-chain acyl-CoA dehydrogenase isoform 1	1.7
Pyruvate metabolism	51011067	pyruvate kinase, muscle, b	1.5
Pyruvate metabolism	47085883	malate dehydrogenase, mitochondrial	1.8
Pyruvate metabolism	41053309	hydroxyacylglutathione hydrolase	1.9
Selenoamino acid metabolism	40363541	adenosylhomocysteinase	3.0
Starch and sucrose metabolism	21426835	glucose phosphate isomerase a	2.1
Metabolism of xenobiotics by cytochrome P450	18858197	glutathione S-transferase pi	1.8

(Table 4.7) and Figure 4.1 shows the steps in which these proteins are involved in glycolysis.

4.3.2 Oxidative Stress

Metals and PAHs, both known constituents of OSPW, are known to induce oxidative stress through the production of reactive oxygen species, such as free radicals and peroxides.[59-55] We observed signs of response to oxidative stress in gills through the upregulation of superoxide dismutase and proteins involved in glutathione metabolism (Table 4.8). Some of these proteins, such as superoxide dismutase, glutathione peroxidase, and NAD(P)H dehydrogenase, quinone 1 were upregulated in the presence of OSPW by at least three-fold.

Superoxide dismutase is an enzyme that repairs cellular damage caused by the free radical superoxide.[56] We observed an upregulation of 3.2-fold of superoxide dismutase, suggesting that OSPW increases superoxide generation and the cell compensates by increasing expression of superoxide dismutase.

Glutathione peroxidase, glutathione-S-transferase, and NAD(P)H dehydrogenase, quinone 1 are detoxifying enzymes that protect against oxidative stress.[57, 58] We observed upregulation of these proteins, along with two other glutathione metabolism pathway proteins (isocitrate dehydrogenase and glutathione reductase).

4.3.3 Immunoregulation

Tissue trauma or invasion by pathogens or parasites induces changes in the quantity of several macromolecules.[59] Based on our evidence supporting gill epithelial damage, it is not surprising that we also detected changes in the

Table 4.7. Differentially-expressed proteins involved in energy production.

KEGG Pathway	gi|	Protein Name	Differential Expression
Glycolysis / Gluconeogenesis	22671688	aldolase A	2.1
Glycolysis / Gluconeogenesis	29477118	aldolase b, fructose-bisphosphate	2.0
Glycolysis / Gluconeogenesis	37590349	enolase 1, (alpha)	2.5
Glycolysis / Gluconeogenesis	21426835	glucose phosphate isomerase a	2.1
Glycolysis / Gluconeogenesis	51011067	pyruvate kinase, muscle, b	1.5
Glycolysis / Gluconeogenesis	47271422	triosephosphate isomerase 1b	2.3
Glycolysis / Gluconeogenesis	169403947	glyceraldehyde-3-phosphate dehydrogenase	2.2
Glycolysis / Gluconeogenesis	38488700	bisphosphoglycerate mutase 1a	1.6
Citrate cycle (TCA cycle)	47085883	malate dehydrogenase, mitochondrial	1.8
Citrate cycle (TCA cycle)	38707983	aconitate hydratase, mitochondrial	2.3
Citrate cycle (TCA cycle)	41054651	isocitrate dehydrogenase 2 (NADP+), mitochondrial	1.9

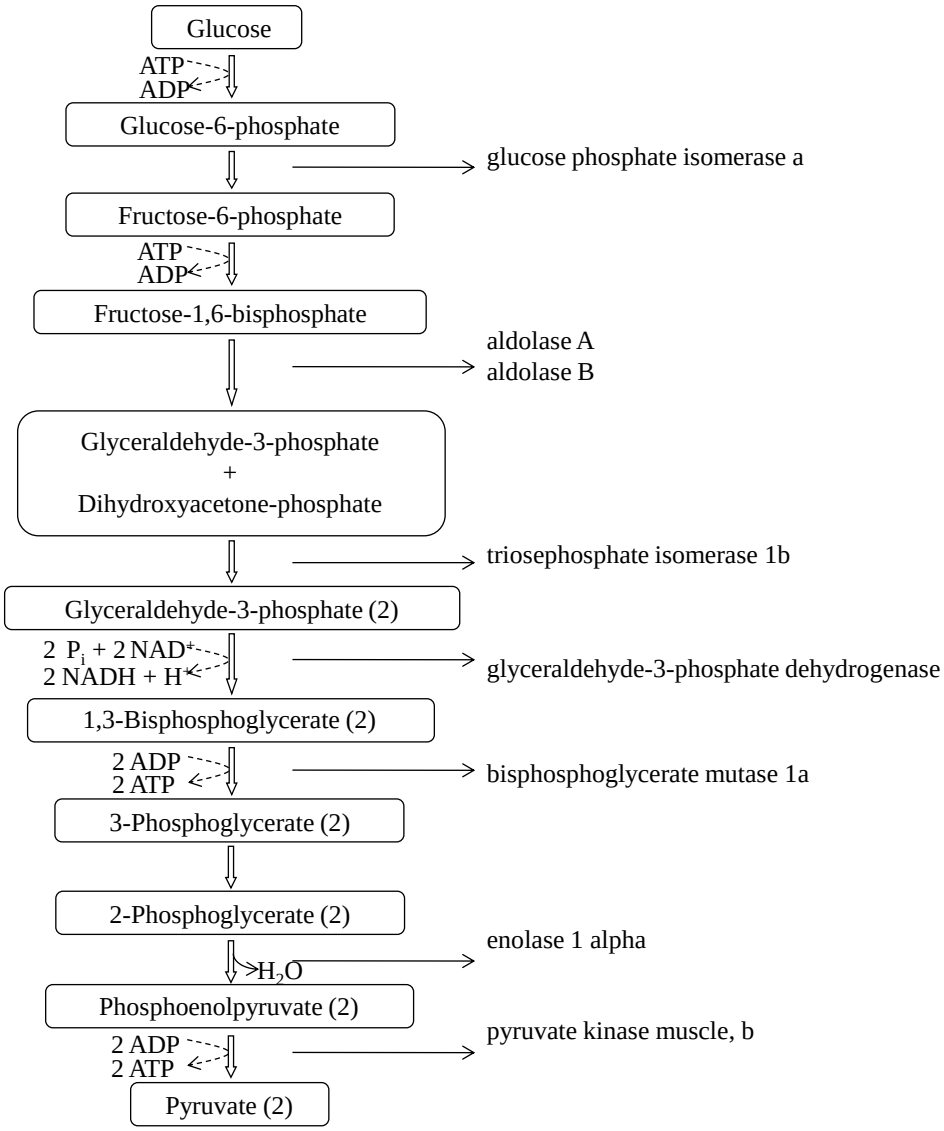

Figure 4.1. Zebrafish gill upregulated proteins involved in glycolysis.

Table 4.8. Differentially-expressed proteins involved in oxidative stress response.

gi|	Protein Name	Differential Expression
18858197	glutathione S-transferase pi	1.8
66472640	glutathione reductase	1.8
41054651	isocitrate dehydrogenase 2 (NADP+), mitochondrial	1.9
29648593	glutathione peroxidase	3.0
56790262	superoxide dismutase	3.2
45387531	NAD(P)H dehydrogenase, quinone 1	3.3

expression of proteins, such as transferrin (gi|27464846) and FK506 binding protein 3 (gi|118150552), involved in immune response.

Transferrin is a glycoprotein that binds and transports iron to cells.[60] Binding of iron to transferrin limits the availability of iron to pathogens, thereby impeding pathogen survival.[59] The upregulation of transferrin in the OSPW exposures could be attributed to an increase in invading pathogens, thereby triggering an increase in immune cells and tissue macrophages or to the increased levels of iron in OSPW. The gill plays an important role in filtering, and tissue damage compromises the gill's ability to protect against invading pathogens. Therefore, increased levels of transferrin decrease the iron available for pathogen survival. Takenka *et al.* detected elevated transferrin levels in cerebral spinal fluid after subarachnoid hemmhorage in patients.[61]

FK506 binding protein 3 (FKBP3), as the name suggests, binds FK506, an immunosuppressant drug used to reduce the risk of organ rejection after a transplant.[62] FKBP3 protein also functions as a molecular chaperone.[63] Studies have observed both upregulation[64, 65] and downregulation[66] of FKBP3 after stress. Zohlnhöfer *et al*. observed upregulation of FKBP3 in patients with neointima compared to controls[64] and Weber *et al*. observed upregulation of the gene coding for FKBP3 proteins with heat shock of the parasite *Entamocba histolytica*.[65] Conversely, Gómez *et al*. showed downregulation of FKBP3 upon mimicking cardiac arrhythmias in adult rat ventricular myocytes.[66] In zebrafish gill exposed to OSPW, a downregulation of FKBP3 was observed. In T-cell lymphocytes, bound FKBP3 inhibits calcium-dependent T-cell activation, thereby functioning

as an immunosuppressor.[67] The observed downregulation of FKBP3 upon OSPW exposure may correspond to a decrease in T-cell lymphocytes.

Although the mechanism of action by which OSPW invokes an immune response is unknown, the differential expression of transferrin and FKBP3 is supportive of an immunoregulatory response to OSPW exposure.

4.3.4 Endocrine Disrupting Effects

There is documentation of endocrine disrupting effects upon exposure to OSPW in slimy sculpin (*Cottus cognatus*),[23] yellow perch (*Perca flavescens*),[68] goldfish (*Carassius auratus*)[3] and fathead minnow (*Pimephales promelas*).[69, 70] In a study involving exposure of adult yellow perch to OSPW, van de Heuvel *et al*. discovered reduced plasma sex steroid hormone levels.[68] Similarly, Lister *et al.* also found testosterone and 17 β-estradiol plasma levels reduced in both male and female goldfish exposed to OSPW, indicating the potential of OSPW to disrupt normal endocrine functioning, but they could not reproduce the effects using NAs extracted from OSPW, suggesting that NAs alone may not mediate endocrine disruption.[3]

Although our study on NAs exposure exhibited potential disruption to reproductive function through the downregulation of vitellogenin proteins (Chapter 3), we were unable to gauge endocrine-disrupting effects in the OSPW exposures. However, our OSPW study was not designed for specific target analysis, so we cannot make the conclusion that OSPW did not affect reproductive function.

4.3.5 Potential Novel Biomarkers

Of the 99 differentially-expressed proteins, 22% were described as hypothetical, predicted, an alphanumeric code (e.g. Zebrafish Genome Collection (ZGC), or no description found. Fourteen of these description-lacking proteins were upregulated and eight were downregulated, but the absence of annotative information limits the information we can glean on the mechanism of action of OSPW on gill as we cannot classify the protein function. The proteins gi|115529393 and gi|189530625 displayed the greatest differential expression, with an upregulation greater than five-fold, but the lack of descriptive names made it difficult to classify these proteins in a meaningful biological context. These proteins do, however, have potential as novel commercial biomarkers to screen for OSPW exposure. The process to commercialization involves four essential components: discovery, qualification, verification, and validation.[71] The discovery process involves identifying candidate biomarkers and qualification involves confirming the differential abundance of candidates. In the verification stage, the specificity of candidates in assessed, and validation consists of establishing sensitivity and specificity.

4.4 Conclusions

Zebrafish exposed to diluted OSPW from the Athabasca Oil Sands exhibited changes in gill protein expression compared to controls. These changes were quantitated using a stable isotope labelling method involving dimethylation of N-termini of peptides after guanidinylation of lysine residues and analysis by

2D LC-ESI MS/MS. Ninety-nine of the quantitated proteins were differentially-expressed in the OSPW-exposed zebrafish gills compared to controls. A number of these proteins are of significance in structure damage and repair, oxidative stress response, immunoregulation, and endocrine-disrupting effects. There is the potential to use a panel of these differentially-expressed proteins as biomarkers for diagnosis and monitoring of OSPW exposure.

4.5 Literature Cited

1. FTFC (Fine Tailings Fundamentals Consortium), *Advances in oil sands tailings research*. Alberta Department of Energy: Edmonton, 1995; Vol. II.

2. Frank, R. A.; Fischer, K.; Kavanagh, R.; Burnison, B. K.; Arsenault, G.; Headley, J. V.; Peru, K. M.; Van Der Kraak, G.; Solomon, K. R., Effect of carboxylic acid content on the acute toxicity of oil sands naphthenic acids. *Environ. Sci. Technol.* **2009,** 43, (2), 266-271.

3. Lister, A.; Nero, V.; Farwell, A.; Dixon, D. G.; Van Der Kraak, G., Reproductive and stress hormone levels in goldfish (Carassius auratus) exposed to oil sands process-affected water. *Aquat. Toxicol.* **2008,** 87, (3), 170-177.

4. MacKinnon, M. D., Development of the tailings pond at Syncrude's oil sands plant: 1978-1987. *AOSTRA J. Res.* **1989,** 5, (2), 109-33.

