Talamantes, Tatjana D., Discovery-driven Proteomics to Advance the Evaluation of the Role of Lipids and Lipid Peroxidation-related Carbonyl Stress in Biological Model Systems. Doctor of Philosophy (Biochemistry and Molecular Biology), May, 2013, 160 pp., 6 tables, 28 illustrations, bibliography, 183 titles.

Research has come a long way with the advancement of sequenced genomes in regards to a variety of organisms. In addition, the technology used to explore biological information encoded by these genome sequences has also been enhanced. However, the knowledge of the role that simple components play in biological functions and cellular processes of the organism, from a global perspective, remains limited. With the advent of proteomics, exploring pathways and networks influenced by cellular components such as lipids, has become possible in the context of systems biology. This dissertation has been divided into two parts. Both will focus on a discovery-driven approach to elucidate the diverse biological functions of lipids as well as stress and its consequences in two biological model systems using mass spectrometry-based proteomics. The first part investigates the broader role of phospholipids (specifically phosphatidylglycerol), essential components of biological membranes in prokaryotes and eukaryotes, in a genetically modified strain of cyanobacteria. Ultimately, understanding the impact of lipid-regulation in this simple organism is expected to reveal previously unknown lipid-signaling mechanisms, which could be further investigated in higher-order organisms. The second part is dedicated to the investigation of lipid-signaling mechanisms generated by electrophilic products of oxidative stress-induced lipid peroxidation, and its consequences in disturbing homeostasis in zebrafish embryos. Exploring free radical-mediated lipid peroxidation

involved in cellular responses promises a comprehensive understanding of the role and/or contribution that electrophilic products of lipid peroxidation play in modulating oxidative stress-related signaling pathways potentially involved in various pathophysiological dysfunctions. Bioinformatics software successfully constructed networks from the proteins identified as being regulated by the induced carbonyl-stress including intra- and intercellular processes involved in eIF2 signaling, glycolysis, and remodeling of epithelial adherens junctions.

DISCOVERY-DRIVEN PROTEOMICS TO ADVANCE THE EVALUATION OF THE ROLE OF LIPIDS AND LIPID PEROXIDATION-RELATED CARBONYL STRESS IN BIOLOGICAL MODEL SYSTEMS

Tatjana D. Talamantes

APPROVED:

_____
Major Professor

_____
Committee Member

_____
Committee Member

_____
Committee Member

_____
University Member

_____
Chair, Department of Molecular Biology and Immunology

_____
Dean, Graduate School of Biomedical Sciences

DISCOVERY-DRIVEN PROTEOMICS TO ADVANCE THE EVALUATION OF THE ROLE OF LIPIDS AND LIPID PEROXIDATION-RELATED CARBONYL STRESS IN BIOLOGICAL MODEL SYSTEMS

DISSERTATION

Presented to the Graduate council of the

Graduate School of Biomedical Sciences

University of North Texas Health Science Center at Fort Worth

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR of PHILOSPHY

By

Tatjana D. Talamantes

Fort Worth, Texas

May 2013

ACKNOWLEDGEMENTS

The following dissertation has been quite the journey. I could not have reached this point without the encouragement and support of many people.

First and foremost, I wish to thank my advisor, Dr. Laszlo Prokai for his patience, knowledge, and guidance. Without the opportunity to work in his lab and for introducing me to the exciting field of mass spectrometry and proteomics my tenure at the University of North Texas Health Science Center would not have been possible.

My gratitude is further expressed towards my advisory committee members, Dr. Katalin Prokai-Tatrai, Dr. Andras Lacko, Dr. Yogesh Awasthi, and Dr. James Caffrey for their assistance and suggestions over the years.

To Dr. Zolán Gombos, Dr. Bettina Ughy, and Dr. Mihály Kis, from the Biological Research Center in Hungary, thank you for your collaborations, knowledge, and guidance. This dissertation would have been incomplete without your assistance.

In addition, I would like to thank my fellow lab members–those who have moved on and those who are current–for their friendliness, feedback, and support. To all of the secretaries in the Molecular Biology and Immunology Department, thank you for your motivation, encouragement, and for lending an ear when I needed it.

It is impossible to individually thank all of my greatest friends–those from college, those from graduate school, and those outside of school–that have left a permanent impression on me

through this journey. Your continuous entertainment, laughter, inspiration, and companionship have kept me shining bright like a diamond.

To my grandparents and my brother, thank you for always believing in me. Last, but not least, a very special thank you goes to my Mom and Pops. Without your love, financial support, praise, and trust, I would not be where I am today. I am forever indebted to you for always pushing me to be the best I can be.

TABLE OF CONTENTS

Page

V. SUMMARY AND DISCUSSION

ABBREVIATIONS

PG, Phosphatidylglycerol

PUFA, Polyunsaturated fatty acid

HNE, 4-Hydroxy-2-nonenal

Cys, Cysteine

His, Histidine

Lys, Lysine

Arg, arginine

ROS, Reactive oxygen species

RNS, Reactive nitrogen species

PCC, Pasteur Culture Collection

m/z, Mas-to-charge

RP-LC/MS, Reversed-phase liquid chromatography mass spectrometry

HPLC, High performance liquid chromatography

LC, Liquid chromatography

LC-MS, Liquid chromatography mass spectrometry

LC-MS/MS, Liquid chromatography tandem mass spectrometry

RP, Reversed phase

DTT, Dithioreitol

iTRAQ, Isobaric tags for relative and absolute quantitation

ICAT, Isotope-coded affinity tags

SILAC, Stable isotope labels with amino acids in cell culture

MS/MS, Tandem mass spectrometry

AUC, Area under the curve

GO, Gene Ontology

PSI, Photosystem I

PSII, Photosystem II

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I. INTRODUCTION

CHAPTER I

INTRODUCTION

Membrane Components: Lipids and Fatty Acid Constituents

As one of the major constituents of cells, lipids are composed of a wide variety of compounds that are characteristically insoluble in water (1). In addition, most lipids have structural properties that give them important biological features. Composed of ionic or polar derivatives of hydrocarbons, these compounds are classified as "amphiphiles". In addition, the ionic or polar groups are denoted as hydrophilic ("water-loving") whereas the nonpolar groups are hydrophobic. These properties contribute to different roles of the lipid including transport, energy storage, and membrane structure. There are three major types of amphipathic structural lipids found in membranes: glycerophospholipids, sphingolipids, and cholesterol (Figure 1). Phosphorylated derivatives of sphingolipids as well as glycerophospholipids are commonly referred to as phospholipids (henceforth referred to as this). Additionally, phospholipids are responsible for organizing the bilayer structure of the membrane, while cholesterol regulates fluidity of the membrane (2).

The different membrane lipids result from various combinations of fatty acid chains and polar head groups. Fatty acids are basic components of many lipid classes, such as free

carboxylic acids, phospholipids, diacylglycerols, and triacylglycerols (3, 4). In essence, phospholipids form the backbone of all cell membranes and contribute to membrane functionality and fluidity (4, 5). These types of lipids contain two long-chain fatty acids attached via an ester linkage to positions 1 and 2 of a glycerol backbone that contains a polar head group attached at position 3 (Figure 2). In addition, the fatty acids that attach to the glycerol backbone vary in length as well as structure. Importantly, fatty acids include either a single bond between the carbons (saturated), forming straight hydrocarbon chains, or they have double bonds between the carbons (unsaturated) causing the chain to bend (2, 4, 5). Fatty acid composition is species- and tissue-specific and about half of the moieties in plants and animals are unsaturated (4).

Lipids: Phospholipids and Physiological Functions

The role of phospholipids is a dynamic phenomenon. Initially, these components were believed to only play a structural role (6). However, as the interest in lipid research has steadily increased, these molecules have been exposed as being active in many aspects of membrane function such as signaling. Moreover, phospholipid membranes keep cells separate from the surrounding environment and organelles separate from the cytosol. The ability for these membranes to form barriers and boundaries allows each main section of the cell to maintain a unique biochemical nature. Along with structural responsibilities, lipids have been shown to regulate key physiological functions including exocytosis, chemotaxis, and cytokinesis, which have been reviewed recently (5).

Lipids, as well as their fatty acid components, are relatively diverse. Not only do fatty acids contribute to the structural and functional duties of cell membranes, but they also participate as precursors for lipid mediators, which affect signal transduction pathways and gene

transcription (7). Moreover, they can be involved in the regulation of biological processes such as inflammatory responses, neuronal signal transmission, and carbohydrate metabolism (8). However, there is a continued need for uncovering additional details about the roles lipids are involved with, in order to have a clearer perspective of their impact in regards to unexplored lipid signaling mechanisms.

Phosphoglycerides make up a large percentage of lipids in biological membranes (3). These constituents belong to the group of glycerol-based phospholipids that contain a polar head group other than phosphate, constructing them as an essential component of biomolecules. Phosphatidylglycerol (PG) (Figure 3) is one of the main phospholipids in most bacterial membranes and can be located in plant and animal membranes (9). Having a phosphate group, this molecule is negatively charged at neutral pH and is classified as an anionic phospholipid (10). For non-photosynthetic prokaryotes, PG is an important but dispensable element of cellular processes (11) whereas its physiological role in non-photosynthetic eukaryotes has not been fully established (10). However, current research has indicated that PG plays a role in mitochondria, as well as several different roles required for respiratory electron transport and oxidative phosphorylation in membranes of mitochondria (10, 12). Furthermore, it is apparent that the complete understanding of lipids and their influence on biological systems remains incomplete. Therefore, questioning the consequences of the function of biological factors along with their removal/inactivation in a biological system may provide a better understanding of membrane components (6).

Phospholipids: Polyunsaturation and Peroxidation

The cause and effect of the imbalance of redox-based changes continues to be a challenge for researchers. Nevertheless, it is widely acknowledged that oxidative stress is important in

regards to the biology and chemistry of several pathologies (13-16). Oxidation of biomembrane lipids (lipid peroxidation), represents a form of redox imbalance in cells and tissues (17). Polyunsaturated fatty acids (PUFAs) contain two or more double bonds and provide readily removable hydrogen molecules (18). This makes them more vulnerable to peroxidation (4, 13, 19).

Lipid peroxidation includes three critical steps: 1) initiation, 2) propagation, and 3) termination (Figure 4A) (18). Resulting lipid radicals rapidly react with molecular oxygen to form peroxyl radicals. Peroxidation then initiates a series of complex, autocatalytic propagation reactions that generate a variety of lipid electrophiles. Such constituents include the α,β-unsaturated hydroxyalkenal, 4-Hydroxy-2-nonenal (HNE) (Figure 4B), which is a peroxidation byproduct of ω-6-polyunsaturated fatty acids (i.e. arachidonic and linoleic acids) (17, 20-23). The highly reactive HNE molecule, along with other lipid electrophiles, has cytotoxic properties due to forming covalent adducts with proteins. These covalent modifications occur on amino acid residues cysteine (Cys), histidine (His), and lysine (Lys) preferentially via Michael addition reaction (Figure 5) (20, 24-26) resulting in protein carbonylation. In addition, protein carbonylation can also be formed via metal-catalyzed oxidized protein carbonyl products containing glutamic and aminoadipic semialdehydes (27-29). Nevertheless, either route to modification can alter structural or enzymatic functions of proteins, and thus, lead to detrimental consequences (30-33). In this end, HNE has been recognized as a broad biomarker of oxidative stress (34).

Oxidative Stress and Protein Carbonylation

In general, either an increase in production of reactive oxygen and nitrogen species (ROS/RNS), lack of antioxidant defenses, or both cause oxidative stress. One significant consequence of oxidative damage is protein damage (14, 26). ROS are generally free radicals. Free radicals have unpaired electron(s) and, in order to reach a more stable energy level, they easily pick up electron(s) or hydrogen atoms from another molecule that, in turn, converts those molecules into free radicals, thus, setting up a chain reaction. Once this process begins, it can cascade, resulting in tissue and cell damage and ultimately lead to cell death (35). In addition to this cascade of events, several inflammatory and age-associated disease states, including macular degeneration, muscular dystrophy, cancer, diabetes, sepsis, cardiovascular disease and neurological disorders, have been linked to this oxidative stress-induced damage (21, 23, 34, 36-41).

The knowledge of the biochemistry and molecular pathology of oxidative stress is progressively increasing. There is evidence that the overproduction of ROS can cause multiple events including direct oxidative modification of proteins and interaction with sugars and lipids (42). One area that has become of interest to study in regards to oxidative stress is the lipid peroxidation byproduct HNE. As mentioned in the previous section, HNE modification has been well characterized and shown to react with specific amino acid residues (Figure 5) (42, 43). Furthermore, the authors of a recent review (17) have highlighted that an excessive amount of HNE can be detected in a variety of pathologies. It has further been implicated that lipid peroxidation may be an underlying contributor towards the progression of these diseases. However, the roles HNE plays in cell death, inflammation, and atypical cell proliferation are poorly understood (17). These processes in an organism are important in maintaining

homeostasis. Therefore, studying how HNE contributes in these processes may provide insight into how HNE affects physiological and pathological conditions (44, 45).

Biological Model Systems: Cyanobacteria and Zebrafish

(1) Cyanobacteria (Synechocystis sp. PCC6803)

Cyanobacteria, also known as blue-green algae, are microorganisms that structurally resemble bacteria. Unlike bacteria, they contain chlorophyll *a* and conduct oxygenic photosynthesis. Cyanobacterial phyla are about 3 billion years old, making them the oldest oxygenic phototrophs on earth (46). These microorganisms can be found virtually anywhere near the surface of earth from Antarctica to hot springs. Consequently, they have learned to adapt and survive in a vast array of conditions and their metabolism along with lifestyle has become responsive to these changing environments (47).

Synechocystis sp. PCC6803 (Figure 6) is a strain of cyanobacterium that was first isolated from fresh water lakes and deposited in the Pasteur Culture Collection (PCC) in 1968. In the 1980s, Professor Sergey Shestakov and coworkers in Moscow, along with Dr. John Williams at Michigan State University and Du Pont, recognized the PCC6803 strain to be spontaneously transformable, able to integrate foreign DNA into its genome by homologous recombination, and able to survive and grow in a wide range of conditions. In 1996, *Synechocystis* PCC6803 was the first photosynthetic organism for which the entire genome had been sequenced and determined as well as made available in a useful format on a website named CyanoBase (48, 49). This model cyanobacterium has been used quite extensively among scientists around the world in many studies, ranging from pigment synthesis and its regulation, to carbon metabolism, to photosynthesis and respiration. Due to the wide array of subject areas currently being researched, cyanobacteria are an appropriate model to study the effects of lipid stress.

(2) Zebrafish (Danio rerio)

Fish are the most numerous and phylogenetically diverse group of vertebrates (50). They can reveal to us important principles about fundamental processes in vertebrate evolution, development, and disease (50-54). The zebrafish *Danio rerio* is one vertebrate model which can be used to detect environmental hazards (50, 53), and can be used in studies that look at the molecular mechanisms of development, pathology, and pharmacology (50, 52).

The zebrafish genome includes genes that have human orthologs (genes in different species that evolved from a common ancestral gene). Furthermore, there are large regions of synteny between the human and zebrafish chromosomes (50, 52), which allow for the relevance of zebrafish to the analysis of higher vertebrates. This exceptional model organism embraces some innate qualities including high fecundity and fertility, optical clarity of the embryo, and the ability to be easily manipulated and maintained (Figure 7) (50, 51, 53, 54). Because of these natural characteristics, studies are being directed towards observing the roles genes and signaling pathways play in neuronal development, apoptosis, and oxidation (50, 53). Accordingly, understanding these roles will potentially provide valuable information for drug screening, target validation, and toxicological studies (50, 53). Moreover, zebrafish models will allow for high confidence proteomic analysis given the availability of its sequenced genome (55).

Experimental Strategy: Discovery-driven Proteomics

The past few decades have seen the advent of new approaches within the scientific community, including, genomics and, most recently, proteomics. Proteomics is the study of the proteome of an organism. This area has become an increasingly developed field not only comprising all proteins of a cell but also incorporating everything “post-genomics.” These “post-

genomic" features include but are not limited to protein isoforms and modifications, proteins interactions, protein structure, and higher-order complexes (56, 57). With a dynamic proteome attracting a wide array of possible variations, challenges are unavoidable, necessitating combinations of different research tools to understand these complexities.

Proteomics approaches involve high-throughput, data-driven methodologies alongside statistical and computational tools (58). Specifically, these tools have included the separation and visualization of proteins using 2-dimensional PAGE and gel-free profiling procedures, both of which are coupled to mass spectrometry for identification. In addition, large protein databases and computer software have contributed to further processing and validating data collected from the aforementioned tools. Fundamentally, this field provides discovery-based tools for biological research, with the capacity to characterize and survey large datasets of proteins and facilitate new regulatory data at the protein level (59).

The following sub-sections will include background to each technique in the discovery-driven proteomics approach applied in this dissertation. In this respect, the methodology discussed below allowed for surveying large datasets from biological systems in order to determine interactions and regulatory information occurring at the protein level.

(1) Mass Spectrometry-based Proteomics: The "Shotgun" Approach

Essentially, a mass spectrometer measures the masses of molecules carried out in the gas phase and separates the ions based on their mass-to-charge ratio (*m/z*). The information acquired from the mass spectrometer enables protein structural information (i.e. peptide masses or amino acid sequences) to identify proteins. Furthermore, mass spectrometry can be used to determine

the locations and types of modifications of the analyte (60). These characteristics make mass spectrometry the most comprehensive and versatile tool in protein chemistry and proteomics.

There are two primary approaches to analyze proteomes by mass spectrometry. These include discovery-based identification and targeted quantification. The goal of discovery-based experiments is to identify as many proteins as possible. With targeted proteomics, experiments are created to monitor only a select few proteins of interest with the high throughput capabilities of mass spectrometry. The standard workflow for either method begins by extracting proteins from a sample using biochemical methods. Then a protease is used to cut up the protein into peptides at defined amino acid residues. The peptides are then separated by liquid chromatography and the fractions are analyzed by mass spectrometry. The mass spectrometer measures and reports the mass-to-charge ratio (m/z) of the peptide ions (61).

As previously mentioned, the cyanobacteria and zebrafish genome have been sequenced and annotated (49, 55). In addition, the availability of the genome sequences has facilitated biological research in many areas concerning these organisms. On the other hand, there is limited knowledge of the proteome of these two biological model systems. Proteins are important biofunctional molecules in living organisms. These molecules regulate how the organism will react in environmental conditions. From this perspective, we decided to utilize the discovery-based mass spectrometry techniques to elucidate the diverse biological functions of lipids as well as stress and its consequences in the two biological model systems at the protein level.

There are two approaches to analyze samples: bottom-up and top-down (Figure 8). The bottom-up method analyzes proteins that have been enzymatically digested into peptides before mass spectrometric analysis. In the top-down approach, an intact protein is analyzed by mass spectrometry (not applied in this dissertation). Fundamentally, the bottom-up technique identifies

proteins based on peptide sequencing. Analysis of a mixture of proteins using the bottom-up approach is called “shotgun” proteomics (62). Generally, proteolytic enzymes (i.e. trypsin) digest proteins into peptides and the peptide mixtures are fractionated by reversed-phase liquid chromatography followed by mass spectrometry (RP-LC/MS) analysis. Tandem mass spectra derived from peptide fragmentation is collected and subsequently compared to databases by appropriate algorithms. This strategy has proven to be suitable for complex samples and identification of posttranslational modifications (62).

(2) Proteomics: From Gel-based to Gel-free Technologies

The success of mass spectrometry in proteomics is based upon the sample preparation prior to instrumental analysis. In short, the less complex a sample is, the better the examination of it from the mass spectrometer will be. Fractionation can be completed either by 1- or 2-dimensional polyacrylamide gel or by chromatographic techniques. These methodologies have become useful for targeted and global proteomic studies (62). For gel-based techniques to be effective, bands or spots containing proteins are excised from the gel matrix and subsequently “in-gel” digested with a protease. Then, peptides can be extracted from the gel for mass spectrometric analysis. It has been acknowledged that this approach has some disadvantages including limited dynamic range, meaning only the most abundant proteins are detected. Therefore, an alternative approach to gel-based methods is gel-free proteomics. Gel-free methods of separation include high performance liquid chromatography (HPLC) coupled to mass spectrometry. In general, liquid chromatography (LC) separates the biological matrices and their individual components, which are then identified by the mass spectrometer. Types of HPLC include ion exchange, reversed phase, and affinity materials. However, the most powerful in

terms of liquid chromatography mass spectrometry (LC-MS) procedures is reversed-phase (RP) chromatography. Since the gel has been avoided in these methods, a greater number of proteins in the mixture can be represented (60). The standard preparation for LC-MS analysis involves denaturing protein mixtures by reduction with dithiothreitol (DTT), followed by alkylation of the Cys residues and subsequent trypsin digestion overnight to generate peptides (Arg or Lys residues at the C terminus). Ultimately, both of these approaches have afforded value to proteomics studies and should therefore be regarded as complements rather than replacements of one another (63).

(3) Quantitative Label-free Proteomics

In the last decade, there has been promising progress in proteomics as a tool for analyzing complex protein samples. As a result, strategies for label-free and label-based (not applied in this dissertation) quantitative proteomics (Figure 9) have been added to the toolbox of methods in this field. Label-based methods include modifying peptides with isobaric tags for relative and absolute quantitation (iTRAQ), protein labeling with isotope-coded affinity tags (ICAT), or stable isotope labels with amino acids in cell culture (SILAC) (64). These strategies are considered more accurate in quantitating protein abundance (65). However, their cost (expensive isotope labels) and limitation in the number of samples that can be analyzed in one experiment make them harder to manage.

Consequently, advances in understanding the correlation between protein abundance and peak areas (66, 67) or the number of tandem mass spectra (MS/MS) (68) have led to the creation of additional analytical procedures for quantitative proteomics (64, 69). For bottom-up proteomics, two groups have been designated for the label-free methods (Figure 10). The first

approach encompasses the area under the curve (AUC) or signal intensity measurement based on the precursor ion spectra. This process measures ion abundances at specific retention times for the given peptides without using stable isotopes; this is referred to as “ion counts”. As ionized peptides are eluted from the reverse-phase column into the mass spectrometer, their ion intensities are measured, and differentially expressed peptides are validated by liquid chromatography tandem mass spectrometry (LC-MS/MS) either subsequently or simultaneously (64). The second method of the label-free technique is denoted as “spectral counting”. The approach is based on the concept that peptides that are more abundant will produce a higher number of MS/MS spectra, and is therefore relative to protein abundance in data-dependent acquisition (68). Both label-free techniques produce large data volumes and require statistical assessment. Furthermore, in order to reduce the long hours of manual validation, sophisticated software algorithms have been designed and made available to compile and process the data faster.

Moving forward from large datasets, biological conclusions from label-free quantitative data can be extracted and analyzed further for biologically meaningful conclusions. Specifically, this necessitates functional analysis by gathering Gene Ontology (GO) annotations for the proteins identified in datasets (70). As the availability of sequenced genomes increases, annotations for core biological process, molecular functions, and cellular components become available for associated proteins as well. Furthermore, data from the annotations collected from associated proteins can be integrated into networks and pathways via bioinformatic tools. The results generated from these software tools may shed light on the complex biological and chemical systems of the organisms being studied (58, 71).

RESEARCH GOALS AND OBJECTIVES

Study of Lipid Stress in Different Model Biological Systems

This dissertation involves undertaking a large-scale, mass spectrometry-based proteomics approach. Overall, we sought to elucidate the broader role of lipids and lipid peroxidation-related carbonyl stress along with its influences in the direction of biological and cellular processes in two biological systems. Aggressive environmental conditions can cause stress effects in many organisms. This strain on the system can then trigger a cascade of events such as the depletion of cellular function. Potentially, these actions can lead to an onset of disease states. The "impact" left on the proteome during stress (i.e. lipid stress) is not fully established and is therefore crucial to study. Integrating a discovery-driven bottom-up proteomics approach, we first aimed to facilitate an understanding of the global role lipids (specifically phosphatidylglycerol) play in a genetically modified cyanobacterial strain. Furthermore, we aspired to observe insights of the underlying biochemistry of lipid electrophile stress, explicitly the lipid peroxidation byproduct HNE and its implications toward human disease in a higher-order vertebrate model (zebrafish). It has become evident that investigating individual genes or proteins one at a time will no longer suffice in understanding a biological system as a whole. Therefore, integrating discovery-driven methodologies we can consider the behavior and relationships of all the components in a particular biological system and understand the details of the organisms' proteomes with respect to the responses to environmental changes (i.e. lipid stress).

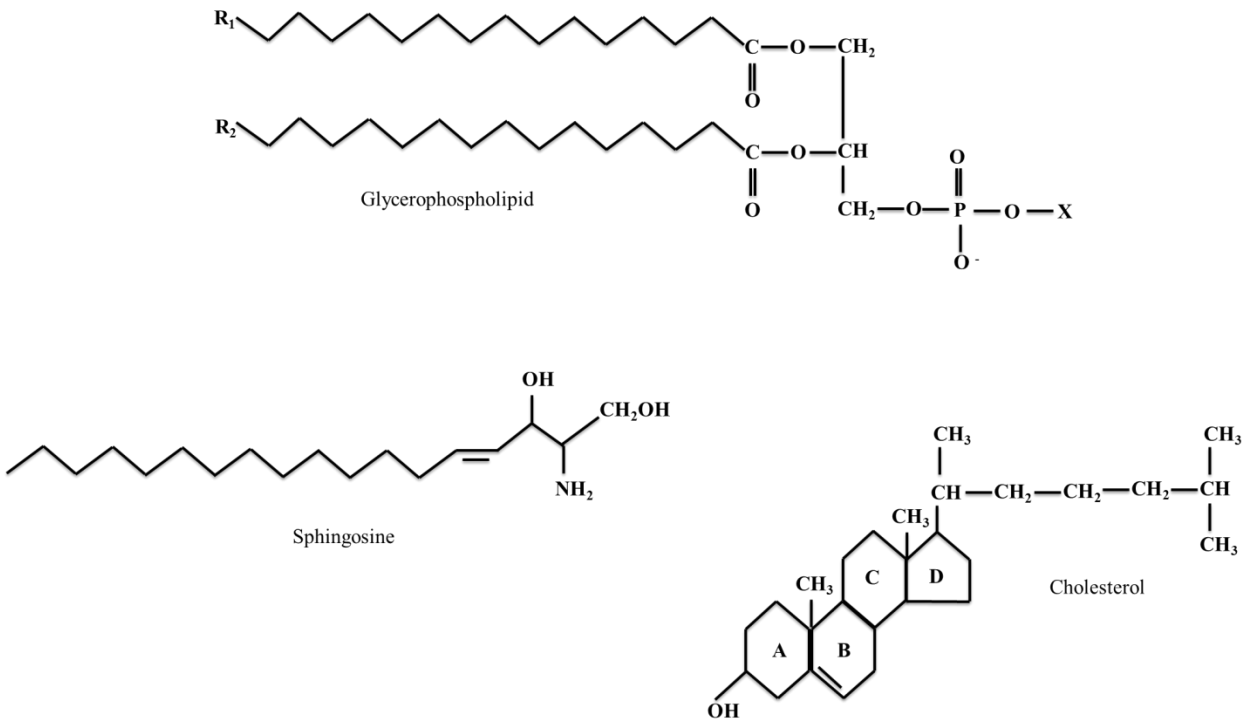


Figure 1. Chemical structures of amphipathic lipids. There are three major types of amphipathic structural lipids found in membranes: glycerophospholipids, sphingolipids, and cholesterol.

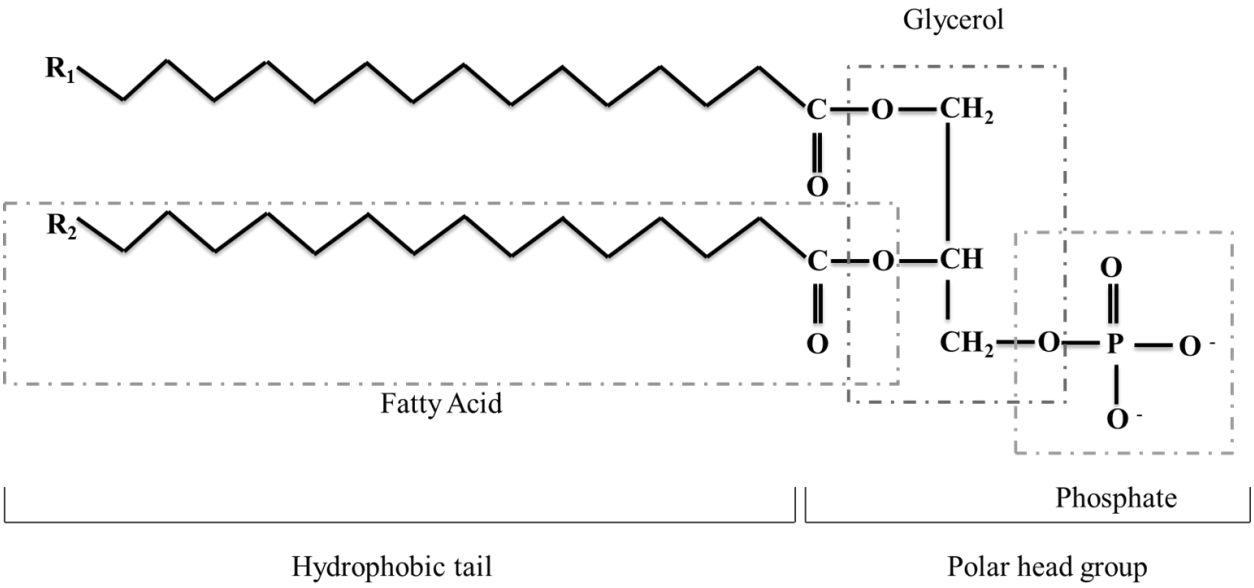


Figure 2. Chemical structure of a phospholipid. Glycerophospholipids contain two fatty acids attached via an ester linkage to positions 1 and 2 of the glycerol backbone. The fatty acids may be different from each other, and are designated R1 and R2. The third carbon of glycerol is linked to a phosphate group forming the polar head group (phosphatidic acid pictured).