5. Madill, R. E. A.; Orzechowski, M. T.; Chen, G.; Brownlee, B. G.; Bunce, N. J., Preliminary risk assessment of the wet landscape option for reclamation of oil sands mine tailings: bioassays with mature fine tailings pore water. *Environ. Toxicol.* **2001,** 16, (3), 197-208.

6. Armstrong, S. A.; Headley, J. V.; Peru, K. M.; Germida, J. J., Differences in phytotoxicity and dissipation between ionized and nonionized oil sands naphthenic acids in wetland plants. *Environ. Toxicol. Chem.* **2009,** 28, (10), 2167-2174.

7. Clemente, J. S.; MacKinnon, M. D.; Fedorak, P. M., Aerobic biodegradation of two commercial naphthenic acids preparations. *Environ. Sci. Technol.* **2004,** 38, (4), 1009-1016.

8. Herman, D. C.; Fedorak, P. M.; MacKinnon, M. D.; Costerton, J. W., Biodegradation of naphthenic acids by microbial populations indigenous to oil sands tailings. *Can. J. Microbiol.* **1994,** 40, (6), 467-77.

9. Holowenko, F. M.; MacKinnon, M. D.; Fedorak, P. M., Characterization of naphthenic acids in oil sands wastewaters by gas chromatography-mass spectrometry. *Water Res.* **2002,** 36, (11), 2843-2855.

10. Lai, J. W. S.; Pinto, L. J.; Kiehlmann, E.; Bendell-Young, L. I.; Moore, M. M., Factors that affect the degradation of naphthenic acids in oil sands wastewater by indigenous microbial communities. *Environ. Toxicol. Chem.* **1996,** 15, (9), 1482-1491.

11. MacKinnon, M. D.; Boerger, H., Description of two treatment methods for detoxifying oil sands tailings pond water. *Water Pollut. Res. J. Can.* **1986,** 21, (4), 496-512.

12. Rogers, V. V.; Wickstrom, M.; Liber, K.; MacKinnon, M. D., Acute and subchronic mammalian toxicity of naphthenic acids from oil sands tailings. *Toxicol. Sci.* **2002,** 66, (2), 347-355.

13. Frank, R. A.; Kavanagh, R.; Burnison, B. K.; Arsenault, G.; Headley, J. V.; Peru, K. M.; Van Der Kraak, G.; Solomon, K. R., Toxicity assessment of collected fractions from an extracted naphthenic acid mixture. *Chemosphere* **2008,** 72, (9), 1309-1314.

14. Leung, S. S.; MacKinnon, M. D.; Smith, R. E. H., The ecological effects of naphthenic acids and salts on phytoplankton from the Athabasca oil sands region. *Aquat. Toxicol.* **2003,** 62, (1), 11-26.

15. Crowe, A. U.; Han, B.; Kermode, A. R.; Bendell-Young, L. I.; Plant, A. L., Effects of oil sands effluent on cattail and clover: photosynthesis and the level of stress proteins. *Environ. Pollut. (Oxford, U. K.)* **2001,** 113, (3), 311-322.

16. Hersikorn, B. D.; Ciborowski, J. J. C.; Smits, J. E. G., The effects of oil sands wetlands on wood frogs (Rana sylvatica). *Toxicol. Environ. Chem.* 92, (8), 1513-1527.

17. Pollet, I.; Bendell-Young, L. I., Amphibians as indicators of wetland quality in wetlands formed from oil sands effluent. *Environ. Toxicol. Chem.* **2000,** 19, (10), 2589-2597.

18. Gentes, M.-L.; Waldner, C.; Papp, Z.; Smits, J. E. G., Effects of oil sands tailings compounds and harsh weather on mortality rates, growth and

detoxification efforts in nestling tree swallows (Tachycineta bicolor). *Environ. Pollut. (Amsterdam, Neth.)* **2006,** 142, (1), 24-33.

19. Harms, N. J.; Fairhurst Graham, D.; Bortolotti Gary, R.; Smits Judit, E. G., Variation in immune function, body condition, and feather corticosterone in nestling tree swallows (Tachycineta bicolor) on reclaimed wetlands in the Athabasca oil sands, Alberta, Canada. *Environ Pollut* 158, (3), 841-8.

20. Colavecchia, M. V.; Backus, S. M.; Hodson, P. V.; Parrott, J. L., Toxicity of oil sands to early life stages of fathead minnows (Pimephales promelas). *Environ. Toxicol. Chem.* **2004,** 23, (7), 1709-1718.

21. Farwell, A.; Nero, V.; Croft, M.; Bal, P.; Dixon, D. G., Modified Japanese medaka embryo-larval bioassay for rapid determination of developmental abnormalities. *Arch. Environ. Contam. Toxicol.* **2006,** 51, (4), 600-607.

22. Peters, L. E.; MacKinnon, M.; Van Meer, T.; van den Heuvel, M. R.; Dixon, D. G., Effects of oil sands process-affected waters and naphthenic acids on yellow perch (Perca flavescens) and Japanese medaka (Oryzias latipes) embryonic development. *Chemosphere* **2007,** 67, (11), 2177-2183.

23. Tetreault, G. R.; McMaster, M. E.; Dixon, D. G.; Parrott, J. L., Using reproductive endpoints in small forage fish species to evaluate the effects of athabasca oil sands activities. *Environ. Toxicol. Chem.* **2003,** 22, (11), 2775-2782.

24. Carvan, M. J., III; Solis, W. A.; Gedamu, L.; Nebert, D. W., Activation of transcription factors in zebrafish cell cultures by environmental pollutants. *Arch. Biochem. Biophys.* **2000,** 376, (2), 320-327.

25. Cheng, J.; Flahaut, E.; Cheng, S. H., Effect of carbon nanotubes on developing zebrafish (Danio rerio) embryos. *Environ. Toxicol. Chem.* **2007,** 26, (4), 708-716.

26. Henry, T. R.; Spitsbergen, J. M.; Hornung, M. W.; Abnet, C. C.; Peterson, R. E., Early life stage toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in zebrafish (Danio rerio). *Toxicol. Appl. Pharmacol.* **1997,** 142, (1), 56-68.

27. Huang, H.; Huang, C.; Wang, L.; Ye, X.; Bai, C.; Simonich, M. T.; Tanguay, R. L.; Dong, Q., Toxicity, uptake kinetics and behavior assessment in zebrafish embryos following exposure to perfluorooctanesulphonic acid (PFOS). *Aquat. Toxicol.* 98, (2), 139-147.

28. Incardona, J. P.; Day, H. L.; Collier, T. K.; Scholz, N. L., Developmental toxicity of 4-ring polycyclic aromatic hydrocarbons in zebrafish is

differentially dependent on AH receptor isoforms and hepatic cytochrome P4501A metabolism. *Toxicol. Appl. Pharmacol.* **2006,** 217, (3), 308-321.

29. Jorgensen, A.; Andersen, O.; Bjerregaard, P.; Rasmussen, L. J., Identification and characterization of an androgen receptor from zebrafish Danio rerio. *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2007,** 146C, (4), 561-568.

30. Kausch, U.; Alberti, M.; Haindl, S.; Budczies, J.; Hock, B., Biomarkers for exposure to estrogenic compounds: gene expression analysis in zebrafish (Danio rerio). *Environ. Toxicol.* **2008,** 23, (1), 15-24.

31. Kusik, B. W.; Carvan, M. J., III; Udvadia, A. J., Detection of mercury in aquatic environments using EPRE reporter zebrafish. *Mar. Biotechnol.* **2008,** 10, (6), 750-757.

32. Ramirez, O. A. B.; Garcia, F. P., Genotoxic damage in zebra fish (Danio rerio) by arsenic in waters from Zimapan, Hidalgo, Mexico. *Mutagenesis* **2005,** 20, (4), 291-295.

33. Schnurstein, A.; Braunbeck, T., Tail moment versus tail length--application of an in vitro version of the comet assay in biomonitoring for genotoxicity in native surface waters using primary hepatocytes and gill cells from zebrafish (Danio rerio). *Ecotoxicol. Environ. Saf.* **2001,** 49, (2), 187-96.

34. Sisman, T.; Geyikoglu, F.; Atamanalp, M., Early life-stage toxicity in zebrafish (Danio rerio) following embryonal exposure to selected polychlorinated biphenyls. *Toxicol. Ind. Health* **2007,** 23, (9), 529-536.

35. Wu, S. M.; Zheng, Y. D.; Kuo, C.-H., Expression of mt2 and smt-B upon cadmium exposure and cold shock in zebrafish (Danio rerio). *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2008,** 148C, (2), 184-193.

36. Perry, S. F., The chloride cell: structure and function in the gills of freshwater fishes. *Annu. Rev. Physiol.* **1997,** 59, 325-347.

37. Rombough, P. J.; Moroz, B. M., The scaling and potential importance of cutaneous and branchial surfaces in respiratory gas exchange in larval and juvenile walleye Stizostedion vitreum. *J. Exp. Biol.* **1997,** 200, (18), 2459-2468.

38. Wood, C. M.; Kelly, S. P.; Zhou, B.; Fletcher, M.; O'Donnell, M.; Eletti, B.; Part, P., Cultured gill epithelia as models for the freshwater fish gill. *Biochim. Biophys. Acta, Biomembr.* **2002,** 1566, (1-2), 72-83.

39. De Souza, A. G.; MacCormack, T. J.; Wang, N.; Li, L.; Goss, G. G., Large-scale proteome profile of the zebrafish (Danio rerio) gill for physiological and biomarker discovery studies. *Zebrafish* **2009,** 6, (3), 229-238.

40. Ji, C.; Guo, N.; Li, L., Differential dimethyl labeling of N-termini of peptides after guanidination for proteome analysis. *J. Proteome Res.* **2005,** 4, (6), 2099-2108.

41. Lo, A. Quantitative proteomics by stable isotope labelling liquid chromatography mass spectrometry. Chapter 4: Automation of 2MEGA labelling chemistry for high throughput proteomics applications. 2011.

42. Wang, N.; Xie, C.; Young, J. B.; Li, L., Off-line two-dimensional liquid chromatography with maximized sample loading to reversed-phase liquid chromatography-electrospray ionization tandem mass spectrometry for shotgun proteome analysis. *Anal. Chem. (Washington, DC, U. S.)* **2009,** 81, (3), 1049-1060.

43. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Anal. Chem. (Washington, DC, U. S.)* **2008,** 80, (12), 4696-4710.

44. Wang, P.; Lo, A.; Young, J.B.; Song, J.; Lai, R.; Kneteman, N.; Hao, C.; Li, L., Targeted quantitative mass spectrometric identification of differentially expressed proteins between Bax-expressing and deficient colorectal carcinoma cells. *J. Proteome Res.* **2009,** 8, (7), 3403-3414.

45. Martin, J. W.; Han, X.; Peru, K. M.; Headley, J. V., Comparison of high- and low-resolution electrospray ionization mass spectrometry for the analysis of naphthenic acid mixtures in oil sands process water. *Rapid Communications in Mass Spectrometry* **2008,** 22, (12), 1919-1924.

46. Bury, N. R.; Grosell, M.; Grover, A. K.; Wood, C. M., ATP-dependent silver transport across the basolateral membrane of rainbow trout gills. *Toxicol. Appl. Pharmacol.* **1999,** 159, (1), 1-8.

47. McGeer, J. C.; Szebedinszky, C.; Gordon McDonald, D.; Wood, C. M., Effects of chronic sublethal exposure to waterborne Cu, Cd or Zn in rainbow trout 2: tissue specific metal accumulation. *Aquat. Toxicol.* **2000,** 50, (3), 245-256.

48. McGeer, J. C.; Szebedinszky, C.; McDonald, D. G.; Wood, C. M., Effects of chronic sublethal exposure to waterborne Cu, Cd or Zn in rainbow trout. 1: Iono-regulatory disturbance and metabolic costs. *Aquat. Toxicol.* **2000,** 50, (3), 231-243.

49. http://www.ccme.ca/publications/

50. Lynn, S.; Yew, F. H.; Chen, K. S.; Jan, K. Y., Reactive oxygen species are involved in nickel inhibition of DNA repair. *Environ. Mol. Mutagen.* **1997,** 29, (2), 208-216.

51. Yilmaz, B.; Ssempebwa, J.; Mackerer, C. R.; Arcaro, K. F.; Carpenter, D. O., Effects of polycyclic aromatic hydrocarbon-containing oil mixtures on generation of reactive oxygen species and cell viability in MCF-7 breast cancer cells. *J. Toxicol. Environ. Health, Part A* **2007,** 70, (13), 1108-1115.

52. Weisman, D.; Alkio, M.; Colon-Carmona, A., Transcriptional responses to polycyclic aromatic hydrocarbon-induced stress in Arabidopsis thaliana reveal the involvement of hormone and defense signaling pathways. *BMC Plant Biol* 10, 59.

53. Gaetke, L. M.; Chow, C. K., Copper toxicity, oxidative stress, and antioxidant nutrients. *Toxicology* **2003,** 189, (1-2), 147-163.