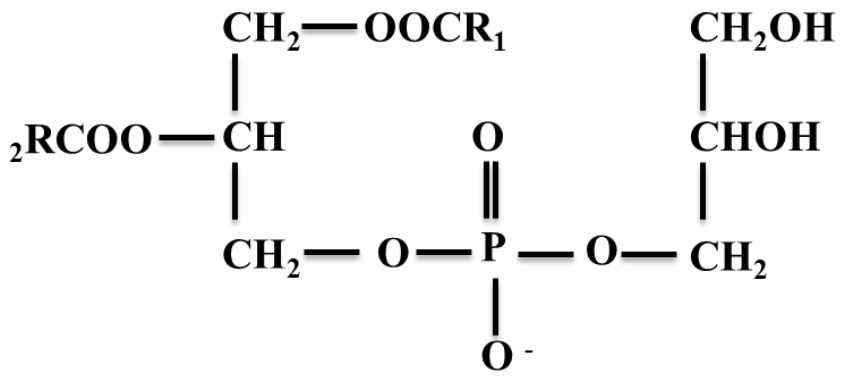
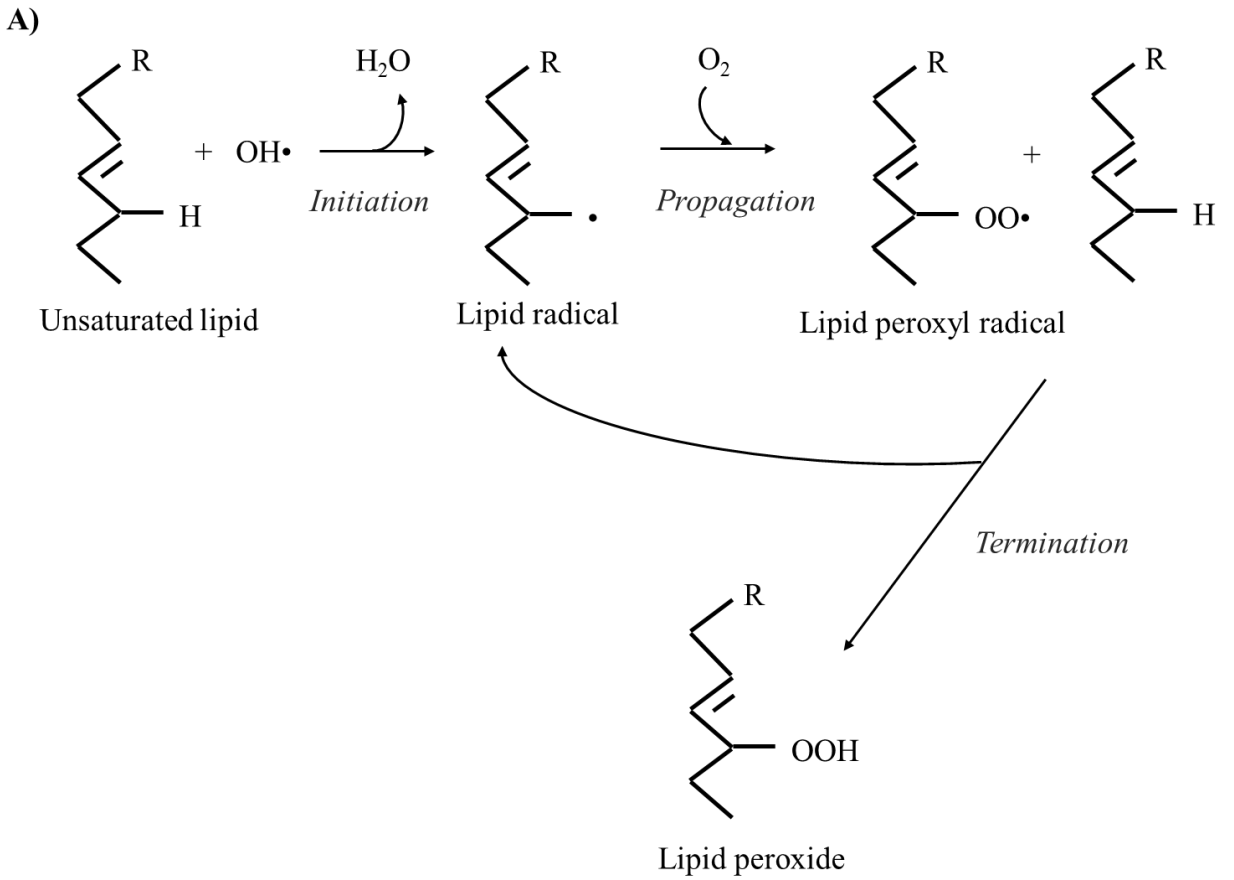
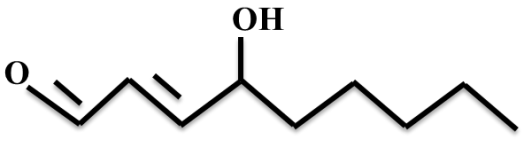


Figure 3. Chemical structure of phosphatidylglycerol (PG). PG is a glycerophospholipid, consisting of an L-glycerol 3-phosphate backbone attached via an ester linkage to either saturated or unsaturated fatty acids on carbons 1 and 2. The phosphate group is negatively charged at neutral pH further classifying it as an anionic phospholipid. PG is also one of the main phospholipids in membranes of bacteria, plants, and animals.



B)



4-Hydroxy-2-nonenal, HNE, $C_9H_{16}O_2$

Figure 4. Lipid peroxidation. A schematic representation of the mechanism of a free radical chemical reaction affecting polyunsaturated fatty acids. The process involves three major steps: initiation, propagation, and termination. Resulting lipid radicals rapidly react with molecular

oxygen to form peroxyl radicals. Peroxidation then initiates a series of complex, autocatalytic propagation reactions that generate a variety of lipid electrophiles (A). Such constituents include the α,β-unsaturated hydroxyalkenal, 4-Hydroxy-2-nonenal (HNE) (B).

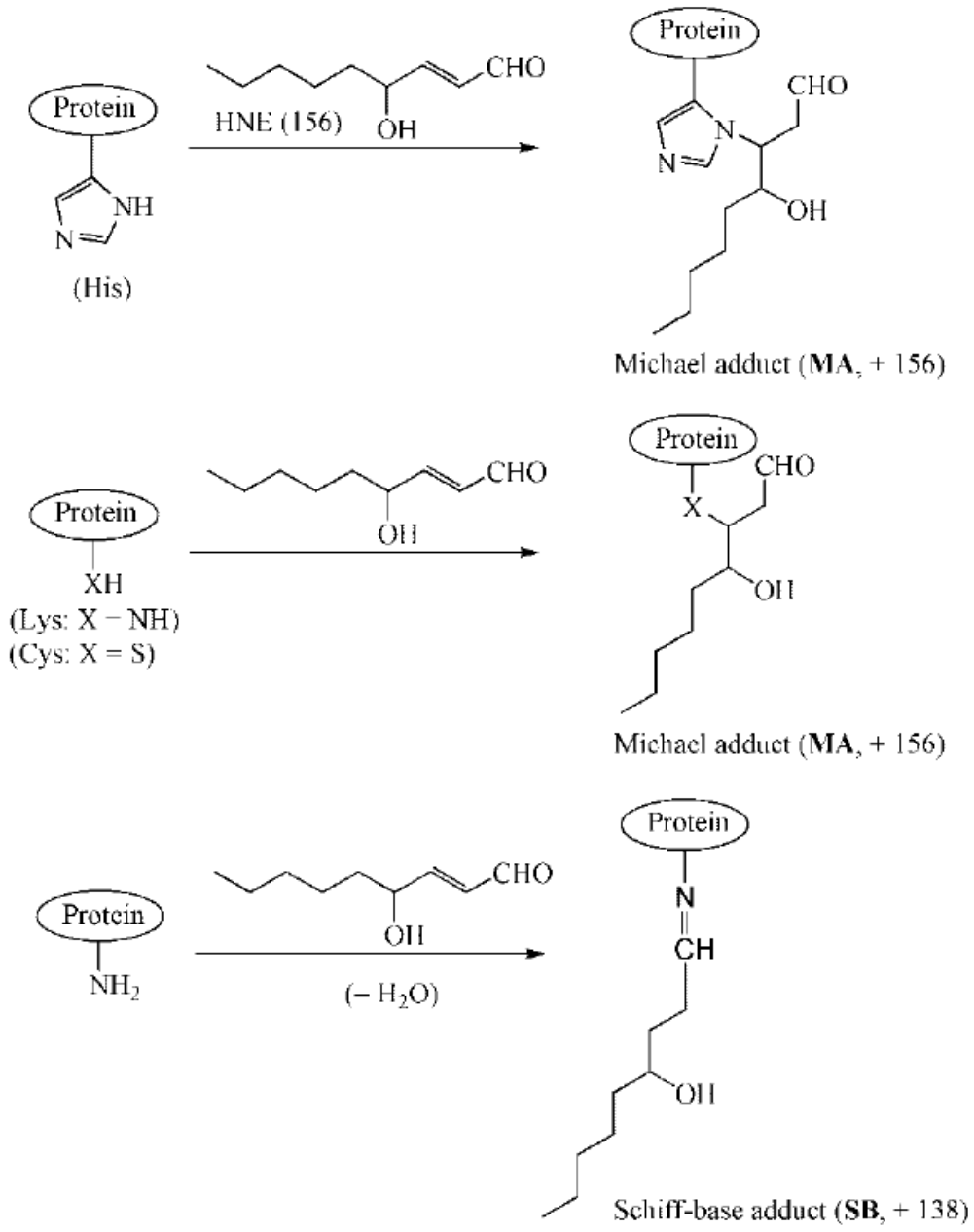


Figure 5. A schematic representation of the chemistry of HNE modification to proteins.

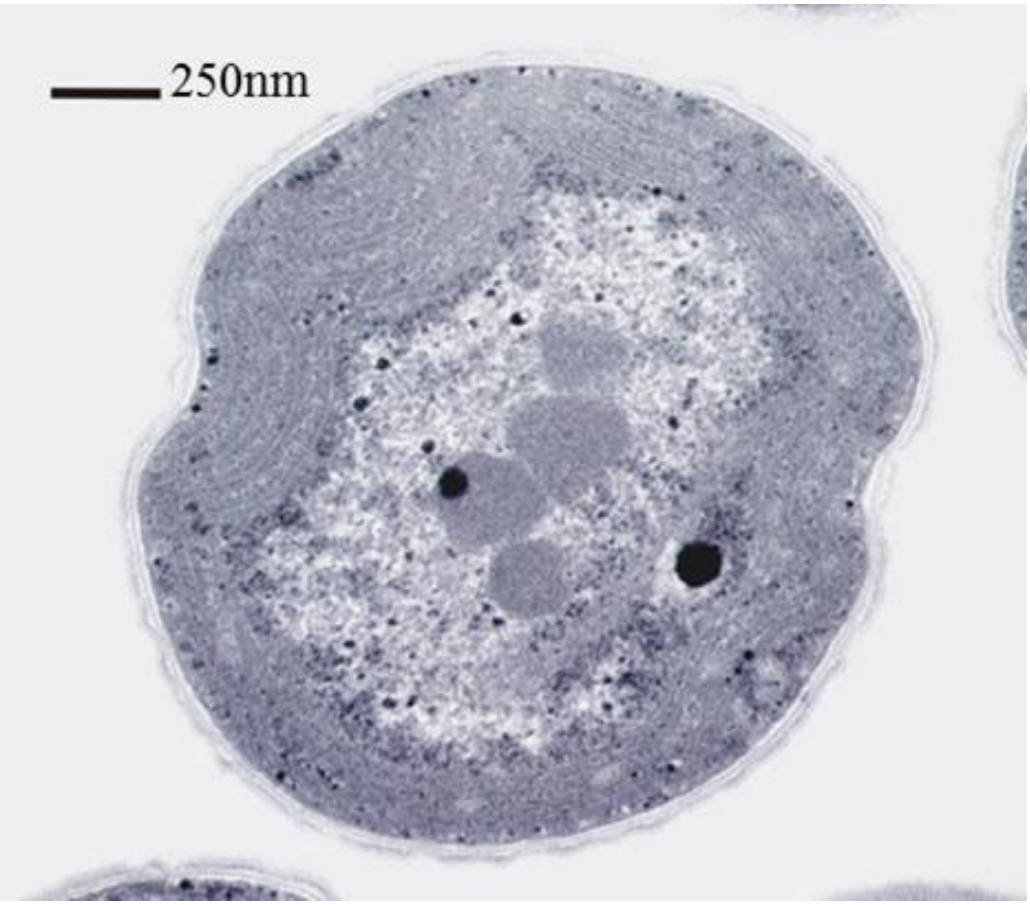


Figure 6. Micrograph of *Synechocystis* sp. PCC6803. Credit to Hajnalka Laczko-Dobos, Institute of Plant Biology, Biological Research Center, Hungarian Academy of Sciences.

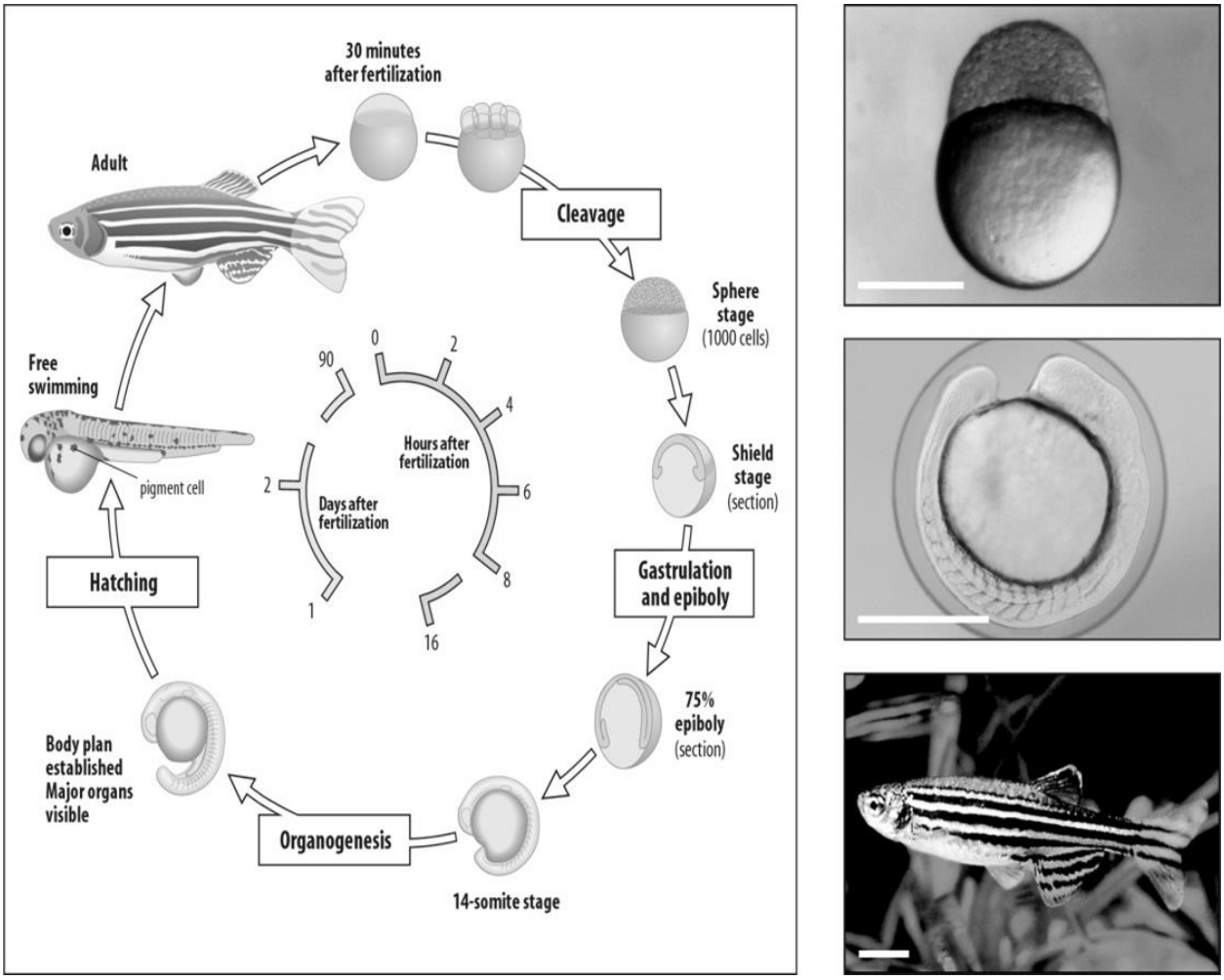


Figure 7. **Diagram of zebrafish (*Danio rerio*) development**. All vertebrate embryos undergo fertilization, cleavage, blastulation, and gastrulation. The eggs have large yolks that provide nutrients to the developing embryo. The zebrafish have a short development cycle of (~12 weeks) and the transparent embryo (0.7 mm) has many advantages (72).

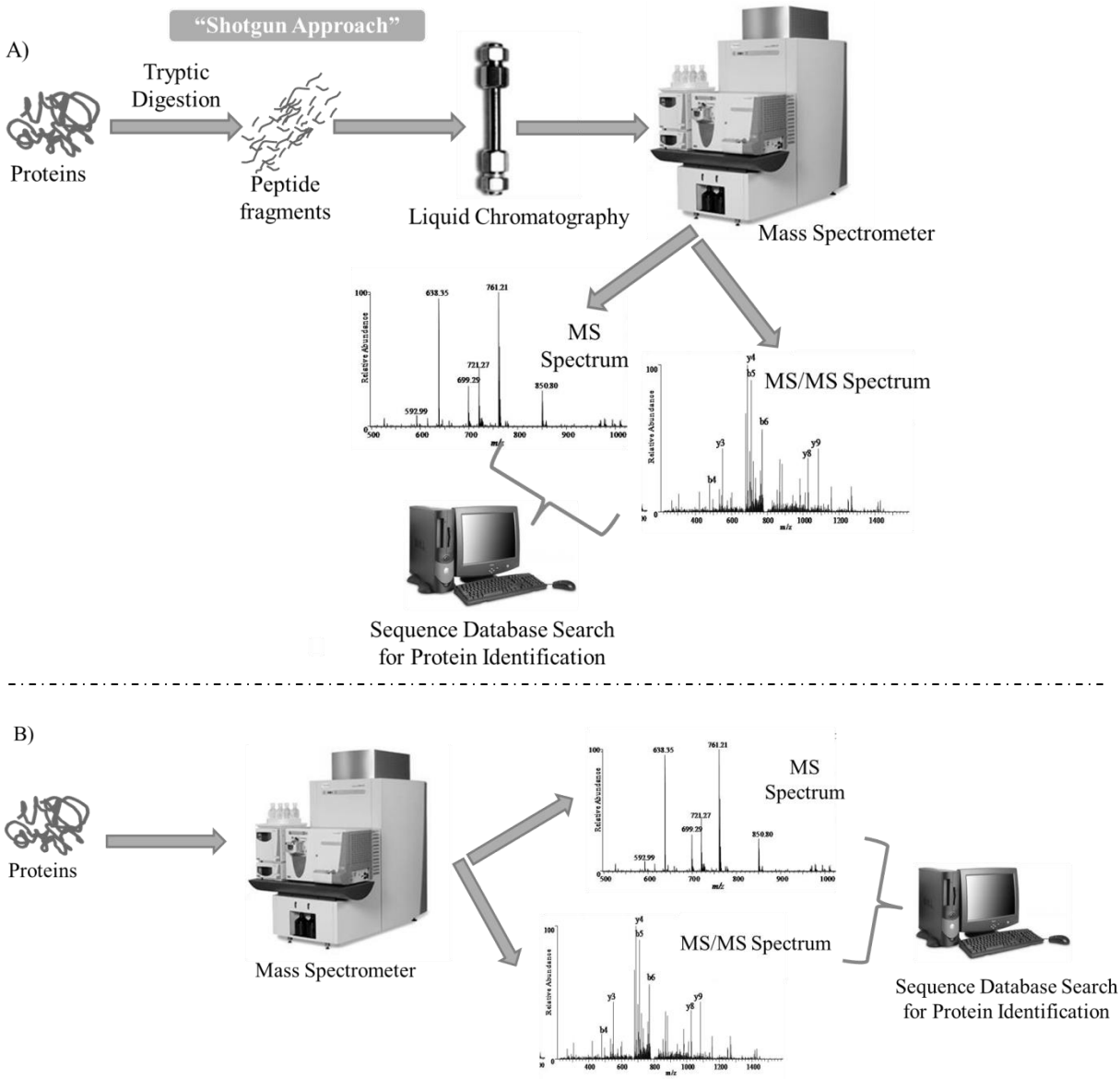


Figure 8. Bottom-up and top-down methodologies. Analysis of a mixture of proteins using the bottom-up approach is called "shotgun" proteomics. First, proteolytic enzymes (i.e. trypsin) digest proteins into peptides and the peptide mixtures are then fractionated by liquid chromatography coupled with mass spectrometry. Tandem mass spectra derived from peptide fragmentation are collected and subsequently compared to databases by appropriate algorithms for identification of peptide sequence information and/or posttranslational modifications (A). The top-down method dissects sequence information of proteins by mass spectrometry. Intact

proteins are directly introduced into the mass spectrometer, fragmented, and compared to a database for identification of protein sequences and/or posttranslational modifications.

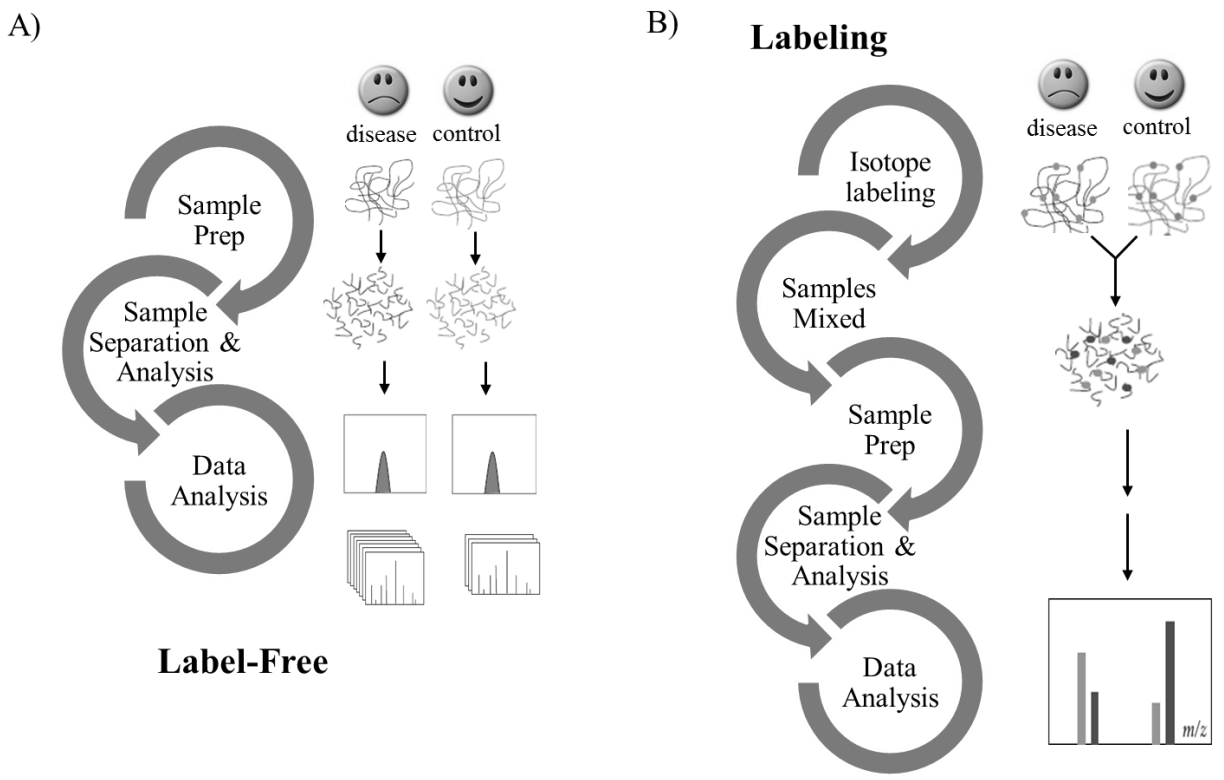


Figure 9. Label-free vs. Labeling Proteomics Approaches. In the label-free approach, each sample is separately prepared and then subjected to individual LC-MS/MS analysis followed by data analysis. Here quantitation is based on either spectral counting of identified proteins or measurements of ion intensity changes (peptide peak areas or peak heights). The change in protein abundance is calculated by direct comparison between different analyses (A). In protein labeling, samples of different conditions are labeled with a specific isotope labeled molecule. After labeling, protein samples are combined together and subjected to sample preparation, which includes protein extraction, reduction, alkylation, and digestion. Samples are then analyzed by LC-MS/MS analysis followed by data analysis (peptide/protein identification, quantification, and statistical analysis).

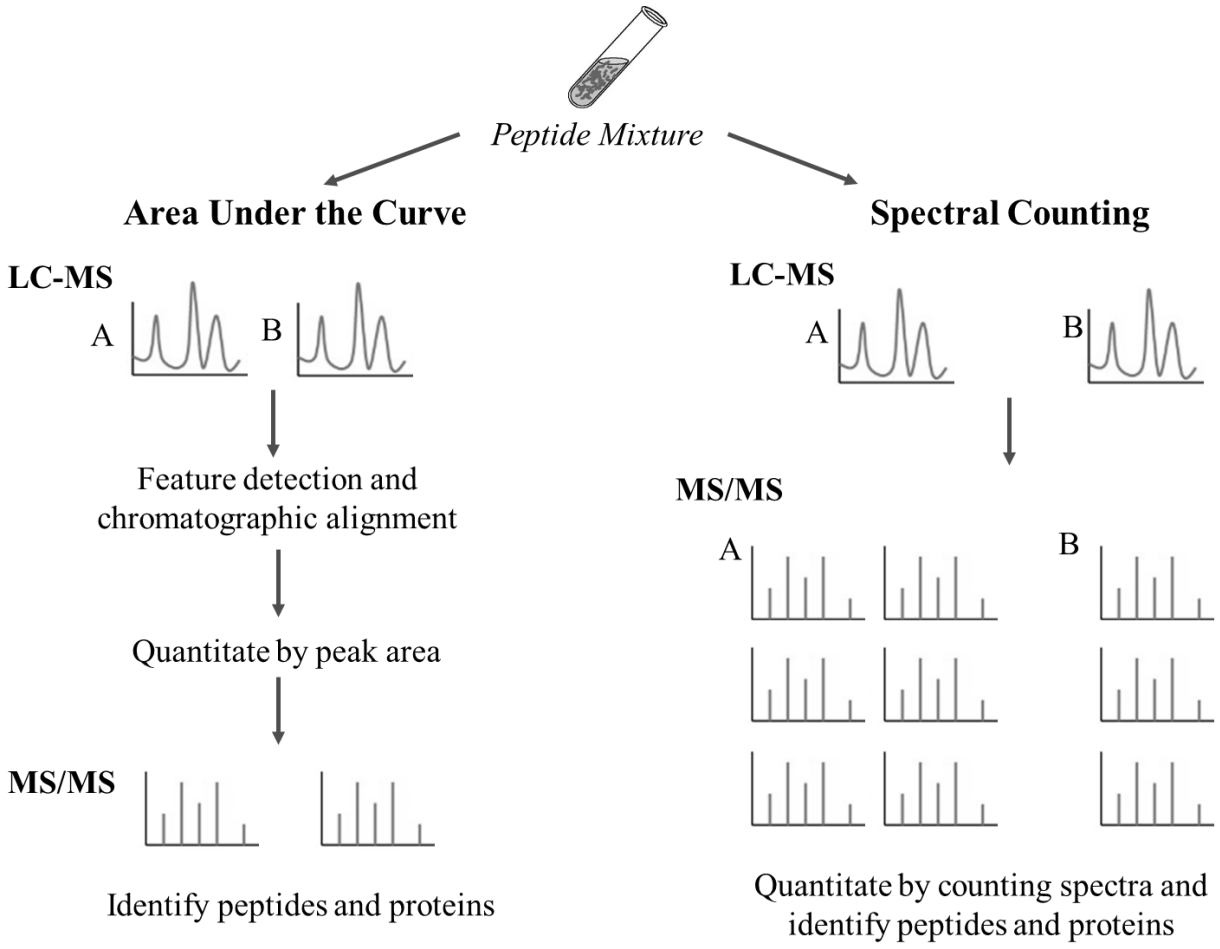


Figure 10. Schematic of the two main label-free strategies. In the LC-MS run, an ion is detected at a specific retention time and recorded with a particular intensity. The signal intensity has been observed to correlate with ion concentration. Therefore, quantifying peptide/protein using the first method of area under the curve (AUC) or signal intensity measurement can be performed. The second method of the label-free technique is referred to as "spectral counting". This approach processes the total number of MS/MS spectra identified for a particular protein as a measure of protein abundance. These two commonly used label-free quantitative approaches are powerful tools for analyzing protein changes in large-scale proteomics studies.

REFERENCES

1. Watson AD. Thematic review series: Systems biology approaches to metabolic and cardiovascular disorders. lipidomics: A global approach to lipid analysis in biological systems. J Lipid Res. 2006 Oct;47(10):2101-11.

2. Nelson DL, Cox MM. Lehninger principles of biochemistry, fifth edition. 5th ed. New York: W.H. Freeman and Company; 2008.

3. Murphy RC, Axelsen PH. Mass spectrometric analysis of long-chain lipids. Mass Spectrom Rev. 2011 Jul-Aug;30(4):579-99.

4. Tvrzicka E, Kremmyda LS, Stankova B, Zak A. Fatty acids as biocompounds: Their role in human metabolism, health and disease--a review. part 1: Classification, dietary sources and biological functions. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2011 Jun;155(2):117-30.

5. Bohdanowicz M, Grinstein S. Role of phospholipids in endocytosis, phagocytosis, and macropinocytosis. Physiol Rev. 2013;93(1):69-106.

6. Takahashi-Umebayashi M, Pineau L, Hannich T, Zumbuehl A, Doval DA, Matile S, et al. Chemical biology approaches to membrane homeostasis and function. Chimia. 2011;65(11):849-52.

7. Kremmyda LS, Tvrzicka E, Stankova B, Zak A. Fatty acids as biocompounds: Their role in human metabolism, health and disease: A review. part 2: Fatty acid physiological roles and applications in human health and disease. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2011 Sep;155(3):195-218.

8. Subramaniam S, Fahy E, Gupta S, Sud M, Byrnes RW, Cotter D, et al. Bioinformatics and systems biology of the lipidome. Chem Rev. 2011 Oct 12;111(10):6452-90.

9. Sakurai I, Hagio M, Gombos Z, Tyystjarvi T, Paakkarinen V, Aro EM, et al. Requirement of phosphatidylglycerol for maintenance of photosynthetic machinery. Plant Physiol. 2003 Nov;133(3):1376-84.

10. Domonkos I, Laczko-Dobos H, Gombos Z. Lipid-assisted protein-protein interactions that support photosynthetic and other cellular activities. Prog Lipid Res. 2008 Nov;47(6):422-35.