54. Pourahmad, J.; O'Brien, P. J.; Jokar, F.; Daraei, B., Carcinogenic metal induced sites of reactive oxygen species formation in hepatocytes. *Toxicol. in Vitro* **2003,** 17, (5/6), 803-810.

55. Zhang, Z.; Huang, C.; Li, J.; Leonard, S. S.; Lanciotti, R.; Butterworth, L.; Shi, X., Vanadate-induced cell growth regulation and the role of reactive oxygen species. *Arch. Biochem. Biophys.* **2001,** 392, (2), 311-320.

56. Valko, M.; Morris, H.; Cronin, M. T. D., Metals, toxicity and oxidative stress. *Curr. Med. Chem.* **2005,** 12, (10), 1161-1208.

57. Korashy, H. M.; El-Kadi, A. O. S., The role of aryl hydrocarbon receptor and the reactive oxygen species in the modulation of glutathione transferase by heavy metals in murine hepatoma cell lines. *Chem.-Biol. Interact.* **2006,** 162, (3), 237-248.

58. The Universal Protein Resource (UniProt). *Nucleic Acids Res.* **2008,** 36, (Database Iss), D190-D195.

59. Bayne, C. J.; Gerwick, L., The acute phase response and innate immunity of fish. *Dev. Comp. Immunol.* **2001,** 25, (8-9), 725-743.

60. Dautry-Varsat, A.; Ciechanover, A.; Lodish, H. F., pH and the recycling of transferrin during receptor-mediated endocytosis. *Proc. Natl. Acad. Sci. U. S. A.* **1983,** 80, (8), 2258-62.

61. Takenaka, K. V.; Sakai, N.; Murase, S.; Kuroda, T.; Okumura, A.; Sawada, M., Elevated transferrin concentration in cerebral spinal fluid after subarachnoid hemorrhage. *Neurol. Res.* **2000,** 22, (8), 797-801.

62. Sinars, C. R.; Cheung-Flynn, J.; Rimerman, R. A.; Scammell, J. G.; Smith, D. F.; Clardy, J., Structure of the large FK506-binding protein FKBP51, an Hsp90-binding protein and a component of steroid receptor complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** 100, (3), 868-873.

63. Dennis, G.; Sherman, B. T.; Hosack, D. A.; Yang, J.; Gao, W.; Lane, H. C.; Lempicki, R. A., DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol.* **2003,** 4, (9), 11.

64. Zohlnhofer, D.; Klein, C. A.; Richter, T.; Brandl, R.; Murr, A.; Nuhrenberg, T.; Schomig, A.; Baeuerle, P. A.; Neumann, F.-J., Gene expression profiling of human stent-induced neointima by cDNA array analysis of microscopic specimens retrieved by helix cutter atherectomy. Detection of FK506-binding protein 12 upregulation. *Circulation* **2001,** 103, (10), 1396-1402.

65. Weber, C.; Guigon, G.; Bouchier, C.; Frangeul, L.; Moreira, S.; Sismeiro, O.; Gouyette, C.; Mirelman, D.; Coppee, J. Y.; Guillen, N., Stress by heat shock induces massive down regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in Entamoeba histolytica. *Eukaryotic Cell* **2006,** 5, (5), 871-875.

66. Gomez, A. M.; Rueda, A.; Sainte-Marie, Y.; Pereira, L.; Zissimopoulos, S.; Zhu, X.; Schaub, R.; Perrier, E.; Perrier, R.; Latouche, C.; Richard, S.; Picot, M.-C.; Jaisser, F.; Lai, F. A.; Valdivia, H. H.; Benitah, J.-P., Mineralocorticoid modulation of cardiac ryanodine receptor activity is associated with downregulation of FK506-binding proteins. *Circulation* **2009,** 119, (16), 2179-2187.

67. Fruman, D. A.; Burakoff, S. J.; Bierer, B. E., Immunophilins in protein folding and immunosuppression. *Faseb J.* **1994,** 8, (6), 391-400.

68. van den Heuvel, M. R.; Power, M.; Richards, J.; MacKinnon, M.; Dixon, D. G., Disease and gill lesions in yellow perch (Perca flavescens) exposed to oil sands mining-associated waters. *Ecotoxicol. Environ. Saf.* **2000,** 46, (3), 334-341.

69. Siwik, P. L.; Van Meer, T.; MacKinnon, M. D.; Paszkowski, C. A., Growth of fathead minnows in oilsand-processed wastewater in laboratory and field. *Environ. Toxicol. Chem.* **2000,** 19, (7), 1837-1845.

70. Kavanagh, R. J.; Frank, R. A.; Oakes, K. D.; Servos, M. R.; Young, R. F.; Fedorak, P. M.; MacKinnon, M. D.; Solomon, K. R.; Dixon, D. G.; Van Der

Kraak, G., Fathead minnow (Pimephales promelas) reproduction is impaired in aged oil sands process-affected waters. *Aquat. Toxicol.* **2011,** 101, (1), 214-220.

71. Rifai, N.; Gillette, M. A.; Carr, S. A., Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat. Biotechnol.* **2006,** 24, (8), 971-983.

Chapter 5

A Method to Standardize Inter-Sample and Intra-Sample Urine to Account for Urinary Volume Variation

5.1 Introduction

Biofluids are routinely used for diagnoses and urine is the least invasive biofluid compared to blood, plasma, serum, and cerebral spinal fluid. The analysis of markers in urine, such as glucose, ketones,[1] leukocytes, nitrates,[2] and proteins,[3-5] have often been used in disease diagnosis and monitoring. Urine is, however, subject to biological variation not only among subjects, but within subjects. Urine volume is subject to kidney filtration and liquid consumption, thereby affecting urinary metabolite concentration.[6] A larger volume of urine will contain a lower concentration of metabolites than a less dilute volume of urine, thereby complicating the standard determination of metabolite concentrations.

Creatinine is a metabolite of creatine phosphate, a molecule stored by muscles and used for the production of ATP.[7] Creatine and creatine phosphate are converted to creatinine (Figure 5.1), which diffuses into the blood and is filtered by the kidneys into urine. In healthy individuals, the rate of creatinine formation and excretion is relatively constant. Creatinine levels in urine and blood are routinely used to assess renal function and monitor cardiovascular disease and diabetes.[3, 8] Although the levels of creatinine in blood are constant, the levels in urine samples vary throughout the day due to urine dilution, therefore a 24 hour urine sample is collected to determine the total concentration of

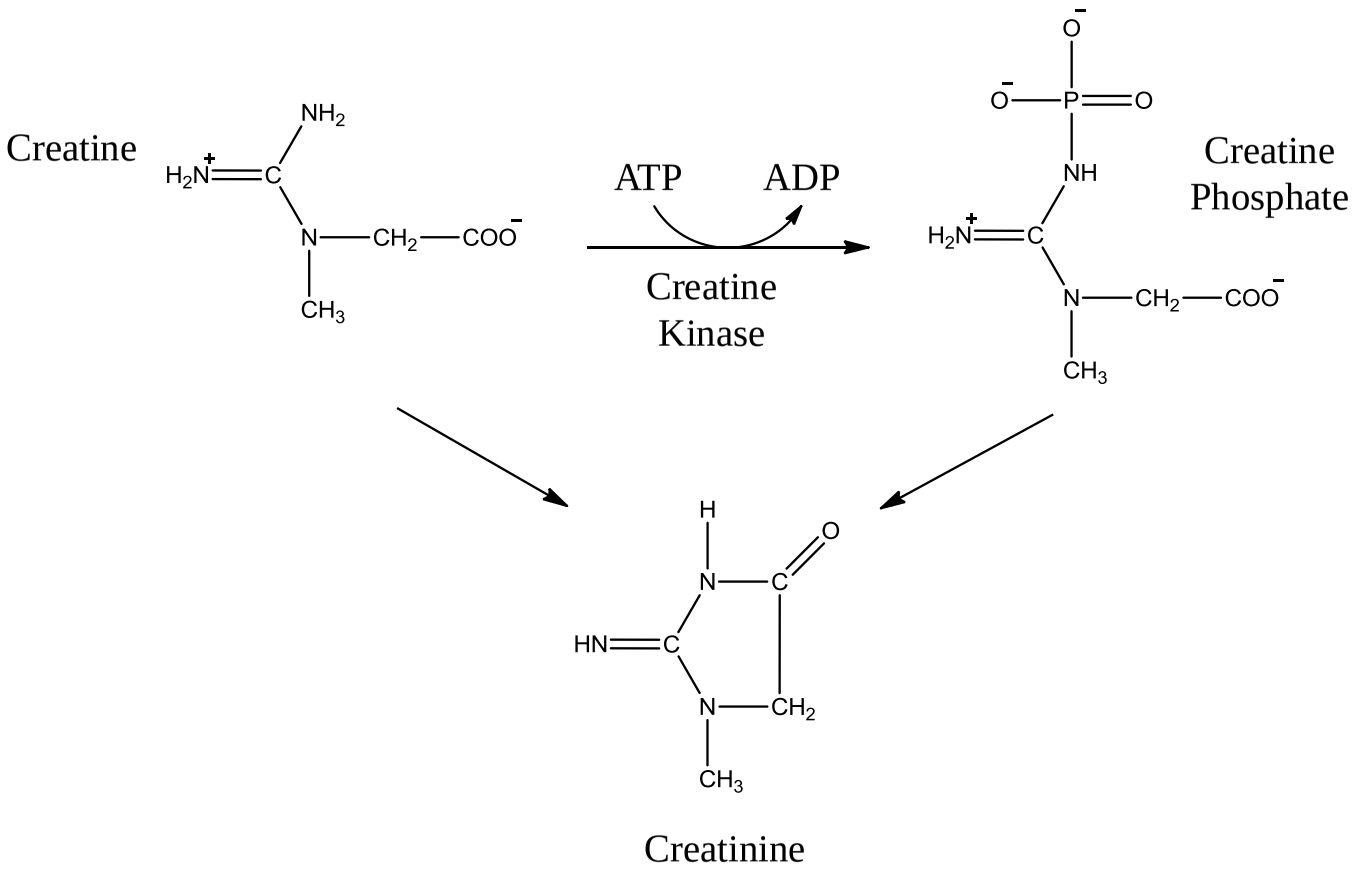

Figure 5.1. Formation of creatinine from creatine and creatine phosphate.

excreted creatinine. If renal function is compromised, creatinine levels in the blood are elevated and levels in urine are reduced.

The inconvenience of sample collection and storage makes 24 hour urine collection cumbersome for the subject. Due to the variability of urine concentration, creatinine is routinely used to normalize levels of other molecules, such as albumin and proteins, in random urine samples.[4, 5] For example, the American Diabetes Association uses the microalbumin:creatinine ratio in random urine samples to screen for microalbuminuria, an early indicator of renal dysfunction.[3]

In nuclear magnetic resonance (NMR) analysis of urine, there are two widely accepted strategies to account for urine volume variation and the subsequent effect on metabolite concentration measurement.[9, 10] In one method, the NMR spectral signal is normalized against a housekeeping metabolite, typically creatinine peak signal. In the other method, the NMR spectra signal is set to total unit intensity and the metabolites of interest are expressed as a percentage of this intensity in a method referred to as normalization to a constant sum. Studies have been performed comparing the two normalization strategies,[11, 12] but no consensus has been reached as to which of the two normalization strategies is superior. Although both methods have been shown to delineate between different cohorts in multivariate analysis (e.g. principal component analysis (PCA)), there are instances where PCA scores and groupings vary significantly for a few metabolites.[11, 12]

While liquid chromatography mass spectrometry (LC/MS) offers the advantage of greater sensitivity to NMR, there is no generally accepted strategy, as there is in NMR, on how to account for urinary volume variation. Some studies have compared metabolite signals to total intensity of spectra;[13] others to spiked creatinine.[14] In some MS metabolite studies, equivolume samples are pooled to create a control to which individual metabolite profiles are compared.[15, 16] A drawback to generating a pooled standard is that a low concentration unique metabolite, potentially one of significance, may be lost in the analysis, such as when using stable isotope labelling to detect peaks pairs. Pooling multiple samples may no longer detect this unique metabolite. With increasing focus on quantitative metabolomics, it is important to focus research efforts on a normalization strategy to relate urinary volumes.

5.2 Experimental

5.2.1 Chemicals and Reagents

The QuantiChrom™ Creatinine Assay Kit was purchased from BioAssay Systems (Hayward, CA, USA). LC-MS grade water and acetonitrile were purchased from Fisher Scientific Canada (Edmonton, AB, Canada). LC-MS grade trifluoroacetic acid (TFA) and formic acid (FA) were purchased from Sigma-Aldrich (Oakville, ON, Canada).

5.2.2 Urine Sample Collection and Processing

This study was conducted in accordance with the Arts, Science & Law Research Ethics Board policy at the University of Alberta and Standard Operating

Procedure 001 (see Appendix A). All participants were provided with information on the study and a consent form (see Appendix B). Eligible participants were instructed to refrain from eating and drinking, with the exception of water, for at least eight hours prior to first void and to continue refraining from eating and drinking, with the exception of water, prior to the second void. Upon collecting the second void urine sample, the participants were instructed to return the sample within two hours for storage at 4 ^{o}C. Further information on participation criteria and instructions can be found in Appendix C.