11. Matsumoto K. Dispensable nature of phosphatidylglycerol in escherichia coli: Dual roles of anionic phospholipids. Mol Microbiol. 2001 Mar;39(6):1427-33.

12. Schlame M. Cardiolipin synthesis for the assembly of bacterial and mitochondrial membranes. J Lipid Res. 2008 Aug;49(8):1607-20.

13. Pratt DA, Tallman KA, Porter NA. Free radical oxidation of polyunsaturated lipids: New mechanistic insights and the development of peroxyl radical clocks. Acc Chem Res. 2011;44(6):458-67.

14. Dalle-Donne I, Scaloni A, Giustarini D, Cavarra E, Tell G, Lungarella G, et al. Proteins as biomarkers of oxidative/nitrosative stress in diseases: The contribution of redox proteomics. Mass Spectrom Rev. 2005 Jan-Feb;24(1):55-99.

15. Ogino K, Wang DH. Biomarkers of oxidative/nitrosative stress: An approach to disease prevention. Acta Med Okayama. 2007 Aug;61(4):181-9.

16. Zarkovic K. 4-hydroxynonenal and neurodegenerative diseases. Mol Aspects Med. 2003 Aug-Oct;24(4-5):293-303.

17. Poli G, Schaur RJ, Siems WG, Leonarduzzi G. 4-hydroxynonenal: A membrane lipid oxidation product of medicinal interest. Med Res Rev. 2008;28(4):569-631.

18. Wagner BA, Buettner GR, Burns CP. Free radical-mediated lipid peroxidation in cells: Oxidizability is a function of cell lipid bis-allylic hydrogen content. Biochemistry. 1994 Apr 19;33(15):4449-53.

19. Rael LT, Thomas GW, Craun ML, Curtis CG, Bar-Or R, Bar-Or D. Lipid peroxidation and the thiobarbituric acid assay: Standardization of the assay when using saturated and unsaturated fatty acids. Journal of Biochemistry and Molecular Biology. 2004;37(6):749-52.

20. Esterbauer H, Schaur RJ, Zollner H. Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes. Free Radic Biol Med. 1991;11(1):81-128.

21. Sayre LM, Lin D, Yuan Q, Zhu X, Tang X. Protein adducts generated from products of lipid oxidation: Focus on HNE and one. Drug Metab Rev. 2006;38(4):651-75.

22. Rauniyar N, Stevens SM, Prokai-Tatrai K, Prokai L. Characterization of 4-hydroxy-2-nonenal-modified peptides by liquid chromatography-tandem mass spectrometry using data-dependent acquisition: Neutral loss-driven MS3 versus neutral loss-driven electron capture dissociation. Anal Chem. 2009 Jan 15;81(2):782-9.

23. Fritz KS, Petersen DR. An overview of the chemistry and biology of reactive aldehydes. Free Radic Biol Med. 2012 Jun 28.

24. Uchida K, Stadtman ER. Modification of histidine residues in proteins by reaction with 4-hydroxynonenal. Proc Natl Acad Sci U S A. 1992 May 15;89(10):4544-8.

25. Nadkarni DV, Sayre LM. Structural definition of early lysine and histidine adduction chemistry of 4-hydroxynonenal. Chem Res Toxicol. 1995 Mar;8(2):284-91.

26. Butterfield DA, Perluigi M, Reed T, Muharib T, Hughes CP, Robinson RAS, et al. Redox proteomics in selected neurodegenerative disorders: From its infancy to future applications. Antioxidants and Redox Signaling. 2012;17(11):1610-55.

27. Requena JR, Chao CC, Levine RL, Stadtman ER. Glutamic and aminoadipic semialdehydes are the main carbonyl products of metal-catalyzed oxidation of proteins. Proc Natl Acad Sci U S A. 2001 Jan 2;98(1):69-74.

28. Levine RL, Stadtman ER. Oxidative modification of proteins during aging. Exp Gerontol. 2001 Sep;36(9):1495-502.

29. Stadtman ER, Levine RL. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. Amino Acids. 2003 Dec;25(3-4):207-18.

30. Stadtman ER. Protein oxidation in aging and age-related diseases. Ann N Y Acad Sci. 2001 Apr;928:22-38.

31. Yan LJ, Sohal RS. Analysis of oxidative modification of proteins. Curr Protoc Cell Biol. 2002 May;Chapter 7:Unit 7.9.

32. Headlam HA, Davies MJ. Markers of protein oxidation: Different oxidants give rise to variable yields of bound and released carbonyl products. Free Radic Biol Med. 2004 May 1;36(9):1175-84.

33. Aldini G, Dalle-Donne I, Colombo R, Maffei Facino R, Milzani A, Carini M. Lipoxidation-derived reactive carbonyl species as potential drug targets in preventing protein carbonylation and related cellular dysfunction. ChemMedChem. 2006 Oct;1(10):1045-58.

34. Dalle-Donne I, Giustarini D, Colombo R, Rossi R, Milzani A. Protein carbonylation in human diseases. Trends Mol Med. 2003 Apr;9(4):169-76.

35. Dizdaroglu M, Jaruga P, Birincioglu M, Rodriguez H. Free radical-induced damage to DNA: Mechanisms and measurement. Free Radic Biol Med. 2002 Jun 1;32(11):1102-15.

36. Nystrom T. Role of oxidative carbonylation in protein quality control and senescence. EMBO J. 2005 Apr 6;24(7):1311-7.

37. Yuan Q, Zhu X, Sayre LM. Chemical nature of stochastic generation of protein-based carbonyls: Metal-catalyzed oxidation versus modification by products of lipid oxidation. Chem Res Toxicol. 2007 Jan;20(1):129-39.

38. Dalle-Donne I, Carini M, Vistoli G, Gamberoni L, Giustarini D, Colombo R, et al. Actin Cys374 as a nucleophilic target of alpha,beta-unsaturated aldehydes. Free Radic Biol Med. 2007 Mar 1;42(5):583-98.

39. Feng J, Xie H, Meany DL, Thompson LV, Arriaga EA, Griffin TJ. Quantitative proteomic profiling of muscle type-dependent and age-dependent protein carbonylation in rat skeletal muscle mitochondria. J Gerontol A Biol Sci Med Sci. 2008 Nov;63(11):1137-52.

40. Grimsrud PA, Xie H, Griffin TJ, Bernlohr DA. Oxidative stress and covalent modification of protein with bioactive aldehydes. J Biol Chem. 2008 Aug 8;283(32):21837-41.

41. Fritz KS, Petersen DR. Exploring the biology of lipid peroxidation-derived protein carbonylation. Chem Res Toxicol. 2011 Sep 19;24(9):1411-9.

42. Stevens SM,Jr, Rauniyar N, Prokai L. Rapid characterization of covalent modifications to rat brain mitochondrial proteins after ex vivo exposure to 4-hydroxy-2-nonenal by liquid chromatography-tandem mass spectrometry using data-dependent and neutral loss-driven MS3 acquisition. J Mass Spectrom. 2007 Dec;42(12):1599-605.

43. Poli G, Schaur RJ. 4-hydroxynonenal in the pathomechanisms of oxidative stress. IUBMB Life. 2000;50(4-5):315-21.

44. Petersen DR, Doorn JA. Reactions of 4-hydroxynonenal with proteins and cellular targets. Free Radical Biology and Medicine. 2004;37(7):937-45.

45. Butterfield DA, Dalle-Donne I. Redox proteomics. Antioxid Redox Signal. 2012 Dec 1;17(11):1487-9.

46. Olson JM. Photosynthesis in the archean era. Photosynth Res. 2006 May;88(2):109-17.

47. Fleming LE, Broad K, Clement A, Dewailly E, Elmir S, Knap A, et al. Oceans and human health: Emerging public health risks in the marine environment. Mar Pollut Bull. 2006;53(10-12):545-60.

48. Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, et al. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. 1996 Jun 30;3(3):109-36.

49. Kaneko T, Tabata S. Complete genome structure of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. Plant and Cell Physiology. 1997;38(11):1171-6.

50. Spitsbergen JM, Kent ML. The state of the art of the zebrafish model for toxicology and toxicologic pathology research--advantages and current limitations. Toxicol Pathol. 2003 Jan-Feb;31 Suppl:62-87.

51. Lemeer S, Pinkse MW, Mohammed S, van Breukelen B, den Hertog J, Slijper M, et al. Online automated in vivo zebrafish phosphoproteomics: From large-scale analysis down to a single embryo. J Proteome Res. 2008 Apr;7(4):1555-64.

52. Lucitt MB, Price TS, Pizarro A, Wu W, Yocum AK, Seiler C, et al. Analysis of the zebrafish proteome during embryonic development. Mol Cell Proteomics. 2008 May;7(5):981-94.

53. Parng C, Ton C, Lin YX, Roy NM, McGrath P. A zebrafish assay for identifying neuroprotectants in vivo. Neurotoxicol Teratol. 2006 Jul-Aug;28(4):509-16.

54. Tay TL, Lin Q, Seow TK, Tan KH, Hew CL, Gong Z. Proteomic analysis of protein profiles during early development of the zebrafish, danio rerio. Proteomics. 2006 May;6(10):3176-88.

55. Zebrafish genome overview [Internet].: Genome Reference Consortium; 2010 [updated Oct. 5, 2010; cited January 2013]. Available from: http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/zebrafish/.

56. Tyers M, Mann M. From genomics to proteomics. Nature. 2003 Mar 13;422(6928):193-7.

57. Buhimschi CS, Weiner CP, Buhimschi IA. Clinical proteomics: A novel diagnostic tool for the new biology of preterm labor, part I: Proteomics tools. Obstetrical and Gynecological Survey. 2006;61(7):481-6.

58. Zhang W, Li F, Nie L. Integrating multiple 'omics' analysis for microbial biology: Application and methodologies. Microbiology. 2010;156(2):287-301.

59. Hayduk EJ, Choe LH, Lee KH. Proteomic tools in discovery-driven science. Curr Sci. 2002;83(7):840-4.

60. Graves PR, Haystead TA. Molecular biologist's guide to proteomics. Microbiol Mol Biol Rev. 2002 Mar;66(1):39,63; table of contents.

61. Doerr A. Mass spectrometry-based targeted proteomics. Nature Methods. 2013;10(1):23.

62. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR,3rd. Protein analysis by shotgun/bottom-up proteomics. Chem Rev. 2013 Feb 26.

63. Abdallah C, Dumas-Gaudot E, Renaut J, Sergeant K. Gel-based and gel-free quantitative proteomics approaches at a glance. International Journal of Plant Genomics. 2012;2012.

64. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, et al. Less label, more free: Approaches in label-free quantitative mass spectrometry. Proteomics. 2011 Feb;11(4):535-53.

65. Gevaert K, Impens F, Ghesquiere B, Van Damme P, Lambrechts A, Vandekerckhove J. Stable isotopic labeling in proteomics. Proteomics. 2008 Dec;8(23-24):4873-85.

66. Bondarenko PV, Chelius D, Shaler TA. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography - tandem mass spectrometry. Anal Chem. 2002;74(18):4741-9.

67. Chelius D, Bondarenko PV. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. Journal of Proteome Research. 2002;1(4):317-23.

68. Liu H, Sadygov RG, Yates III JR. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem. 2004;76(14):4193-201.

69. Zhu W, Smith JW, Huang CM. Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol. 2010;2010:840518.

70. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: Tool for the unification of biology. Nat Genet. 2000;25(1):25-9.

71. McKenna NJ. Discovery-driven research and bioinformatics in nuclear receptor and coregulator signaling. Biochimica et Biophysica Acta - Molecular Basis of Disease. 2011;1812(8):808-17.

72. Vertebrate development I: Life cycles and experimental techniques [Internet].; cited January 2013]. Available from: http://www.mun.ca/biology/desmid/brian/BIOL3530/DEVO_03/devo_03.html.

CHAPTER II

PROTEOMIC ANALYSIS IMPLICATES PHOSPHATIDYLGLYCEROL AS A BROAD REGULATOR OF CELLULAR PROCESSES IN *SYNECHOCYSTIS* SP. PCC6803

Bottom-up label-free methods have become an emerging strategy for the straightforward and convenient quantitative survey of a complex proteome. In this chapter, we aspired to identify phosphatidylglycerol-regulated proteins on a global scale (i.e., directly in biological samples of high complexity) from a crude cell lysate of cyanobacteria utilizing this proteomics approach, for the first time.

INTRODUCTION

Photosynthesis is the basic energy source of cyanobacteria and other photosynthetic organisms. Phosphatidylglycerol (PG) is the only phospholipid that is an indispensable constituent of photosynthetic complexes (1-3). It is important for both the formation and functioning of the photosynthesis machinery (4). Studies have shown that cyanobacterial cells are unable to divide in the absence of PG, further supporting the claim that this phospholipid is a requirement for the formation and definition of the division site (5). Additionally, compiled supplementary data, using mutants, further demonstrated that PG is located at functionally and

structurally important sites of the reaction centers of Photosystems I (PSI) and II (PSII) of the photosynthetic apparatus (5, 6). Photosystems are protein complexes where light is absorbed and transferred into energy and electrons. They are located in thylakoid membranes of plants, algae, and cyanobacteria. X-ray crystallography analyses also revealed the presence of an enrichment of lipid molecules bound to PSII as compared to PSI, thus suggesting that PSII may play a significant role in the regulation of the second photosystem complex (5, 6). As an essential element of the photosystems, PG may also assist in specific and imperative roles towards repair and assembly processes (5, 6). Changing environmental conditions also affects the biosynthesis of membrane lipids in cyanobacteria. This further suggests that membrane lipids participate in active roles of adaptive processes (7).

Biochemical and molecular genetics approaches have successfully identified genes that encode enzymes responsible for carrying out the biosynthesis of PG, in addition to the generation of mutants defective in the biosynthesis of the lipid. Together, these outcomes have contributed significantly to understanding the function of this phospholipid in photosynthetic organisms such as cyanobacteria and eukaryotic plants (5, 6). The *pgsA* mutant of *Synechocystis* sp. PCC6803 is a representative strain of cyanobacteria. The mutant has shown that it is unable to produce yet able to grow in the presence of exogenously added PG and survives for an extended period after the phospholipid is withdrawn (8). Therefore, this genetically modified strain has become a valuable model system to study the importance of PG in photosynthesis and other cellular processes. However, many of the functional details in these processes remain limited. Therefore, the next logical step in understanding functional aspects of these cells is to study the proteins encoded by this cyanobacterial genome

Mass spectrometry has become an essential tool for proteomics (9). This technology allows for the profiling of proteins in a biological system on a global-scale ("shotgun proteomics"), and has provided another approach for obtaining information about the level of protein expression (10). Additionally, detecting differentially expressed proteins in a complex biological sample also reveals details of functional regulation on the protein level of the organism. In this chapter, we aimed to facilitate the numerous studies mentioned above with a discovery-driven approach utilizing bottom-up label-free methods to reveal any dynamic changes in the *Synechocystis* proteome on a global scale during PG-regulation (i.e., PG-depletion). The benefit of this discovery-driven approach demonstrates the biological importance and diverse role of this phospholipid in cyanobacteria.

EXPERIMENTAL STRATEGY

Organism and Culture Conditions

Cells of the Δ*pgsA* mutant strain from *Synechocystis* sp. PCC6803 were grown photoautotrophically as previously described (4). The *pgsA* mutant cells were harvested at 21 days with or without exogenous PG (PG+ and PG-, respectively).

Protein Extraction and One-dimensional SDS-gel Separation

Total protein content was determined by the microLowry method (11). Cyanobacterial cells corresponding to 2.5 mg of protein were mixed with sample buffer containing 3% SDS, ground in liquid nitrogen using a mortar and pestle, and subsequently heated at 60 °C for 10 min. For one-dimensional separation, proteins were loaded on a 10% denaturing polyacrylamide gel; no markers were used. Gels were washed 3 times for 10 min in deionized water and then stained

with Bio-safe CBB (Bio-Rad, Hercules, CA) for 1h. Following this, gels were destained twice for 1h in deionized water and subsequently cut into 5 almost equal protein-containing pieces. One of the advantages of this was that most of the phycobiliproteins (water-soluble protein components of the light harvesting antennae of cyanobacteria and certain algae) are accumulated in one fraction. The idea came from (12) and in-gel digestion was based on this.

In-gel Digestion

Destained gel pieces were further subjected to in-gel digestion as previously reported (13). Briefly, gel pieces were dried and proteins were reduced, alkylated, and digested with Trypsin (Applied Biosystems, Foster City, CA) overnight. Tryptic peptides were extracted from the gel pieces and dried in a vacuum concentrator before mass spectrometric analyses.

Liquid Chromatography–tandem Mass Spectrometry

Data-dependent MS and MS/MS acquisitions for in-gel digests were generated using a hybrid linear quadrupole ion trap–Fourier transform ion cyclotron resonance (7-T) mass spectrometer (LTQ-FT, Thermo Finnigan, San Jose, CA) equipped with a nano-electrospray ionization source and operated with Xcalibur (version 2.2) and Tune Plus (version 2.2) data acquisition software. Online reversed-phase high performance liquid chromatography (RP-HPLC) was performed with an Eksigent nano-LC-2D (Eksigent, Dublin, CA) system following standard protocol previously reported (14).

Database Search

MS/MS data generated by data dependent acquisition via the LTQ-FT were extracted by BioWorks version 3.3 and searched against a composite *Synechocystis* protein sequence database of cyanobacteria (*Synechocystis*.par) using Mascot version 2.2 (Matrix Science, Boston, MA) search algorithm. Search filters in Mascot were set with a fragment ion mass tolerance of 0.80 Da, a parent ion tolerance of 25 ppm, and one missed cleavage. Fixed modification included carbamidomethylation of cysteine, while oxidation of methionine was specified as a variable modification.

Data Compilation, Relative Quantification, and Analysis of Genome Information

The analysis of PG- and PG+ cells were achieved by triplicate injections of both samples (tryptic digests of PG- and PG+, respectively) on the LTQ-FT mass spectrometer. Spectral counts (MS/MS) were determined through Scaffold (version Scaffold 3.0, Proteome Software Inc., Portland, OR), while extracted peptide intensity (MS) features were generated with Progenesis LC-MS (version 3.0, Nonlinear Dynamics, Durham, NC) for comparative evaluation of the methods. In Scaffold, initial peptide identifications were accepted if they could be established at >95% probability as specified by Peptide Prophet Algorithm (15). In addition, protein identifications were accepted if they could be established at >99% probability, and contained at least two identified unique peptides. The Protein Prophet algorithm assigned protein probabilities, and these identification criteria typically established a <0.01% false discovery rate based on a decoy database search strategy at the protein level (16). The method for label-free relative quantitation followed previously published protocol (17).

Progenesis LC-MS generated extracted peptide intensity (MS) features for additional evaluation of differential protein expression from the same set of raw data files acquired from PG-/PG+ samples. Alignment of the data, based on LC retention time of each analysis, created a single aggregate run containing all MS data with representing peptide ions and, then, generated feature outline maps for detection and quantification of peptide ions from individual analyses. Carrying out peptide quantifications, the software first summed then normalized peak intensities, then filtered the data to retain only ions with positive charges (z) of two and three, and exported peak lists to query against the *Synechocystis* protein sequence database using Mascot (see *Database Search* subsection above). By means of analysis of variance (ANOVA, $p<0.05$), proteins were accepted as differentially expressed based on at least a two-fold change in expression, and passing validation by Peptide Prophet and Protein Prophet using protein identifications from Scaffold.

Further analysis of genome information, was completed by subjecting differentially expressed identifications to CyanoBase (http://genome.kazusa.or.jp/cyanobase), a comprehensive and freely accessible web database for cyanobacteria (18). On the GeneView page, annotations including function, biological processes, as well as transmembrane data can be retrieved from the database.

RESULTS AND DISCUSSION

Currently a narrow focus of cyanobacterial research has been dedicated to shotgun proteomics (19). Initial efforts to incorporate this technical approach using this organism were first published in 2005 (12). The study generally focused on evaluating the efficiency of different

fractionation methods. However, these contributions along with other subsequent studies later published further emphasized the need to incorporate shotgun proteomics studies in cyanobacterial exploration (19, 20). As there is an increase in the number of completed cyanobacterial genomes, the next logical approach is to study their proteomes (i.e. according to mutants). Therefore, our scheme of experiments integrated a discovery-driven label-free method to elucidate changes in the proteome of the *pgsA Synechocystis* strain during PG-regulation.

Differential Expression of Proteins during PG-depletion of *pgsA Synechocystis* Strain

Synechocystis PCC6803 *pgsA* mutant cells were harvested at 21 days with or without exogenous PG. Previously, it was reported that cultivation for 4 weeks in PG-deficient medium was lethal for this mutant (4). Samples were then analyzed by in-gel tryptic digestion followed by liquid chromatography-tandem mass spectrometry (GeLC-MS/MS) to further survey the proteome for protein expression levels during PG-regulation. Briefly, SDS-PAGE was performed on crude cellular protein extracts. Gels were then cut horizontally into pieces for subsequent in-gel tryptic digestion and analyzed by RP-LC-MS/MS. This GeLC-MS/MS approach permitted the global-scale identification of PG-regulated proteins by label-free methods. After searching against a composite protein sequence database of *Synechocystis*, 720 proteins were identified in this genetically modified strain of cyanobacteria (list not shown).

Label-free quantification was first investigated in this study by spectral counts from the software Scaffold, using our previously validated method (17). Consequently, considering at least >2-fold change using the total number of MS/MS spectra identified for a particular protein (Scaffold), our investigation obtained four PG-regulated proteins after G-test (a likelihood-ratio

statistical test for independence, $p<0.05$) (Table 1). In addition, all four proteins were shown to be down-regulated in PG-depleted mutants.

Introducing Progenesis LC-MS into our workflow of label-free proteomics has resulted in a far superior performance in terms of revealing a higher proportion of regulated proteins as compared to methods we have employed earlier (17, 21). Spectral counting requires a minimum of spectral counts (protein abundance) and uses stringent G-tests (17) for statistical analyses. This discards possibly significant lower-abundant species, which could be regulated in the given environment. On the other hand, after the same set of raw data files was evaluated for differential protein expression by Progenesis LC-MS software, we revealed that lower-abundance proteins were not compromised during statistical analysis. Following analysis of variance (ANOVA, $p<0.05$), requiring at least a 2-fold change in expression, and excluding hypothetical proteins, 71 proteins were revealed to be differentially expressed (Table 2).

Moreover, Photosystem I reaction center subunit XI (slr1655), Photosystem I reaction center subunit VII (ssl0563), along with the two probable porin proteins (slr1908 and slr1841) identified as differentially expressed (via spectral counting) were consecutively confirmed with peak integrations using Progenesis LC-MS (denoted by $^{+}$ in Table 2). Examples of differentially modulated peptides, which correlate to Photosystem I reaction center subunit XI and probable porin protein, can be found in Figure 1 and Figure 2, respectively.

Although label-free proteomics using both MS/MS spectral counting (Scaffold) and peptide peak intensities (Progenesis LC-MS) on a linear ion trap-FTMS hybrid instrument were suitable to identify differential protein expression for key cellular proteins in the *pgsA* mutant of *Synechocystis* PCC6803, the latter (i.e., an MS-based) approach had far superior performance in terms of revealing PG-regulated species (Figure 3a). In addition, identified proteins had to meet

criteria set by Scaffold (i.e., minimum peptide confidence score >95% and minimum protein confidence score >99% with a minimum of 2 peptides) (Figure 3b). This global-scale survey of modulated PG content has identified even more proteins of the cyanobacterial proteome. This becomes essential to understand PG-regulation in various cellular processes.

Regulation of Protein Distribution in the Photosynthetic Complexes

Previously, it has been recognized that PG is the only phospholipid that is an indispensable constituent of photosynthetic membranes. This lipid has also been demonstrated to influence the efficiency of electron transfer (5). Moreover, PG and digalactosyldiacylglycerol (DGDG) (a glycolipid in thylakoid membranes) interact in photosynthetic processes and construct protein subunits to form PSI and PSII reaction centers. However, many questions concerning PG involvement in photosynthesis remain unanswered. Subsequently, our focus was primarily dedicated to photoautotrophic, oxygen-producing cyanobacteria under stress conditions to identify PG-regulated proteins on a global-scale.

Using the label-free method, 71 differently expressed proteins within this mutant strain were determined. This large dataset still required analysis to further understand the role of PG in this specific proteome. Therefore, genome information of *Synechocystis* (a part of CyanoBase, an online resource for cyanobacterial genomes) was accessed to investigate annotations associated with the PG-regulated proteins. As a result, differentially expressed proteins in PG-depleted cyanobacteria were found to be involved in processes of photosynthesis and respiration, translation, energy metabolism, transport, and binding (Figure 4).

A majority of the proteins discovered from our label-free method were shown to be down-regulated, which is in good correlation with studies reviewed previously (5). The proteins

that showed the strongest changes in abundance were, as expected, subunits of the photosynthetic complexes. Moreover, of the 71 differentially expressed proteins, 11 of them were membrane proteins. These included 2 up-regulated proteins: Periplasmic protein (slr0708) and Cell division protein (slr1604), and 9 down-regulated proteins: Photosystem II manganese-stabilizing polypeptide (sll0427), Photosystem I reaction center subunit III precursor [PSI-F], plastocyanin [cyt c553] docking protein (sll0819), Photosystem II 12 kDa extrinsic protein (sll1194), ATP synthase B chain (subunit I) of CF (0) (sll1324), Photosystem II 11 kD protein (slr1645), Photosystem I reaction center subunit XI (slr1655), Photosystem I reaction center subunit Ib (slr1835), Cytochrome b559 alpha subunit (ssr3541), and Photosystem II extrinsic protein (sll1638).

Out of the nine membrane proteins down-regulated in PG-depleted cells, several have been studied and shown to participate as subunits of the PSI complex. For example, PsaB (slr1835) forms the heterodimeric core of PSI along with its counterpart PsaA (22). Additionally, of the four smaller subunits of the PSI complex, PsaD, PsaE, PsaF (sll0819), and PsaL (slr1655), two were revealed as down-regulated in our study. Similarly, during depletion of PG, PsaL, which is linked to structural features of the PSI complex, has been shown to dissociate the PSI trimer into the monomer (22-25). Furthermore, the other down-regulated membrane proteins included PSII extrinsic proteins PsbO (sll0427) and PsbU (sll1194), both of which contribute to the core assembly at the lumenal side of the PSII complex (5). We also identified Photosystem II extrinsic protein (sll1638, down-regulated), another extrinsic protein of PSII, which has been shown to participate in the regulation and stabilization of the water oxidation complex (26-29). Moreover, if this is a true identifier for PsbQ and participates in photosynthesis, then the

regulation we uncovered here relates with the down-regulation of other proteins involved in photosynthesis (30).

Inherently, PSI and PSII complexes are the key locations where light-driven energy production takes place (22). In addition, photoautotroph organisms such as cyanobacteria utilize energy from light. As a result, dysfunctions in photosynthetic protein expression levels can be detrimental in the survival of these organisms. Similarly, when photosynthesis is not properly working, the organism seeks alternative routes to get energy. *Synechocystis* cells are able to grow heterotrophically (31), but for this, they need sugar (i.e. carbohydrates). In addition, for heterotrophs to grow, they need oxidative electron transport. Subsequently, our data has shown the up-regulation of three proteins, Flavoprotein (sll0550), NADH dehydrogenase subunit 7 (slr0261), and Ferrodoxin-NADP oxidoreductase (slr1643), which are generally produced under specific growth conditions and utilized in the transfer of electrons (respiration).

Also similarly related to the oxygenic electrotransport proteins just mentioned, we found several up-regulated proteins involved in metabolism of carbohydrates. These included 6-phosphogluconate dehydrogenase (sll0329), Putative sugar-nucleotide epimerase/dehydratease (sll0576), NAD(P)-dependent glyceraldehyde-3-phosphate dehydrogenase (sll1342), phosphoribulokinase (sll1525), Phosphoglycerate kinase (slr0394), and Glucose 6-phosphate dehydrogenase assembly protein (slr1734). Intriguingly, four proteins involving carbohydrate metabolism were also down-regulated (sll1028, sll1029, slr0012, slr0623). However, further investigations need to be completed to figure out a network system that involves photosynthesis and the oxidative electron transport, along with their cross-regulation. These results indicate the complexity of the effect of PG-depletion.

In accordance with the intricate effect of PG-depletion, we further noticed cross-regulation of proteins involved in the synthesis of ATP. Up-regulated proteins included ATP synthase alpha chain (sll1326) and Phoshpoglycerate kinase (slr0394) whereas down-regulated proteins included ATP synthase B chain (subunit I) of CF(0) (sll1324) and Soluble inorganic pyrophosphatase (slr1622). Additionally, protein synthesis and degradation, along with nitrogen metabolism and amino acid metabolism proteins, were primarily presented as down-regulated. This shows that the protein synthesis decreased and reorganized due to PG-depletion. The latter groups of proteins regulate general metabolic processes. These outcomes implicate real systems biology, i.e., various processes are cross-regulated and one could easily modify the other. Yet again, this denotes further investigation is required to explore these systems.

Lipid metabolism was also revealed, encompassing Glucosyl-glycerol-phosphate synthase (sll1566, up-regulated) along with Soluble inorganic pyrophosphatase (slr1622, down-regulated) as the two proteins regulated under PG-depletion. Lipid metabolism can be associated with carbohydrate metabolism since the other three glycolipid components of photosynthetic membranes (Monogalactosyldiacylglycerol (MGDG), DGDG, and sulfoquinovosyldiacyglycerol (SQDG)) contain glucose, whereas phosphate metabolism is more closely related to PG (contains phosphate). PG and SQDG are both anionic lipids, and thus could be cross-regulated despite their different head groups (phosphate and sulfur, respectively). Specifically, another strain of cyanobacteria, *Synechococcus* (produced by inactivating the *cdsA* gene, which encodes cytidine 5'-diphosphate-diacylglycerol synthase, a key enzyme in PG synthesis) displayed that PG-depletion increased the SQDG content (3). This could mean that PG can be substituted for by SQDG, therefore maintaining the amount of anionic charged lipids in the cells. However, this

could not compensate the negative effect of PG-depletion on photosynthesis and on other processes (3).

There were additional differentially expressed protein levels discovered in carotenoid biosynthesis, nucleotide synthesis, cell division, ion uptake and transport and others. From earlier physiological studies, PG has been acknowledged as a key component of membrane structure and function of photosynthetic processes (PSI and II). It becomes evident that the effect of PG is multifaceted. We believe that the knowledge of the functional information of protein expression levels provides a strong basis for future follow-up experimental studies. Our results designated numerous processes of this microorganism are either regulated and/or cross, further suggesting that these dynamic changes are insights into cellular adaptations under PG-modulated conditions. Further studies on the isolated membranes from the cells (i.e., thylakoid membranes) could provide additional awareness in relation to cellular adaptation.