Urine samples were processed in a Containment Level 2 laboratory prior to long-term storage in a Containment Level 1 laboratory. Samples were centrifuged at 4000 rpm for 10 minutes and the resulting supernatant filtered through a 0.2 µm filter. The filtered urine was then re-filtered through another 0.2 µm filter. The double-filtered urine was aliquoted and stored at -80 ^{o}C. Prior to analysis, urine was thawed and kept on ice.

5.2.3 Creatinine Assay

The creatinine assay was performed following the manufacturer's protocol and is based on an improved Jaffé method that utilizes picrate to form an orange-red-coloured complex with creatinine. The intensity of the colour, measured at 510 nm, is directly proportional to creatinine concentration in the sample. The manufacturer claims that the optimized formulation substantially reduces interference by substances in the untreated sample. The kit contained 50 mg/dL creatinine standard, Reagent A, and Reagent B. Briefly, creatinine standard at 50 mg/dL and urine diluted two-fold in water were transferred to a 96-well plate in

triplicate. Working solutions were prepared by mixing equal volumes of Reagents A and B, which was quickly added to the standard and sample wells, followed by brief mixing. The optical density was measured at 1 min and 5 min at 510 nm. Creatinine concentration of the samples was calculated using the following equation: $\frac{OD_{sample\ 5} - OD_{sample\ 1}}{OD_{standard\ 5} - OD_{standard\ 1}} \times n \times 50\ (mg/dL)$, where $OD_{sample\ 5}$ and $OD_{sample\ 1}$ are the optical density values of the sample at 5 minutes and 1 minute, respectively, $OD_{standard\ 5}$ and $OD_{standard\ 1}$ are the optical density values of the standard at 5 minutes and 1 minute, respectively, and *n* is the dilution factor.

5.2.4 LC/UV

Samples were diluted 10 times in Mobile Phase A (0.1% TFA in water). An Agilent 1100 series LC equipped with a UV detector was used for the analysis. Fifty microliters of diluted urine was injected onto a Polaris C18A column (Palo Alto, CA, USA, 3 µm, 50 mm x 2.0 mm). Mobile Phase A consisted of 0.1% TFA in water and Mobile Phase B consisted of 0.1% TFA in acetonitrile. The elution gradient was 2.5% B for 5.00 min, 2.5-85% B from 5.00 to 5.01 min, 85% B from 5.01 to 10.00 min, 85-2.5% B from 10.00 to 10.01 min, and held at 2.5% B from 10.01 to 15.00 min at a flow rate of 0.200 mL/min. UV chromatograms were obtained at 210, 214, 254, and 280 nm. Samples were run in quadruplicate.

5.2.5 LC/MS

Samples were diluted 10 times in Mobile Phase A (0.1% FA in water). LC/MS analyses were performed on an Agilent 1100 series HPLC system coupled to an esquire3000plus ion trap (Bruker Daltonics, Bremen, Germany) equipped with an electrospray source operating in positive mode. Chromatographic separation was achieved on an ACE® C18 column (3 µm, 100 Å, 50 mm x 2.1 mm). Mobile Phase A consisted of 0.1% FA in water and Mobile Phase B consisted of 0.1% FA in acetonitrile. Fifty microliters of diluted urine was separated at 4.0-4.0% B from 0.00-1.00 min, 4.0-50.0% B from 1.00-80.00 min, 50.0-100.0% B from 80.00-85.00 min, held at 100.0% B for 10.00 min, 100.0 to 4.0% B from 95.0 to 95.1 min, and held at 4.0% B until 105 min at an eluent flow rate of 100 µL/min.

5.3 Results and Discussion

Concentrations of endogenous metabolites are affected by a variety of factors, including age, gender, and diet.[17] While a study can control age and gender, control over diet is nontrivial, even when the subjects are rats.[6] Aside from inter-subject biological variability, metabolomic analysis of urine also faces the challenges of intra-subject biological variability, volume variation, sample type (e.g. first void, second void, 24 hour, random), and preanalytical handling and storage.[16, 18] Urine volume is subject to water consumption, physiological factors, such as kidney function, and pathophysiological factors. Consequently,

the concentrations of endogenous metabolites in urine will also vary, necessitating the need to normalize samples.

In LC/MS analysis of urine, treatment to account for volume variation varies from pooling samples,[15, 16] to addition of internal standards,[6] to normalizing to total ion intensity,[13] to normalizing to creatinine intensity.[14] With no standard methodology, study comparisons are limited to samples analyzed within a study.

It was our intent to determine if a relationship existed between concentration of creatinine, a common marker, and different measured parameters of urine (LC/UV absorbance signal, LC/MS TIC intensity, and LC/MS creatinine EIC intensity). We realize that a single reference compound is not representative of the wide range of metabolites present in urine, but a point of commencement was necessary. To the best of our knowledge, a study comparing these specific parameters has not been conducted.

5.3.1 Creatinine Assay Analysis

Creatinine levels in urine samples were measured using a commercially available kit. The commercial creatinine assay kit is based on the Jaffé reaction where creatinine reacts with alkaline picrate to form an orange-red creatinine-picrate complex. Further modifications have developed the Jaffé reaction into a kinetic assay that substantially reduces interferences by other biological components.[19]

Second-void samples were obtained from ten male and ten female volunteers on three consecutive days and the measured creatinine values are listed in Table 5.1. Each sample was assayed in triplicate and the associated correlation

Table 5.1. Creatinine values measured by the creatinine assay kit. Each sample was measured in triplicate with correlation of variation (CV, %) in brackets.

Sample Code	Sex	Creatinine (mg/dL)			AVG (mg/dL)	CV (%)
		Day 1	Day 2	Day 3		
001	F	137 (14)	256 (10)	95 (45)	162	52
004	F	184 (5)	20 (7)	28 (12)	77	120
005	F	45 (7)	44 (10)	15 (9)	35	49
008	F	128 (9)	187 (1)	122 (8)	146	25
009	F	61 (9)	16 (25)	122 (14)	66	80
010	M	55 (4)	27 (19)	18 (12)	33	57
011	M	62 (21)	92 (6)	134 (13)	96	38
015	M	190 (2)	231 (14)	201 (11)	207	10
016	M	139 (2)	56 (5)	79 (8)	91	47
018	F	33 (2)	46 (5)	26 (4)	35	28
019	F	27 (41)	15 (13)	27 (20)	23	31
021	M	250 (7)	31 (11)	111 (11)	131	85
023	M	25(41)	63 (5)	25 (7)	38	58
024	F	97 (2)	8 (9)	14 (10)	40	124
025	M	63 (14)	47 (6)	74 (13)	62	22
026	M	68 (3)	18(17)	109 (23)	65	71
027	F	45 (32)	62 (5)	40 (10)	49	24
029	M	136 (5)	93 (12)	49 (5)	93	47
030	M	159(5)	109 (9)	183 (9)	150	25
032	F	106 (4)	108 (14)	117 (8)	110	5

of variation (CV) indicated in the table. The range of creatinine values between subjects is not surprising since urinary creatinine depends on age, gender, and lean muscle mass.[20] We limited our study to subjects between the age of 20 and 30 to reduce age-related biological variation. The measured creatinine values among one person over the three days varied from as low as 5% to as high as 124%, with an average of 50% . An assay reanalyzing the creatinine amounts for those samples with CVs greater than 50% over the three days produced similar results. The variation in urinary creatinine concentration within subjects was not surprising; changes in urine volume for the second void over three days could account for the variability. In order to gauge a more consistent urinary creatinine concentration, a 24 hour urine sample, where the subject would collect the entire bladder void contents over a 24 hour period, would give a clearer indication of urinary creatinine levels. However, we chose to analyze a second void sample for the convenience of sample collection for the participants and to minimize variation in preprocessing storage conditions. We chose to relate all other findings (LC/UV and LC/MS) to the assayed creatinine values, since the creatinine assay is a standardized assay.

5.3.2 LC/UV Analysis

Metabolites have diverse properties, therefore we chose to monitor UV absorbance at a range of wavelengths (210, 214, 254, and 280 nm) to account for some of the different chromophores in urine. For example, carboxyls absorb at 210 nm, amidos at 214 nm, conjugated systems (e.g. double bonds) at 254 nm, and carbonyls and nitros at 280 nm.[21] The four wavelengths monitored is by no

means a comprehensive UV analysis, not withstanding that there are metabolites that do not contain chromophores.

Diluted urine samples were run in quadruplicate and their UV absorbance measured at 210, 214, 254, and 280 nm. The chromatogram peaks were integrated and the areas under the curve (AUC) measured between 1 and 12 minutes.

Figure 5.2 shows the correlation of AUC values as a function of the creatinine assay values for each of the three days. The AUC for the samples were normalized against the AUC for blanks, also run in quadruplicate. Across the three days, the square of the correlation coefficient (R^2, referred to as correlation for the remainder of this chapter) between UV absorbance and creatinine concentration was the weakest at 210 and 254 nm ($R^2 < 0.65$, with the exception of the R^2 value at 210 nm on Day 2), and consistently greater than 0.81 at 214 nm.

It was suspected that a separation on the basis of gender might improve correlation values because creatinine values differ based on a variety of factors, including gender and lean muscle mass. Separating the plots by gender (Figures 5.3 and 5.4) showed consistent improved correlation for females at all wavelengths (improvement ranging from 3 to 124% with an average improvement of 27%), but was generally weaker for males when compared to the correlation values obtained when treating both genders as one group (Table 5.2). Even with separation by gender, metabolic profiles can vary vastly due to diet. In fact, even when fed a control diet, Sprague Dawley rats exhibited different metabolic

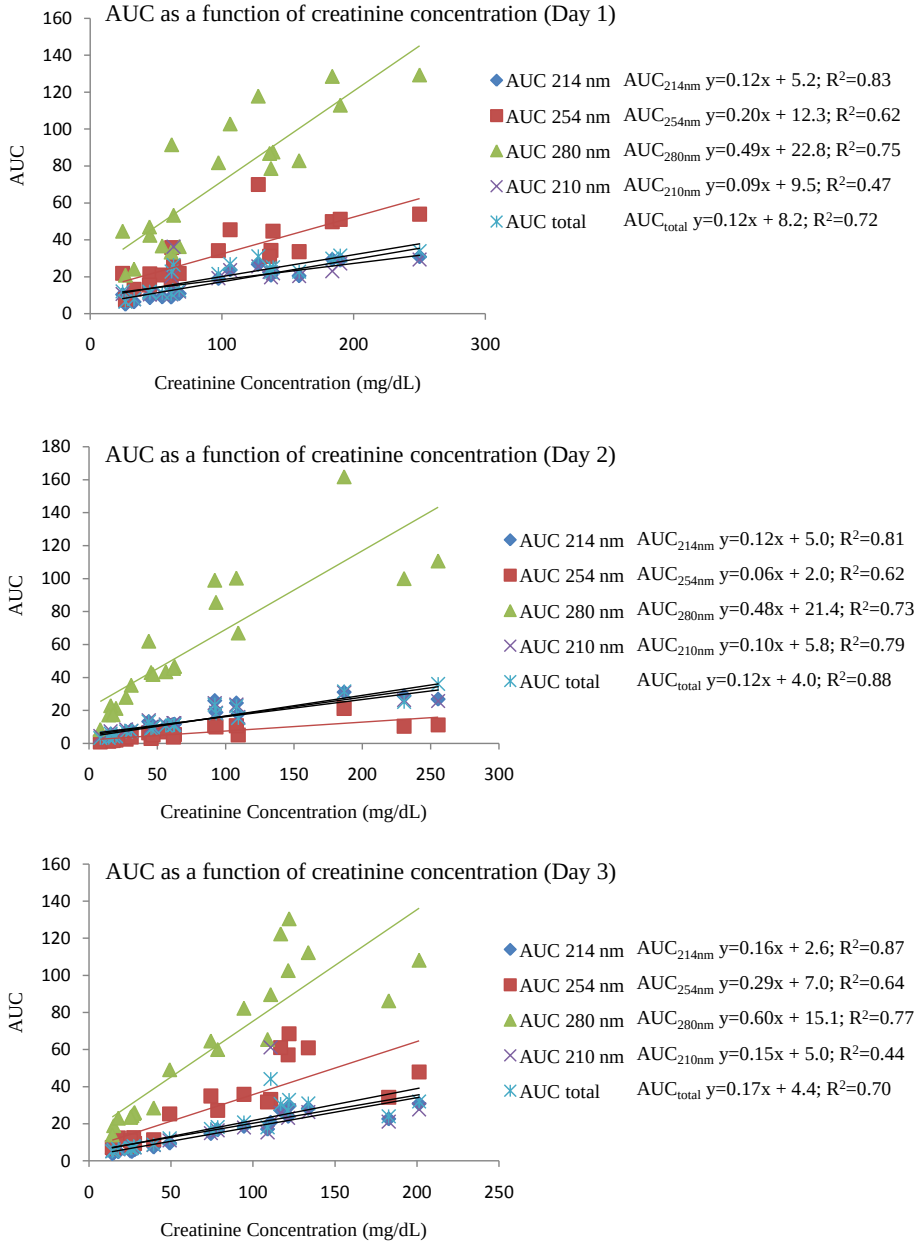

Figure 5.2. Plots of AUC as a function of creatinine concentration for all 20 samples over the course of three days. AUC was normalized to blanks.