CONCLUSION

Cyanobacteria are major contributors of photosynthetic activity. Furthermore, it is estimated that they contribute more than half of the primary production of oxygenic photosynthesis that is essential for sustaining life on earth (32). Although they are not as structurally complex as higher-order species, they are the most genetically diverse microorganisms, occupying a wide variety of habitats from freshwater to hot springs. As a result of their adaptability, their roles in evolutionary, ecological, environmental, and biotechnological aspects have been studied (20, 32-34). Regardless of their importance, cyanobacterial physiology remains to be fully understood. This discovery-driven label-free proteomics approach using peptide peak intensities on a linear ion trap-FTMS hybrid instrument permitted a successful

quantitative survey of key cellular proteins in the *Synechocystis* mutant. Furthermore, these results demonstrated the biological importance and the impact of PG at the protein level in this mutant strain of cyanobacteria. Within this proteome, PG has proven to play a diverse role in functions associated with photosynthesis and respiration, translation, energy metabolism, transport, and binding.

Table 1. Differentially expressed proteins in the *Synechocystis* PCC6803 *pgsA* strain after phosphatidylglycerol-regulation identified by spectral counting using Scaffold software. Minimum protein: 99%, minimum peptide: 95% confidence scores with a minimum of 2 peptides. G-test was accepted at $p < 0.05$ with at least a 2-fold change.

Protein Name[±]	**Accession Number**	**PG-supplemented Average Normalized Spectral Count ± Standard Deviation**	**PG-depleted Average Normalized Spectral Count ± Standard Deviation**	**G-test ($p<0.05$)**[+]	**Fold Change**
Photosystem I reaction center subunit XI[*] (*psaL*)	slr1655	1.02E+01 ± 1.44E+00	2.91E+00 ± 1.04E+00	0.04	3.5
Probable porin	slr1841	2.91E+01 ± 2.25E+00	4.44E+00 ± 7.16E-01	< 0.001	6.6
Probable porin	slr1908	9.48E+00 ± 1.55E+00	1.00E+00 ± 0.00E+00	0.005	9.5
Photosystem I reaction center subunit VII (*psaC*)	ssl0563	1.14E+01 ± 6.99E-01	3.32E+00 ± 7.00E-01	0.03	3.5

[+] Spectral counts from each fraction in PG-supplemented and PG-depleted groups were subjected to a summation-based G statistical test

[±] Gene symbols are in parentheses

[*] Indicated as a membrane protein via crosscheck in CyanoBase

Table 2. Examples of differentially expressed proteins in cyanobacteria after phosphatidylglycerol-regulation identified by spectral counting (Scaffold) and extracted peptide intensity (MS) features generated with Progenesis LC-MS (Nonlinear Dynamics) for comparative evaluation of the methods.

Protein Name[±]	Accession Number	Number of Unique Peptides	Sequence Coverage (%)	PG-Supplemented Average Normalized Abundance ± Standard Deviation	PG-depleted Average Normalized Abundance ± Standard Deviation	ANOVA ($p<0.05$)	Max Fold Change
A) Up-regulated in PG-depleted pgsA mutant cells of *Synechocystis* sp. PCC6803							
Putative arsenical pump-driving ATPase	sll0086	2	5.3	1.81E+03 ± 9.16E+01	8.81E+03 ± 1.09E+03	< 0.001	4.9
6-phosphogluconate dehydrogenase (*gnd*)	sll0329	8	10	2.07E+04 ± 4.78E+03	1.59E+05 ± 1.64E+04	< 0.001	7.7
Flavoprotein (*flv3*)	sll0550	3	5.8	6.38E+02 ± 5.31E+01	2.82E+03 ± 3.31E+02	< 0.001	4.4
Putative sugar-nucleotide epimerase/dehydratease	sll0576	4	7.4	1.22E+04 ± 1.02E+02	3.16E+04 ± 1.19E+03	< 0.001	2.6
Phosphate-binding periplasmic protein precursor, PBP (*pstB*)	sll0680	4	8.6	7.11E+05 ± 9.40E+03	2.48E+06 ± 3.42E+05	0.001	3.5
GDP-fucose synthetase (*wcaG*)	sll1213	3	7.4	2.87E+04 ± 4.60E+03	7.87E+04 ± 4.75E+03	0.004	2.7
Periplasmic protein	sll1306	6	17	2.31E+04 ± 3.77E+03	4.06E+03 ± 7.73E+02	0.003	5.7
ATP synthase alpha chain (*atpA*)	sll1326	14	11	1.20E+05 ± 2.73E+04	7.70E+05 ± 4.51E+04	0.002	6.4
ATP synthase gamma chain (*atpC*)	sll1327	-	-	6.88E+04 ± 2.62E+03	1.35E+05 ± 2.32E+03	< 0.001	2.0
NAD(P)-dependent glyceraldehyde-3-phosphate dehydrogenase (*gap2*)	sll1342	14	26	3.21E+05 ± 6.35E+04	6.54E+05 ± 2.72E+04	0.03	2.0
Ferredoxin-dependent glutamate synthase (*glsF*)	sll1499	8	3	1.21E+05 ± 1.17E+03	5.85E+05 ± 1.40E+05	0.007	4.8
Phosphoribulokinase (*prk*)	sll1525	6	14	4.90E+04 ± 1.27E+04	1.43E+05 ± 2.23E+04	0.03	2.9
Glucosylglycerolphosphate synthase (*ggpS*)	sll1566	2	4	1.68E+03 ± 1.02E+02	5.73E+03 ± 7.57E+02	0.001	3.4

Glutamate decarboxylase (*sll1641*)	sll1641	3	6.6	1.48E+03 ± 1.69E+02	3.88E+03 ± 3.13E+02	0.002	2.6
50S ribosomal protein L2 (*rpl2*)	sll1802	6	14	7.53E+04 ± 1.77E+04	3.61E+05 ± 9.28E+04	0.01	4.8
30S ribosomal protein S11 (*rps11*)	sll1817	5	23	3.53E+05 ± 1.23E+04	7.75E+05 ± 6.57E+04	0.001	2.2
Arginine biosynthesis bifunctional protein (*argJ*)	sll1883	5	11	1.10E+05 ± 1.64E+04	2.87E+05 ± 5.25E+04	0.02	2.6
NADH dehydrogenase subunit 7 (*ndhH*)	slr0261	6	11	2.55E+05 ± 2.35E+04	7.91E+05 ± 1.07E+05	0.003	3.1
Two-component hybrid sensor and regulator (*hik43*)	slr0322	-	-	2.65E+04 ± 1.79E+03	5.27E+04 ± 6.09E+03	0.007	2.0
Phosphoglycerate kinase (*pgk*)	slr0394	9	15	8.80E+04 ± 1.74E+04	4.04E+05 ± 2.34E+04	0.002	4.6
Periplasmic protein *	slr0708	3	7.7	6.74E+03 ± 7.22E+02	3.70E+04 ± 2.11E+03	< 0.001	5.5
Alcohol dehydrogenase [NADP+]	slr0942	2	8.6	2.44E+03 ± 6.76E+02	7.48E+03 ± 4.40E+02	0.03	3.1
30S ribosomal protein S1 (*rps1a*)	slr1356	3	6.4	2.33E+03 ± 3.36E+02	1.94E+04 ± 1.01E+04	0.02	8.3
Elongation factor EF-G (*fus*)	slr1463	8	9.5	4.82E+04 ± 7.38E+03	1.27E+05 ± 3.48E+04	0.04	2.6
Cell division protein FtSH (*ftsH*) *	slr1604	5	6	2.69E+06 ± 3.80E+05	1.09E+07 ± 2.36E+05	< 0.001	4.0
Ferredoxin-NADP oxidoreductase (*petH*)	slr1643	9	16	1.70E+04 ± 1.05E+03	6.15E+04 ± 4.92E+03	< 0.001	3.6
Glucose 6-phosphate dehydrogenase assembly protein (*opcA*)	slr1734	4	8.6	2.75E+04 ± 1.69E+03	1.09E+05 ± 1.49E+04	< 0.001	4.0

± Gene symbols are in parentheses

* Indicated as a membrane protein via crosscheck in CyanoBase

Continued from Table 2

Protein Name[±]	Accession Number	Number of Unique Peptides	Sequence Coverage (%)	PG-Supplemented Average Normalized Abundance ± Standard Deviation	PG-depleted Average Normalized Abundance ± Standard Deviation	ANOVA (p<0.05)	Max Fold Change
B) Down-regulated in PG-depleted pgsA mutant cells of *Synechocystis* sp. PCC6803							
Probable ribonuclease D (*rnd*)	sll0320	5	20	4.69E+04 ± 4.02E+03	1.03E+04 ± 6.71E+02	< 0.001	4.5
Photosystem II manganese-stabilizing polypeptide (*psbO*) [*]	sll0427	6	15	6.88E+04 ± 1.62E+04	1.47E+04 ± 9.30E+02	0.007	4.7
Photosystem I reaction center subunit III precursor [PSI-F], plastocyanin [cyt c553] docking protein (*psaF*) [*]	sll0819	5	42	1.13E+06 ± 1.64E+05	1.69E+05 ± 3.48E+04	0.002	6.7
Carbon dioxide concentrating mechanism protein (*ccmK2)*	sll1028	2	31	7.38E+05 ± 6.12E+04	1.57E+05 ± 1.76E+04	< 0.001	4.7
Carbon dioxide concentrating mechanism protein (*ccmK1*)	sll1029	6	46	2.59E+05 ± 2.18E+04	4.39E+04 ± 3.69E+03	< 0.001	5.9
Geranylgeranyl hydrogenase (*chlP*)	sll1091	6	7.6	1.00E+05 ± 1.28E+04	3.38E+04 ± 2.37E+03	0.003	2.6
30S ribosomal protein S10 (*rps10*)	sll1101	4	29	1.94E+05 ± 2.84E+04	5.50E+04 ± 1.91E+04	0.02	3.5
Photosystem II 12 kDa extrinsic protein (*psbU*) [*]	sll1194	4	25	8.34E+04 ± 2.87E+03	2.55E+04 ± 1.55E+03	< 0.001	3.3
Elongation factor TS (*tsf*)	sll1261	11	39	1.53E+05 ± 3.17E+04	5.58E+04 ± 2.97E+03	0.02	2.7
ATP synthase B chain (subunit I) of CF(0) (*atpF*) [*]	sll1324	4	13	2.52E+04 ± 1.42E+03	1.05E+04 ± 6.15E+02	< 0.001	2.4
Bacterioferritin	sll1341	5	28	2.39E+05 ± 9.10E+03	8.66E+04 ± 1.59E+04	0.007	2.8
AhpC/TSA family protein (*aphC*)	sll1621	8	30	3.22E+05 ± 7.47E+04	1.14E+05 ± 2.65E+04	0.04	2.8
LexA repressor (*lexA*)	sll1626	11	49	1.08E+05 ± 4.17E+03	4.60+0E4 ± 3.59E+03	< 0.001	2.4
DNA binding protein HU	sll1712	6	46	1.37E+05 ± 3.13E+04	1.67E+04 ± 1.66E+03	0.001	8.2
50S ribosomal protein L11 (*rpl11*)	sll1743	3	25	1.92E+05 ± 2.79E+04	4.58E+04 ± 2.74E+03	0.001	4.2
50S ribosomal protein L1 (*rpl1*)	sll1744	8	20	5.06E+04 ± 7.50E+03	1.63E+04 ± 2.33E+02	0.002	3.1
50S ribosomal protein L12 (*rpl12*)	sll1746	5	31	3.48E+05 ± 1.21E+04	5.74E+04 ± 6.27E+03	< 0.001	6.1

RNA polymerase beta prime subunit (*rpoC2*)	sll1789	17	7.3	6.91E+04 ± 1.98E+03	2.71E+04 ± 2.90E+03	0.001	2.6
50S ribosomal protein L6 (*rpl6*)	sll1810	6	11	7.10E+04 ± 6.49E+03	2.76E+04 ± 2.64E+03	0.002	2.6
30S ribosomal protein S13 (*rps13*)	sll1816	2	13	5.72E+04 ± 1.86E+03	2.42E+04 ± 4.39E+02	< 0.001	2.4
50S ribosomal protein L13 (*rpl13*)	sll1821	2	21	7.32E+04 ± 1.41E+04	8.61E+03 ± 1.24E+03	0.001	8.5
D-3-phosphoglycerate dehydrogenase (*serA*)	sll1908	7	8	2.00E+05 ± 8.58E+03	5.76E+04 ± 5.27E+03	< 0.001	3.5
Ribulose bisphosphate carboxylase small subunit (*rbcS*)	slr0012	2	22	5.88E+05 ± 3.09E+03	1.73E+05 ± 6.62E+03	< 0.001	3.4
ATP-dependent Clp protease proteolytic subunit (*clpP3*)	slr0165	4	13	5.74E+04 ± 1.29E+04	2.45E+04 ± 1.52E+03	0.04	2.3
Orotate phosphoribosyltransferase (*umpS*)	slr0185	3	9.1	3.59E+04 ± 8.62E+03	7.33E+03 ± 4.80E+02	0.003	4.9
Thioredoxin (*trxA*)	slr0623	6	48	2.43E+05 ± 1.27E+04	1.53E+04 ± 1.26E+03	< 0.001	15.9
Antioxidant protein (*1-cys prx*)	slr1198	6	18	7.61E+04 ± 2.23E+04	4.75E+03 ± 6.14E+02	0.003	16.0
NADH dehydrogenase subunit 1(*ndhJ*)	slr1281	4	20	5.42E+03 ± 2.87E+02	2.24E+02 ± 3.68E+01	< 0.01	24.2
Iron transport system substrate-binding protein (*futA1*)	slr1295	4	4.7	1.67E+04 ± 2.96E+03	3.28E+03 ± 2.96E+02	0.002	5.1
Sulfate transport system substrate-binding protein (*sbpA*)	slr1452	9	22	1.65E+05 ± 1.19E+04	1.92E+04 ± 1.40E+03	< 0.01	8.6
Soluble inorganic pyrophosphatase (*ppa*)	slr1622	3	20	5.33E+04 ± 1.60E+03	2.36E+04 ± 1.53E+03	< 0.001	2.3
Photosystem II 11 kD protein (*psb27*) *	slr1645	3	15	1.42E+06 ± 8.20E+04	5.32E+05 ± 2.03E+04	< 0.001	2.7
P700 apoprotein subunit lb (*psaB*) *	slr1835	10	7.4	2.95E+05 ± 2.24E+04	4.10E+04 ± 7.56E+03	< 0.001	7.2
Nitrogen regulatory protein P-II (*glnB*)	ssl0707	9	48	2.72E+06 ± 5.09E+04	1.03E+06 ± 7.99E+04	< 0.001	2.6
50S ribosomal protein L29 (*rpl29*)	ssl3436	4	48	9.49E+04 ± 1.43E+04	3.22E+03 ± 7.06E+02	< 0.001	29.5
30S ribosomal protein S16 (*rps16*)	ssr0482	3	48	5.72E+04 ± 1.55E+04	4.27E+03 ± 1.33E+03	0.005	13.4
30S ribosomal protein S18 (*rps18*)	ssr1399	4	37	2.16E+03 ± 1.66E+02	2.73E+02 ± 9.14E+01	0.006	7.9
Cytochrome b559 alpha subunit (*psbE*) *	ssr3451	3	32	8.09E+05 ± 2.58E+04	2.44E+05 ±1.98E+04	< 0.001	3.3