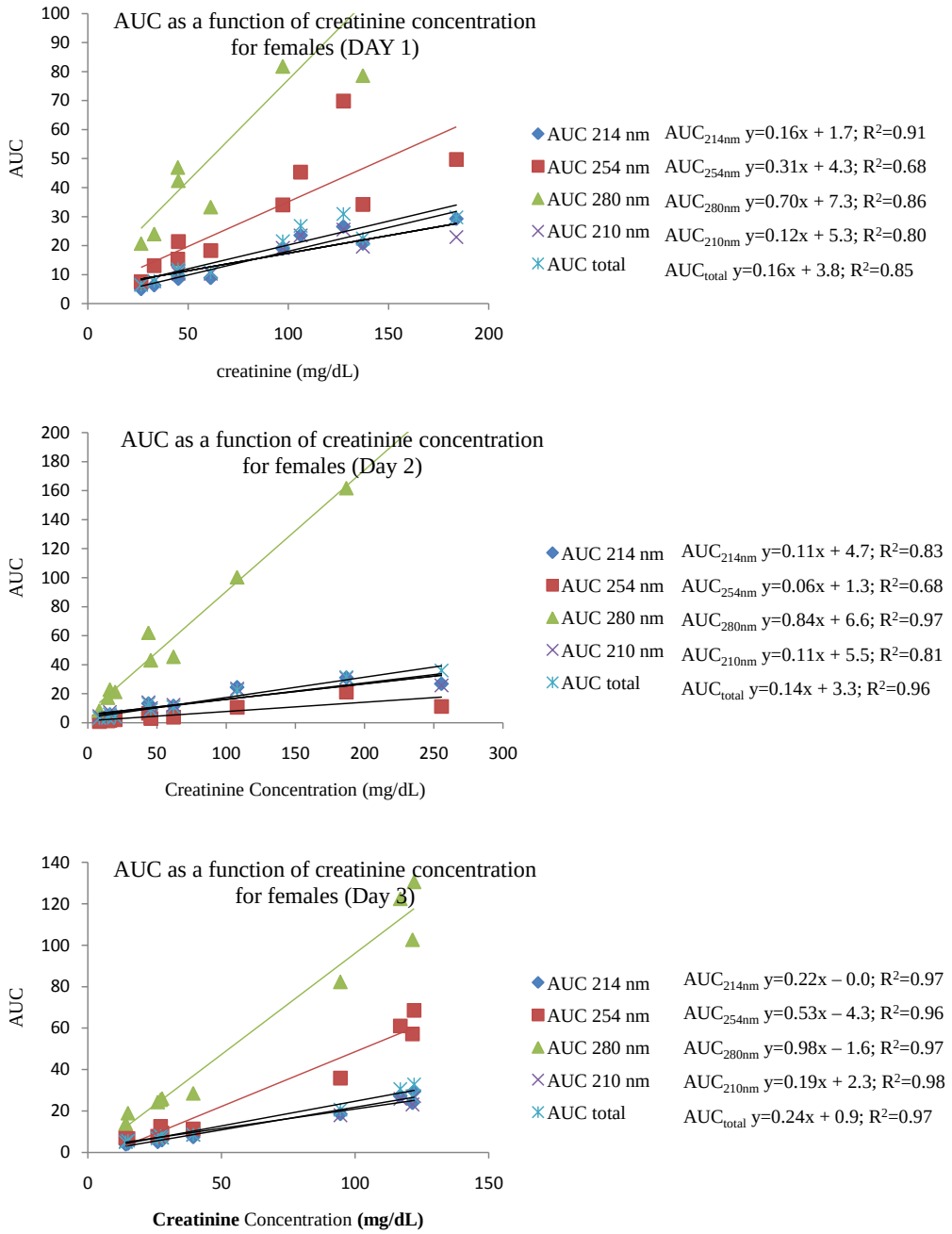

Figure 5.3. Plots of AUC as a function of creatinine concentration for 10 samples from females over the course of three days. AUC was normalized to blanks.

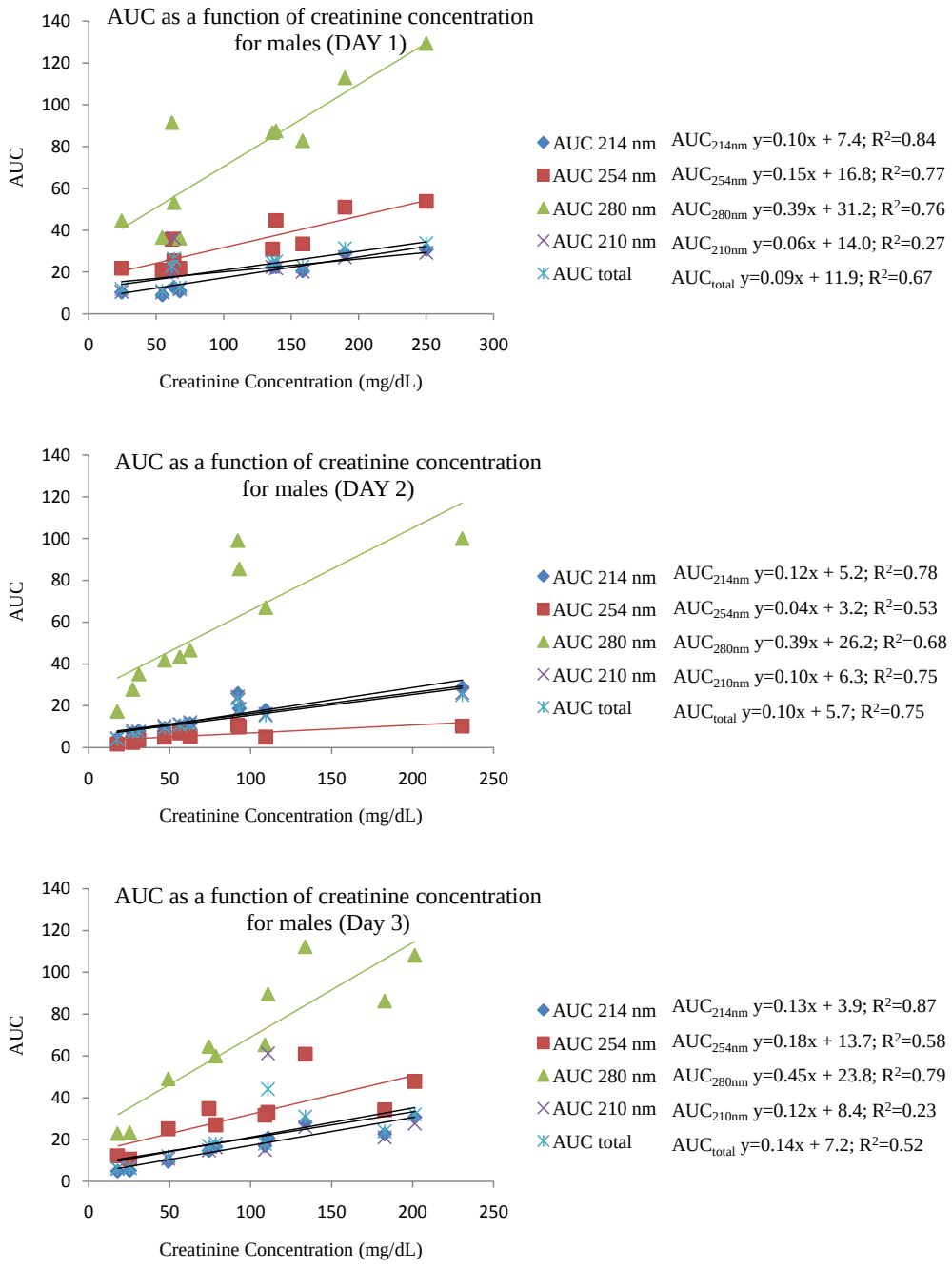

Figure 5.4. Plots of AUC as a function of creatinine concentration for 10 samples from males over the course of three days. AUC was normalized to blanks.

Table 5.2. Comparison of R^2 values upon gender differentiation.

Parameter	Day	λ (nm)	M & F[1] R^2	M		F	
				R^2	% Change[2]	R^2	% Change[2]
AUC	1	214	0.83	0.84	1.1	0.91	9.3
		254	0.62	0.77	25	0.68	10
		280	0.75	0.76	1.6	0.86	14
		210	0.47	0.27	-42	0.80	70
		Sum	0.72	0.67	-7.5	0.85	17
	2	214	0.81	0.78	-3.5	0.83	2.5
		254	0.62	0.53	-14	0.68	9.9
		280	0.73	0.68	-6.9	0.78	6.6
		210	0.79	0.75	-4.7	0.81	2.7
		Sum	0.88	0.75	-15	0.96	8.9
	3	214	0.87	0.87	0.23	0.97	13
		254	0.64	0.58	-9.0	0.96	49
		280	0.77	0.79	3.5	0.97	26
		210	0.44	0.23	-47	0.98	120
		Sum	0.70	0.52	-26	0.97	399
EIC	1	n/a	0.88	0.85	-3.2	0.94	6.6
	2	n/a	0.88	0.78	-12	0.95	7.9
	3	n/a	0.83	0.86	3.4	0.77	-7.4
TIC	1	n/a	0.74	0.79	6.9	0.73	-1.9
	2	n/a	0.71	0.65	-8.8	0.78	10
	3	n/a	0.82	0.89	9.2	0.89	8.6

[1] M & F refer to the plots containing all 20 data points; M= male; F = female
[2] % Change = (measured – "expected")/"expected", where measured refers to the R^2 value from M or F plots and "expected" refers to the R^2 value from the M & F plots.

different metabolic profiles,[6] and the diets of people are subject to less control.

5.3.3 LC/MS Analysis

The MS data were analyzed using Bruker Daltonics' DataAnalysis (Version 3.2) software. An extracted ion chromatogram (EIC) of m/z 114.1, corresponding to ionized creatinine, with a window of ± 0.5 Da, was obtained and the peak intensity plotted against the assayed creatinine concentration. The plots for Days 1, 2, and 3 are shown in Figure 5.5 and there is generally an increased correlation between the assayed creatinine concentrations and the creatinine EIC signals compared to the assayed creatinine concentrations and the absorbance signals, summarized in Table 5.2. We further investigated the correlation between assayed creatinine concentrations and the total ion intensity of chromatographed peaks from the ion spectra (Figure 5.6), which generally exhibits a weaker correlation when compared to the EIC. This was not surprising as the electrospray ionization efficiency will depend on the analytes competing for ionization and the complexity of the urine sample insinuates increased competition for ionization. Different urine samples will contain not only different concentrations of the same metabolites, but also different metabolites.

5.3.4 Overall Analysis

Our data shows that the highest correlation is between the creatinine assay results and the creatinine EIC signal, with slightly greater correlation among the female samples than the male samples.

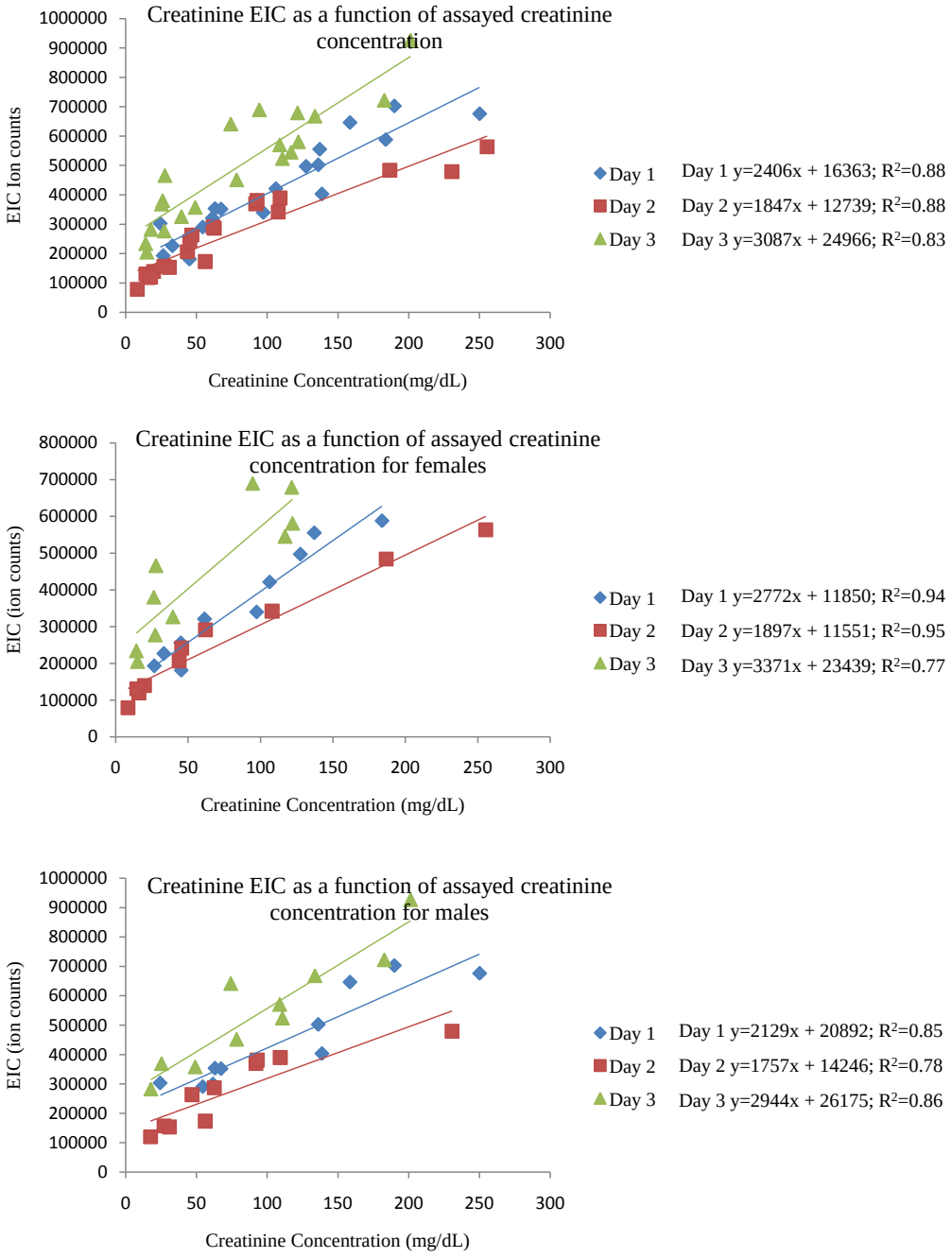

Figure 5.5. Plots of creatinine EIC signal as a function of creatinine concentration.

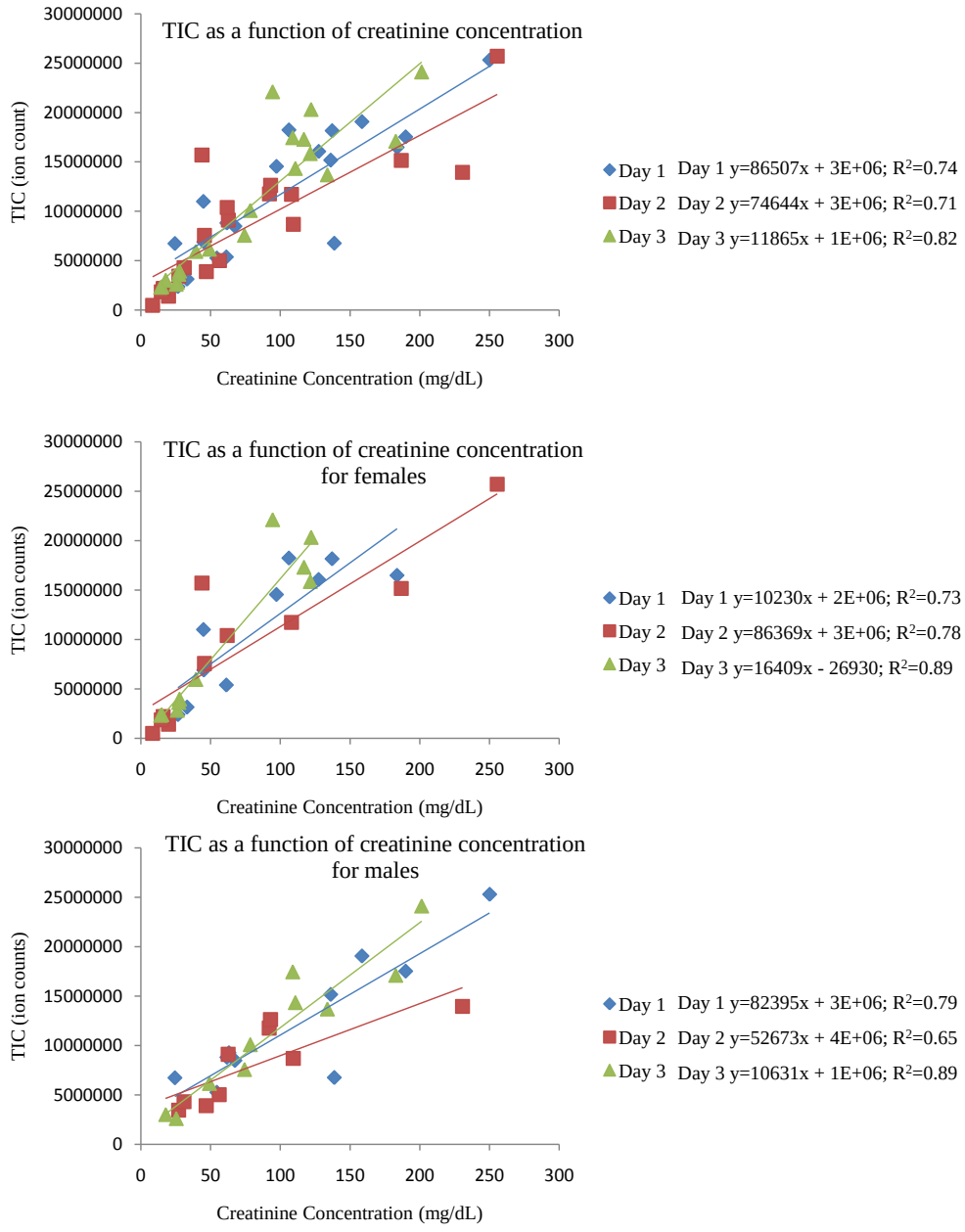

Figure 5.6. Plots of TIC signal as a function of creatinine concentration.

The correlation between assayed creatinine values and normalized AUC values were weaker in comparison; when comparing all 20 samples, the greatest correlation among the wavelengths was at $AUC_{214\ nm}$ (R^2 = 0.83, 0.81, and 0.87 for Days 1, 2, and 3, respectively). A study by Kemperman *et al.* evaluated relating urine injection volumes to the AUC of chromatogram traces obtained at 214 nm and to creatinine EIC for the purpose of urinary protein analysis.[16] For the purpose of their analysis, they found that the lowest CVs were obtained using AUC at 214 nm, not unexpected since the peptide backbone absorbs at that wavelength. Their study detected spiked proteins in urine, therefore the biological variation would be minimal due to lower natural protein content in healthy individuals.

5.4 Conclusions

Although the study was designed to be diligent in the time between sample collection and preprocessing storage, the time was controlled by the participant. Urine at room temperature, and even at 4 ^{o}C is known to undergo changes due to metabolite stability[22-24] and, consequently, these sample changes will affect the concentration and types of metabolites detected. As mentioned previously, a 24 hour urine study would be best for studying urinary creatinine, but at the expense of studying other urinary metabolites and hindering our goal of standardizing urine samples. Although this study did not specifically examine any metabolites other than creatinine, we do want to extend the use to any urine sample in general and metabolite.

From our study of ten male and ten female participants, each of whom contributed second void urine samples on three consecutive days, we attempted to determine a relationship between urinary creatinine values as measured by a commercial creatinine assay kit and LC/UV chromatograms and LC/MS ion chromatograms to account for urinary volume variation. We noticed that separating our analysis of the data on the basis of gender gave a better indication of relationships between creatinine assay results and other measurements, giving credence to the factors affecting creatinine content, such as lean body mass. From our results, it appears that the relationship between creatinine assay values and EIC creatinine gives the best correlation as can be seen in Table 5.2. Studies involving metabolite labelling strategies[15, 25] to compare other metabolites would be beneficial to furthering this study.

5.5 Literature Cited

1. Ueda, H.; Nagumo, F., Determination of urine ketone bodies. *Med. Technol. (Tokyo)* 1983, 11, (9), 873-877.

2. Soler Rodriguez, F.; Miguez Santiyan, M. P.; Pedrera Zamorano, J. D., Evaluation of reagent strips for the rapid diagnosis of nitrite poisoning. *J. Anal. Toxicol.* 1992, 16, (1), 63-65.

3. Gross Jorge, L.; de Azevedo Mirela, J.; Silveiro Sandra, P.; Canani Luis, H.; Caramori Maria, L.; Zelmanovitz, T., Diabetic nephropathy: diagnosis, prevention, and treatment. *Diabetes Care* 2005, 28, (1), 164-176.

4. Miller, R. C.; Brindle, E.; Holman, D. J.; Shofer, J.; Klein, N. A.; Soules, M. R.; O'Connor, K. A., Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. *Clin. Chem. (Washington, DC, U. S.)* 2004, 50, (5), 924-932.

5. Molitch Mark, E.; DeFronzo Ralph, A.; Franz Marion, J.; Keane William, F.; Mogensen Carl, E.; Parving, H.-H.; Steffes Michael, W., Nephropathy in diabetes. *Diabetes Care* 2004, 27 Suppl 1, S79-83.

6. Warrack, B. M.; Hnatyshyn, S.; Ott, K.-H.; Reily, M. D.; Sanders, M.; Zhang, H.; Drexler, D. M., Normalization strategies for metabonomic analysis of urine samples. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* 2009, 877, (5-6), 547-552.

7. Wallimann, T.; Wyss, M.; Brdiczka, D.; Nicolay, K.; Eppenberger, H. M., Intracellular compartmentation, structure and function of creatine kinase isoenzymes in tissues with high and fluctuating energy demands: the 'phosphocreatine circuit' for cellular energy homeostasis. *Biochem. J.* 1992, 281, (1), 21-40.

8. Anavekar, N. S.; McMurray, J. J. V.; Velazquez, E. J.; Solomon, S. D.; Kober, L.; Rouleau, J.-L.; White, H. D.; Nordlander, R.; Maggioni, A.; Dickstein, K.; Zelenkofske, S.; Jeffrey, D. O.; Leimberger, D.; Califf, R. M.; Pfeffer, M. A., Relation between renal dysfunction and cardiovascular outcomes after myocardial infarction. *N. Engl. J. Med.* 2004, 351, (13), 1285-1295.

9. Craig, A.; Cloarec, O.; Holmes, E.; Nicholson, J. K.; Lindon, J. C., Scaling and Normalization Effects in NMR Spectroscopic Metabonomic Data Sets. *Anal. Chem.* 2006, 78, (7), 2262-2267.

10. Gartland, K. P. R.; Bonner, F. W.; Nicholson, J. K., Investigations into the biochemical effects of region-specific nephrotoxins. *Mol. Pharmacol.* 1989, 35, (2), 242-250.

11. Akira, K.; Masu, S.; Imachi, M.; Mitome, H.; Hashimoto, M.; Hashimoto, T., 1H NMR-based metabonomic analysis of urine from young spontaneously hypertensive rats. *J. Pharm. Biomed. Anal.* 2008, 46, (3), 550-556.

12. Schnackenberg, L. K.; Sun, J.; Espandiari, P.; Holland, R. D.; Hanig, J.; Beger, R. D., Metabonomics evaluations of age-related changes in the urinary compositions of male Sprague Dawley rats and effects of data normalization methods on statistical and quantitative analysis. *BMC Bioinf.* 2007, 8, (Suppl. 7), doi:10.1186/1471-2105-8-S7-S3.

13. Jonsson, P.; Bruce, S. J.; Moritz, T.; Trygg, J.; Sjoestroem, M.; Plumb, R.; Granger, J.; Maibaum, E.; Nicholson, J. K.; Holmes, E.; Antti, H., Extraction, interpretation and validation of information for comparing samples in metabolic LC/MS data sets. *Analyst (Cambridge, U. K.)* 2005, 130, (5), 701-707.

14. Neale, J. R.; Dean, B. J., Liquid chromatography-tandem mass spectrometric quantification of the dehydration product of tetranor PGE-M, the major urinary metabolite of prostaglandin E2 in human urine. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* 2008, 871, (1), 72-77.

15. Guo, K.; Li, L., Differential 12C-/13C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal. Chem. (Washington, DC, U. S.)* 2009, 81, (10), 3919-3932.

16. Kemperman, R. F. J.; Horvatovich, P. L.; Hoekman, B.; Reijmers, T. H.; Muskiet, F. A. J.; Bischoff, R., Comparative urine analysis by liquid chromatography-mass spectrometry and multivariate statistics: method development, evaluation, and application to proteinuria. *J. Proteome Res.* 2007, 6, (1), 194-206.

17. Lenz, E. M.; Bright, J.; Wilson, I. D.; Hughes, A.; Morrisson, J.; Lindberg, H.; Lockton, A., Metabonomics, dietary influences and cultural differences: a 1H NMR-based study of urine samples obtained from healthy British and Swedish subjects. *J. Pharm. Biomed. Anal.* 2004, 36, (4), 841-849.

18. Miller, W. G.; Bruns, D. E.; Hortin, G. L.; Sandberg, S.; Aakre, K. M.; McQueen, M. J.; Itoh, Y.; Lieske, J. C.; Seccombe, D. W.; Jones, G.; Bunk, D. M.; Curhan, G. C.; Narva, A. S., Current issues in measurement and reporting of urinary albumin excretion. *Clin. Chem. (Washington, DC, U. S.)* 2009, 55, (1), 24-38.

19. Cook, J. G., Factors influencing the assay of creatinine. *Ann Clin Biochem* 1975, 12, (6), 219-232.

20. Barr, D. B.; Wilder, L. C.; Caudill, S. P.; Gonzalez, A. J.; Needham, L. L.; Pirkle, J. L., Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. *Environ. Health Perspect.* 2005, 113, (2), 192-200.

21. Skoog, D. A., *Principles of Instrumental Analysis. 3rd Ed*. 1985; p. 410.

22. Collins, A. C.; Sethi, M.; MacDonald, F. A.; Brown, D.; Viberti, G. C., Storage temperature and differing methods of sample preparation in the measurement of urinary albumin. *Diabetologia* 1993, 36, (10), 993-997.

23. Guzman, N. A.; Hernandez, L.; Advis, J. P., Effect of low temperature storage (on collected urine specimens) in the identification and quantitation of urinary metabolites by capillary electrophoresis. *J. Liq. Chromatogr.* 1989, 12, (13), 2563-2574.

24. Lomeo, A.; Bolner, A., Stability of several biochemical markers of bone metabolism. *Clin. Chem. (Washington, D. C.)* 2000, 46, (8, Pt. 1), 1200-1202.

25. Guo, K.; Li, L., High-performance isotope labeling for profiling carboxylic acid-containing metabolites in biofluids by mass spectrometry. *Anal. Chem. (Washington, DC, U. S.)* 82, (21), 8789-8793.

Chapter 6

Conclusions and Future Work

The majority of this thesis focussed on using mass spectrometry in aquatic toxicoproteomics studies through the qualitative and quantitative characterization of the zebrafish gill proteome. Toxicoproteomics endeavours to identify critical proteins and pathways in biological systems that are affected by and respond to adverse chemical and environmental exposures using proteomic technologies.

In Chapter 2, we used a combination of techniques (sequential solubilization and 2D LC-ESI MS/MS with precursor ion exclusion) to maximize proteome coverage of zebrafish gill and identified 5716 unique proteins. Although annotative information was not available for 30% of the proteins, the identified proteome did exhibit excellent coverage of biochemical pathways relevant to the numerous physiological functions of the gill. This comprehensive proteome profile served as a preface for using zebrafish gill proteome for future studies (Chapters 3 and 4) in aquatic toxicology.

In Chapter 3, the effect of NAs exposure (3 mg/L) on the proteins of zebrafish gill was investigated using the 2MEGA stable isotope labelling approach. Gill digests of the NH_4HCO_3 soluble fraction of exposed zebrafish was divided and the lysine amino residues guanidinylated, followed by dimethylation of the N-termini of peptides with either CH_2O or $^{13}CD_2O$. Controls were similarly treated. Labelled digests were mixed such that a CH_2O labelled control and a $^{13}CD_2O$ labelled sample constituted the forward case and a $^{13}CD_2O$ labelled control and a CH_2O labelled sample constituted the reverse case. The two cases

were analyzed using 2D LC-ESI MS/MS with precursor ion exclusion. Subsequent data processing revealed 15 differentially-expressed proteins in at least two of the three Biological Replicates. Among the differentially-expressed proteins were two classes of downregulated proteins, vitellogenin and MHC class I. These results suggest that NAs disrupt both the reproductive and immune systems, evident by the downregulation of vitellogenin proteins, and downregulation of vitellogenin and MHC class I proteins, respectively. The downregulation of vitellogenin and MHC class I proteins suggests that NAs have an immunosuppressive response. Further studies involving same-sex fish are required to give credence to the hypothesis that NAs affect reproduction.

Investigating NAs response on zebrafish led to a curiosity of the effects of OSPW on zebrafish since the interest in NAs stems from their presence in OSPW. OSPW has other constituents, such as PAHS, heavy metals, residual bitumen and dissolved organics and inorganics, which may contribute to its toxicity. In Chapter 4, we investigated the effects of exposure to OSPW (diluted to contain 3 mg/L NA) on the proteome of zebrafish gill. The same methodology used in Chapter 3 was followed and we found 99 differentially-expressed proteins in a minimum of two of the three Biological Replicates upon OSPW exposure. The upregulation of proteins in the adherens junction, focal adhesion, regulation of actin cytoskeleton and tight junction pathways were suggestive of structural damage to the gill tissue. As a consequence, an increase in metabolism and ATP production to provide energy and resources to repair damage would be expected and was observed with the upregulation of numerous proteins involved in various

metabolic pathways and in the glycolysis/gluconeogenesis pathway. Evidence of immunoregulation and oxidative stress response was also observed.

Finally, in Chapter 5, an investigation into standardizing urine analysis by relating the amount of urine to creatinine concentration was conducted. Although a useful biofluid for metabolite analysis, metabolite concentration in urine varies as a function of kidney filtration and liquid consumption. We proposed using urinary creatinine concentration as a standard normalization factor for LC/MS when analyzing urine samples. We used a commercial creatinine assay kit to measure creatinine in each of the samples and related these values to the responses from LC-UV and LC/MS. The best correlation was with the LC/MS EIC creatinine signal, and this signal could be used as a normalization factor when measuring other metabolites. Further studies involving measuring total 24 hour urine volume and creatinine would be beneficial to understanding the relation between creatinine and urinary dilution.

While establishing the proteome profile of the gill was a necessary step in the pathway to utilizing zebrafish gill as a marker in aquatic toxicity, a greater understanding of protein-protein interactions and protein-pathway interactions is needed to provide a larger picture. Our research efforts were limited by the lack of annotative information, but we were still able to glean important pathway information regarding physiological function. This information was a useful basis for our quantitative toxicoproteomics endeavours. Bioinformatics, much like the proteomics field, is evolving, therefore the ease of proteomics analysis can improve with bioinformatics.

Initially, we chose to quantify only the NH_4HCO_3 protein soluble portion of the toxicant exposures. We can gain more information on toxicant effects if we look at the whole proteome (e.g. SDS and urea soluble protein fractions) than a single portion of the proteome.

While the NAs and OSPW exposure studies have identified differentially-expressed proteins, some of which could be classified into KEGG pathways, there is still a need to validate these putative biomarkers of exposure beyond biological replicates. We have plans to use Western Blots for further validation.

Our efforts focussed on monitoring proteomic changes at relatively low toxicant values for a short period of time (seven days). Acute and chronic exposure studies at different toxicant exposures would be interesting to monitor how the organism adapts in the short- and long-term, and to garner an understanding of mechanisms of action of toxicity. Furthermore, tailings ponds vary in composition and concentration of their components and may exhibit different effects on aquatic organisms. A study comparing differential-expression of zebrafish gill after exposure to different sites could further aid in understanding the mechanisms of action of toxicity of OSPW.

To gain a dynamic picture of the biology involved with toxicant exposure, a systems biology approach, where transcriptomics, proteomics, and metabolomics strategies are used, could provide complementary information. Our laboratory has developed stable isotope labelling strategies for the quantitative analysis of metabolites (Figures 6.1 to 6.3) and applying these strategies to gill

metabolite extracts could show interesting correlations between proteins and their associated metabolite products.

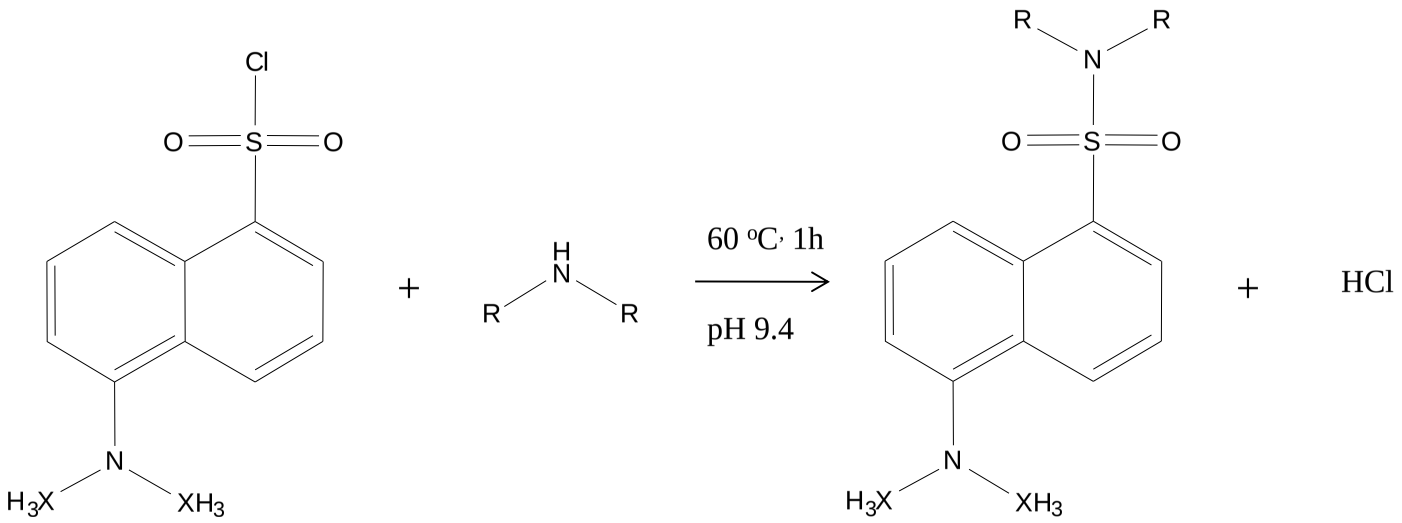

Figure 6.1 Reaction scheme for isotopic labelling of amine- and phenol-containing metabolites using dansyl chloride. (X = ^{12}C for light isotope label and ^{13}C for heavy isotope label).

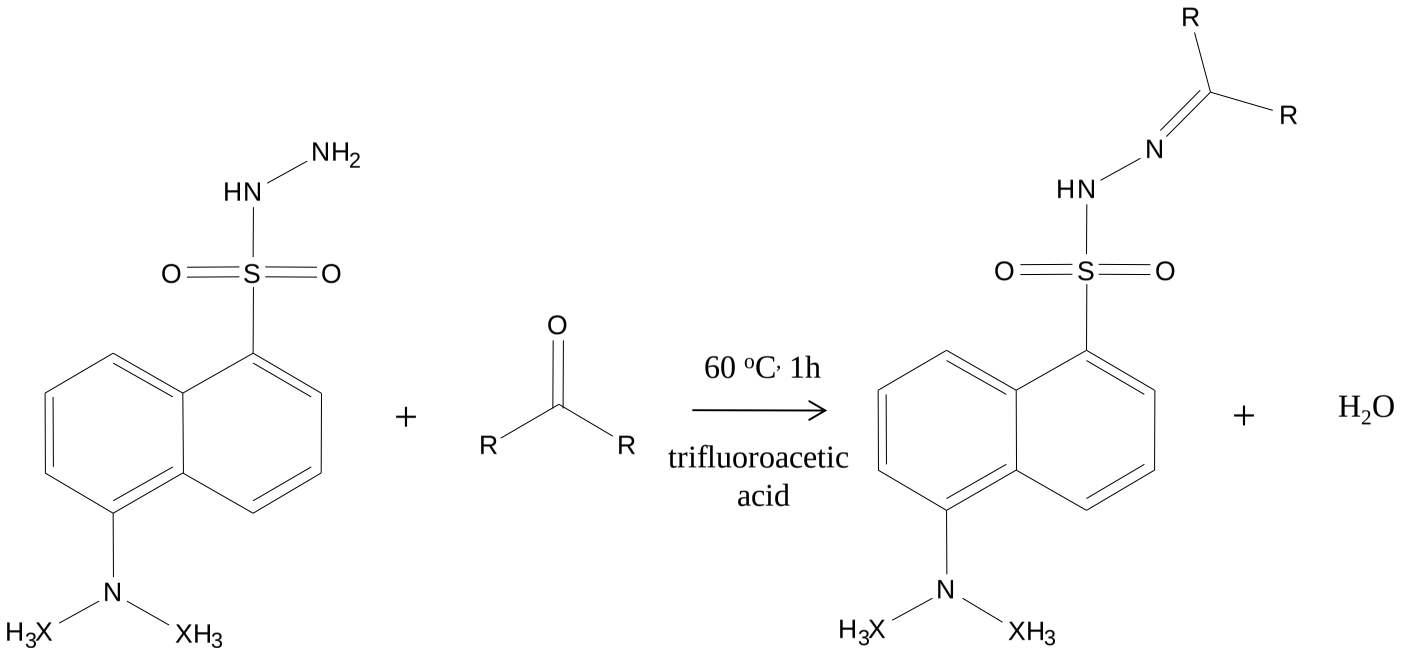

Figure 6.2 Reaction scheme for isotopic labelling of ketone- and aldehyde-containing metabolites using dansyl hydrazine. (X = ^{12}C for light isotope label and ^{13}C for heavy isotope label).

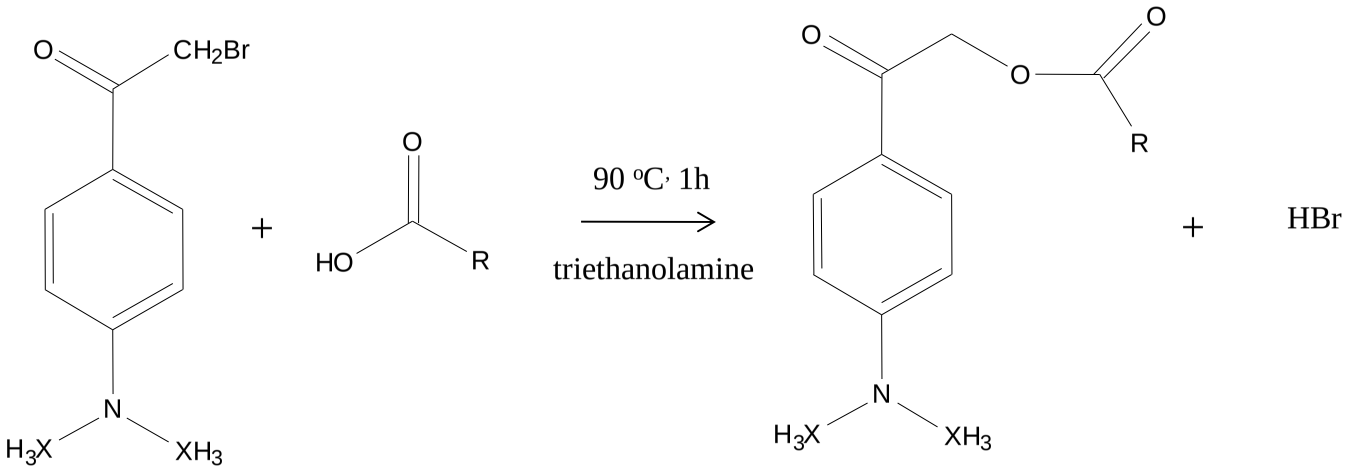

Figure 6.3 Reaction scheme for isotopic labelling of carboxylic acid--containing metabolites using p=dimethylaminephenacyl bromide. (X = ^{12}C for light isotope label and ^{13}C for heavy isotope label).

Appendix A

Obtaining and Storage of Human Urine Samples Standard Operating Procedure
SOP#001 Version #1

1. Scope

This document applies to the collection and storage of human urine samples for the Li research group.

2. Objective

Illustrate the procedures used for obtaining and storage of urine samples for the Li research group.

3. Materials

3.1. Resealable bag

3.2. Towelettes

3.3. BD Falcon 50 mL Polypropylene Conical Tube, 30 x 115 mm

3.4. Safety Equipment

- 3.4.1. Lab Gloves
- 3.4.2. Lab Coat
- 3.4.3. Lab Safety Glasses

3.5. LC-MS grade water

3.6. LC-MS grade methanol

3.7. Micropipettes and pipette tips

3.8. Eppendorf microcentrifuge tubes (1.5 mL)

3.9. Freeze and solvent resistant labels and markers

3.10. Kimwipes

3.11. 70% Ethanol solution

3.12. Nylon syringe filters, 0.2 micron

3.13. Syringes

3.14. Ice

3.15. Ice bucket

3.16. Dry Ice

3.17. Dry Ice bucket

4. Equipment

4.1. -80 ^{o}C freezer

4.2. Centrifuge (Containment Level 2)

5. Safety

5.1. All lab personnel should be familiar with the contents of the University of Alberta Biosafety Manual.

5.2. Appropriate safety apparel, including safety glasses, lab coat, and lab gloves should be worn.

5.3. Urine samples must be properly processed in the Containment Level 2 laboratory (W4-10) prior to storage in a Containment Level 1 laboratory.

6. Sample Handling

6.1. Keep urine samples on ice at all times that it is not stored at -80 ^{o}C.

7. Obtaining Urine Samples

7.1. Urine samples are to be obtained only from participants who have reviewed and signed the appropriate Research Information and Participants' Consent Form.

7.2. Once the Research Information and Participants' Consent Form has been read and signed, provide the participant with a resealable bag, towelettes, and BD Falcon 50 mL Polypropylene Conical Tubes, 30 x 115 mm

7.3. Participants are requested to refrain from eating or drinking for eight hours except water prior to collecting urine. Specific instructions for collecting midstream urine outlined in SOP 002 and SOP 003.

7.4. Once the participant has collected the urine specimen, they are instructed to return the specimen in a resealable bag within two hours of collection to a designate, as outlined in the appropriate Research Information and Participants' Consent Form.

8. Processing Urine Samples

8.1. Upon receipt of the specimen, move the sample to the Containment Level 2 laboratory located in Chemistry W4-10. Record the sample ID, gender, age, BMI, sample collection time and date, and operator (your name).

8.2. Wipe the centrifuge with 70% ethanol before use. Make sure that the centrifuge is balanced. Centrifuge the sample at 4000 rpm for 10 minutes. Wipe the centrifuge with 70% ethanol after use.

8.3. Turn on the biosafety hood and allow to run for at least 15 minutes before beginning work. Disinfect the hood before and after each use by cleaning surface, including front grill, with 70% ethanol. Spray items with ethanol before placing them in the hood. Spray gloves in ethanol and rub hands together. Lab coats with loose sleeves should be cuffed by tying them up. This prevents you from dragging contaminated air into the hood.

8.4. Transfer the supernatant of the urine using a pipette to a 0.2 micron filter-capped syringe. Filter the urine into a clean BD Falcon tube. Refilter the urine using a new 0.2 micron filter. The double-filtered urine is now treated for removal into a Containment Level 1 laboratory.

8.5. Place any used items (e.g. gloves, pipette tips, BD Falcon tubes), with the exception of sharps, into the biohazardous waste containers in W4-10. Sharps should be disposed of in the sharps container.

8.6. Make certain to disinfect the centrifuge and biosafety hood with 70% ethanol.

9. Storage of Urine Samples

9.1. Once the urine samples have been filtered twice through a 0.2 micron filter, they can be moved to a Containment Level 1 laboratory.

9.2. Filtered urine is to be stored in 1 mL aliquots at -80 ^{o}C in Chemistry W3-03.

10. Responsibility.

This SOP was written by: Andrea G. De Souza _____

This SOP was formatted by: Andrea G. De Souza _____

This SOP was reviewed by: Liang Li _____

11. Date: 19 January 2010

"I have read and understand this SOP. I agree to fully adhere to its requirements."

Date (YY/MM/DD)	Last Name	First Name	Signature

Appendix B

Research Information and Participant's Consent Form

Name of Study: Method Development for Inter-sample and Intra-sample Urine Comparison

Purpose. You are invited to participate in a research study (*Method Development for Inter-sample Urine Comparison*) being conducted by Professor Liang Li of the Department of Chemistry, University of Alberta. This study intends to develop an analytical technique for minimizing the effects of variable dilution in urine samples using liquid chromatography and mass spectrometry. The analytical equipment to be used for the analyses is located in the Department of Chemistry (room numbers W3-03, W3-06, W4-10, and EB-38). The primary objective of this study is to standardize inter-sample and intra-sample urine to account for urinary volume variations and, hence, urinary metabolite concentration. The samples may be used for other studies involving method development and validation for urine analysis.

Your participation. Your participation involves collecting a second void urine sample in the container provided following specific instructions provided in an attached document. The only identifiers on your sample will be an arbitrary code, your gender and BMI, and the sample collection time and date. Once the sample has been collected, please proceed directly to Chemistry W3-28, where you will be given access to place your sample in Chemistry W4-10. Further instructions are provided in an attached document.

Your rights. Your decision to participate in this study is entirely voluntary and you may decide at any time to withdraw. Your participation will remain confidential and anonymous. Our data file will NOT contain any personal identifiers (i.e., names or student ID numbers). The results of this study may be presented at scholarly conferences, published in professional journals, or presented in class lectures. The data will be securely stored by Professor Liang Li for a minimum of five years.

Benefits and risks. This research contributes to the development of new analytical methods for characterizing metabolites. Better characterization of metabolites can potentially contribute to the advancement of our understanding of how metabolites affect human health. There are no foreseeable risks to this study, but if any risks should arise, the researcher will inform the participants immediately. If you should experience any adverse effects, please contact Professor Liang Li or Andrea De Souza immediately.

Contact information. If you have any questions or comments on the study, or if you wish a clarification of rights as a research participant, you can contact Professor Liang Li or Andrea De Souza or the Human Research Ethics Committee at the numbers and addresses below.

Professor Liang Li
Department of Chemistry
University of Alberta
Chemistry Centre W3-39
Edmonton, AB T6G 2G2
(780) 492-3250
liang.li@ualberta.ca

Andrea De Souza
Department of Chemistry
University of Alberta
Chemistry Centre W3-03
Edmonton, AB T6G 2G2
(780) 492-6103
desouza@ualberta.ca

Dr. Tom Johnson, Chair
Arts, Science, and Law Research Ethics Board
Faculty of Arts
University of Alberta
Edmonton, AB T6G 2E9
(780) 492-2834
tom.johnson@ualberta.ca

Participant's Signature _____ Date _____

Researcher's Signature _____ Date _____

Appendix C

Instructions for Participants

Name of Study: Method Development for Inter-sample and Intra-sample Urine Comparison

Criteria

1. Participants should be between the ages of 20 and 30 years old.
2. Participants should not consume prescription medications (ie. Antibiotics) with the exception of birth control.
3. Participants should not consume over the counter pharmaceuticals (ie. Cough/cold medicine, pain medicine) with the exception of natural supplements (e.g. vitamins)
4. Participants should be non-smokers.
5. Participants should not be menstruating on the days of sample collection. Blood proteins interfere with analysis.
6. Participants are instructed to provide second void specimens for three consecutive days.
7. Participants are instructed to refrain from eating or drinking (except water) for at least 8 hours prior to first void and to continue refraining from eating or drinking (except water) prior to the second void.

Materials Provided

For each day of urine collection, you will be provided with the following:

1. 1 brown paper bag
2. 1 50 mL collection tube
3. 1 re-sealable bag
4. 2 towelettes
5. 1 label with a pre-assigned arbitrary code

Please do not use your own towelettes or collection containers

Second Void Specimen Collection Instructions

1. When you first arise from a night's sleep, void your bladder. DO NOT collect this specimen. Continue abstaining from food and drink (except water).
2. When you are ready to void your bladder again (second void), wash and dry your hands.
3. Clean your urogenital area with the towelette provided. Note: Women are to clean from front to back only.
4. Begin urinating into the toilet. Stop the flow of urine, and begin urinating into the container until at least 20 milliliters of urine (or as much urine as possible, if less than 20 milliliters) has been collected. Stop the flow of urine, remove the container, and continue urinating into the toilet.

5. Screw the lid securely on the container, taking care not to touch the rim or the inside of the container with your fingers. Wipe the outside of the container with the second towelette.
6. Place the container in the re-sealable plastic bag, then place this bag in the brown bag.
7. Write the following information on the provided label and place the label in the brown paper bag:
   Gender (M or F)
   BMI (calculation instructions are on this sheet)
   Date of collection
   Time of collection
8. Immediately go to Chemistry W4-28 where Chiao-Li (Dibs) Tseng, Yiman (Delia) Wu, or Jiamin (Tammy) Zheng will take you to Chemistry W4-10 and unlock the door for you. You will walk into the room and place your brown paper bag in the refrigerator.

Calculating BMI (Body Mass Index)

Any of the following three equations may be used to calculate BMI:

1. $BMI = \frac{mass\ (kg)}{(height\ (m))^2}$
2. $BMI = \frac{mass\ (lb) \times 703}{(height\ (in))^2}$
3. $BMI = \frac{mass\ (lb) \times 4.88}{(height\ (ft))^2}$

Contact information. If you have any questions, you can contact Chiao-Li (Dibs) Li, Yiman (Delia) Wu, or Jiamin (Tammy) Zheng.

Chiao-Li (Dibs) Tseng	Yiman (Delia) Wu	Jiamin (Tammy) Zhang
Chemistry W3-28	Chemistry W3-28	Chemistry W3-28
chiaoli@ualberta.ca	yiman1@ualberta.ca	jiamin3@ualberta.ca