Photosystem I reaction center subunit XI (*psaL*) [*+]	slr1655	5	31	6.32E+05 ± 2.46E+04	8.29E+04 ± 4.66E+04	0.01	7.6
Probable porin [+]	slr1841	15	20	2.64E+06 ± 3.79E+05	2.44E+05 ± 4.50E+04	0.001	10.8
Probable porin [+]	slr1908	9	15	5.72E+05 ± 2.84E+04	5.62E+04 ± 3.97E+03	< 0.001	10.2
Photosystem I reaction center subunit VII (*psaC*) [+]	ssl0563	5	69	5.19E+05 ± 4.66E+04	7.91E+04 ± 1.01E+03	< 0.001	6.6
Photosystem II extrinsic protein (*psbQ*) [*]	sll1638	8	39	2.62E+05 ± 2.41E+04	2.47E+04 ± 3.49E+03	< 0.001	10.6
Periplasmic protein (*tolC*)	slr1270	10	8	6.01E+04 ± 1.04E+04	2.34E+04 ± 2.40E+03	0.01	2.6

[±] Gene symbols are in parentheses

* Indicated as a membrane protein via crosscheck in CyanoBase

[+] Denotes confirmation of protein expression via both spectral counting (Scaffold) and peak integration (Progenesis LC-MS)

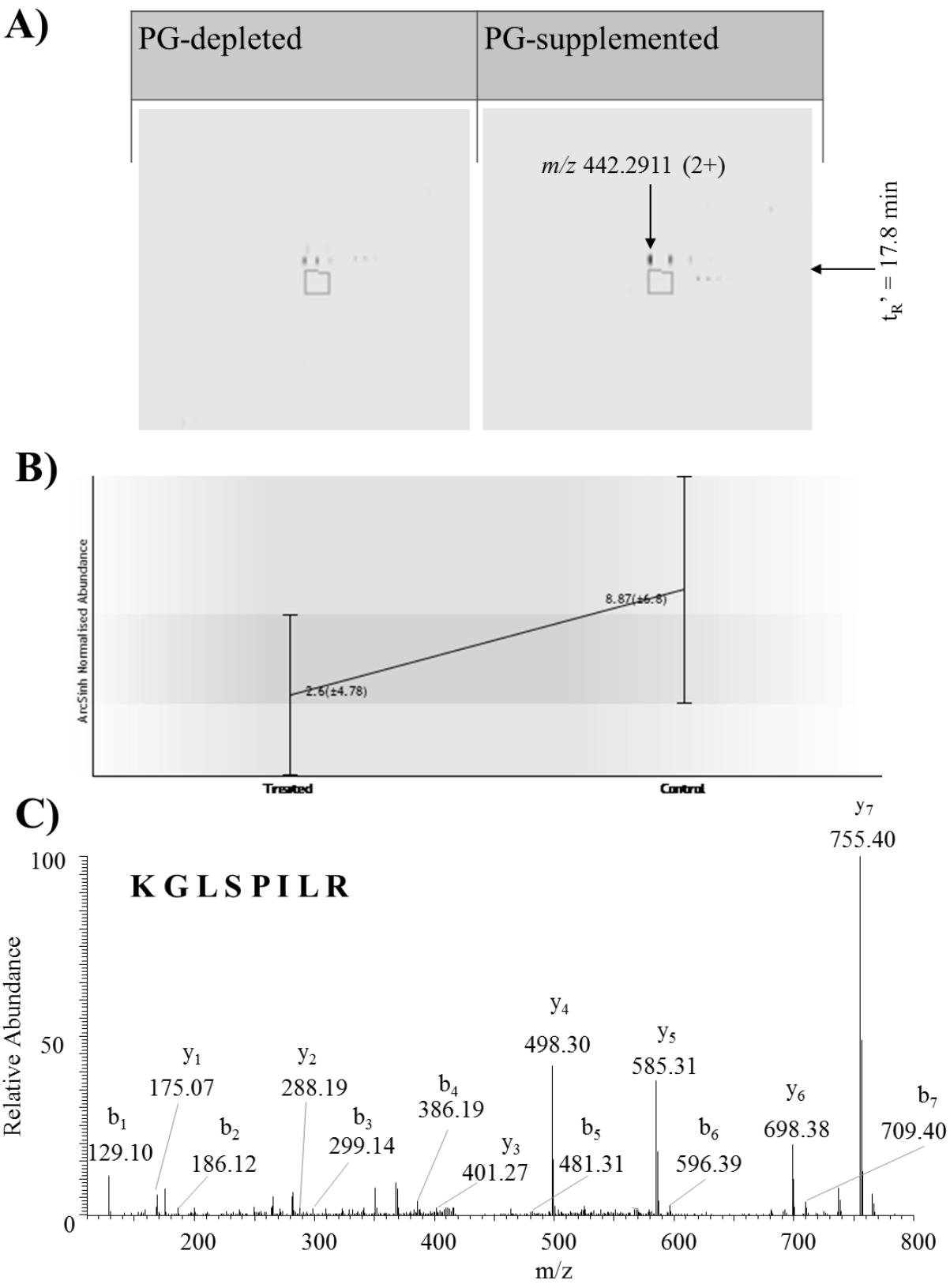

A)
PG-depleted
PG-supplemented
m/z 442.2911 (2+)
t_R' = 17.8 min
B)
ArcSinh Normalised Abundance
8.87(±6.8)
2.5(±4.78)
Treated
Control
C)
K G L S P I L R
Relative Abundance
m/z
y7 755.40
y4 498.30
y5 585.31
y1 175.07
y2 288.19
b4 386.19
y6 698.38
b7 709.40
b1 129.10
b2 186.12
b3 299.14
y3 401.27
b5 481.31
b6 596.39

Figure 11. Example of a down-regulated peptide identified in PG-depleted mutant. Based on extracted peptide intensities obtained from MS scans (FTICR), a feature showing difference between *Synechocystis* PCC 6803 *pgsA*- mutant cells harvested 21 days after depriving from *versus* maintaining exogenous PG was obtained at *m/z* 442.29 (2+) with the aligned retention time of 17.8 min (A). Average normalized abundances shown here tag this protein as being down-regulated in PG-depleted *pgsA* mutant cells of *Synechocystis* sp. PCC6803 (B). One of the sequence-identifying MS/MS (CID) spectra for the sequence KGLSPILR identified by Mascot database search, corresponded to Photosystem I reaction center subunit XI (slr1655) with a score of 51.26 (C).

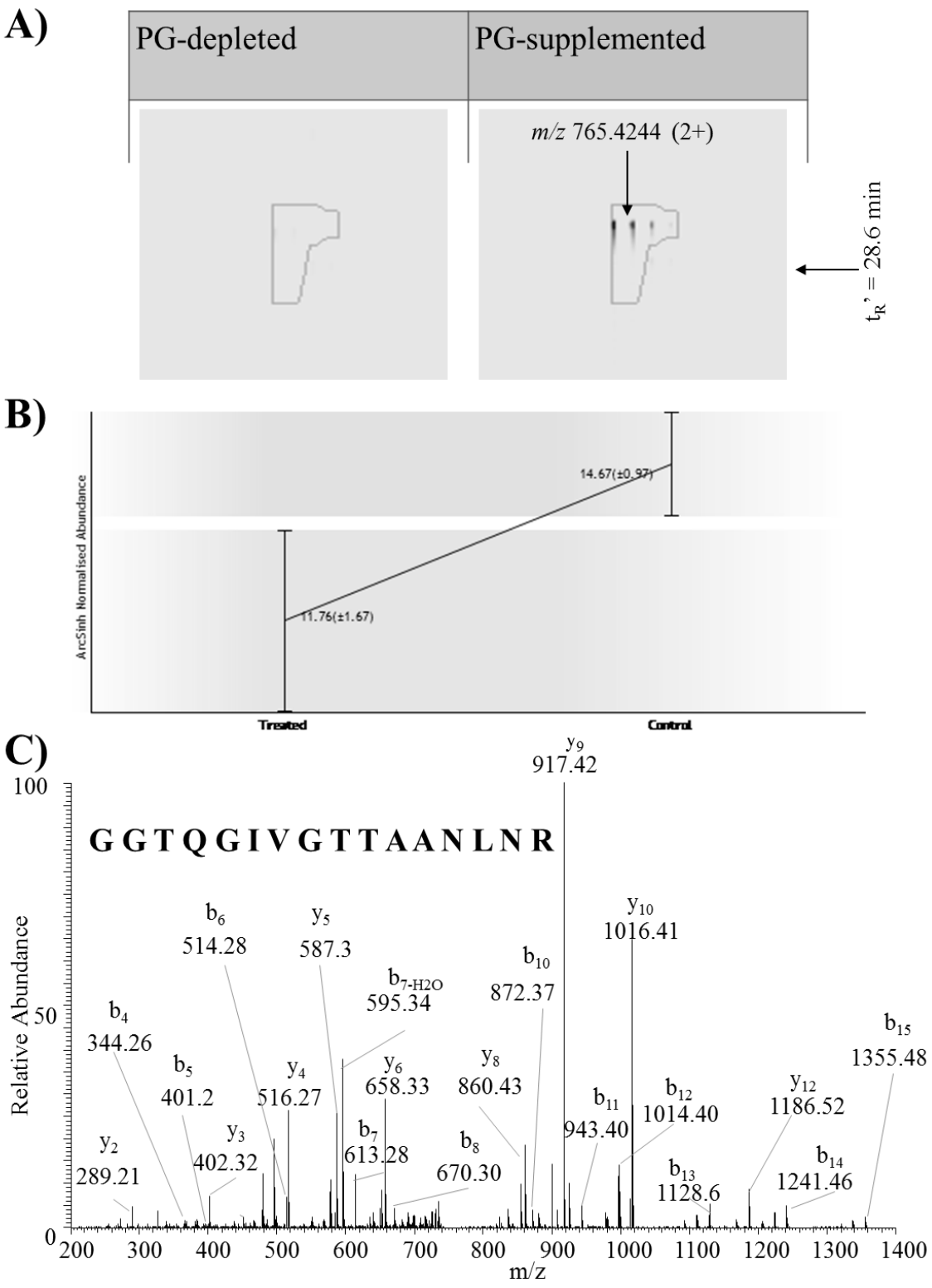
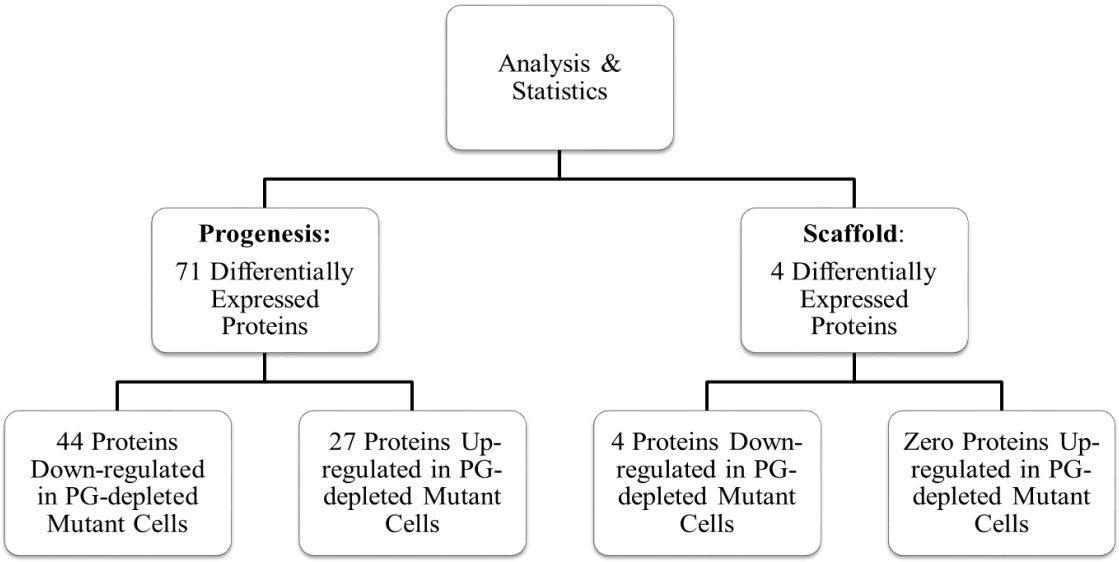

A)
PG-depleted
PG-supplemented
m/z 765.4244 (2+)
t_R' = 28.6 min
B)
ArcSinh Normalised Abundance
14.67(±0.97)
11.76(±1.67)
Treated
Control
C)
G G T Q G I V G T T A A N L N R
Relative Abundance
m/z
y9 917.42
y10 1016.41
b6 514.28
y5 587.3
b7-H2O 595.34
b10 872.37
b4 344.26
b5 401.2
y4 516.27
y6 658.33
y8 860.43
b15 1355.48
b11 943.40
b12 1014.40
y12 1186.52
y2 289.21
y3 402.32
b7 613.28
b8 670.30
b13 1128.6
b14 1241.46

Figure 12. Example of a down-regulated peptide identified in PG-depleted mutant. Based on extracted peptide intensities obtained from MS scans (FTICR), a feature showing difference between *Synechocystis* PCC 6803 *pgsA*- mutant cells harvested 21 days after depriving from *versus* maintaining exogenous PG was obtained at *m/z* 765.42 (2+) with the aligned retention time of 28.6 min (A). Average normalized abundance shown here tag this protein as being down-regulated in PG-depleted *pgsA* mutant cells of *Synechocystis* sp. PCC6803 (B). One of the sequence-identifying MS/MS (CID) spectra for the sequence GGTQGIVGTTAANLNR identified by Mascot database search, corresponded to Probable porin (slr1841) with a score of 80.22.

a)



b)

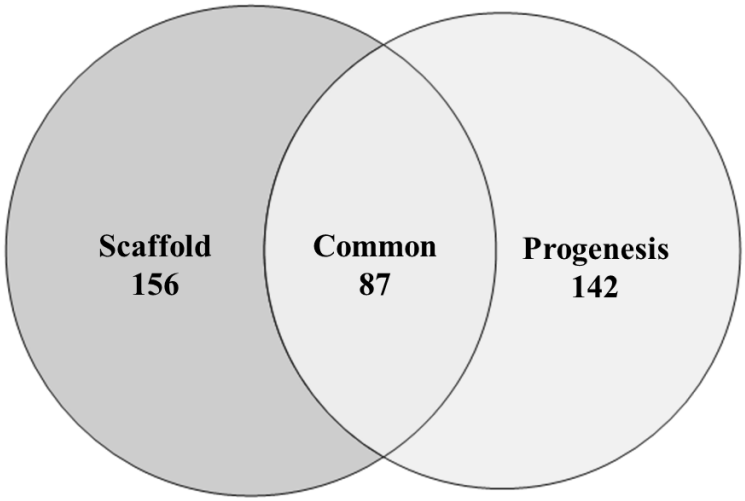


Figure 13. Experimental Scheme and Analysis. a) Flow diagram for identification of differentially expressed proteins using spectral counting by Scaffold (Proteome Software) and extracted peptide intensity features by Progenesis LC-MS (Nonlinear Dynamics). b) Unique and common identified proteins between Scaffold and Progenesis software. Criteria included minimum peptide score >95%, minimum protein score >95% with a minimum of 2 peptides. ANOVA was accepted at $p<0.05$ with a >2-fold change.

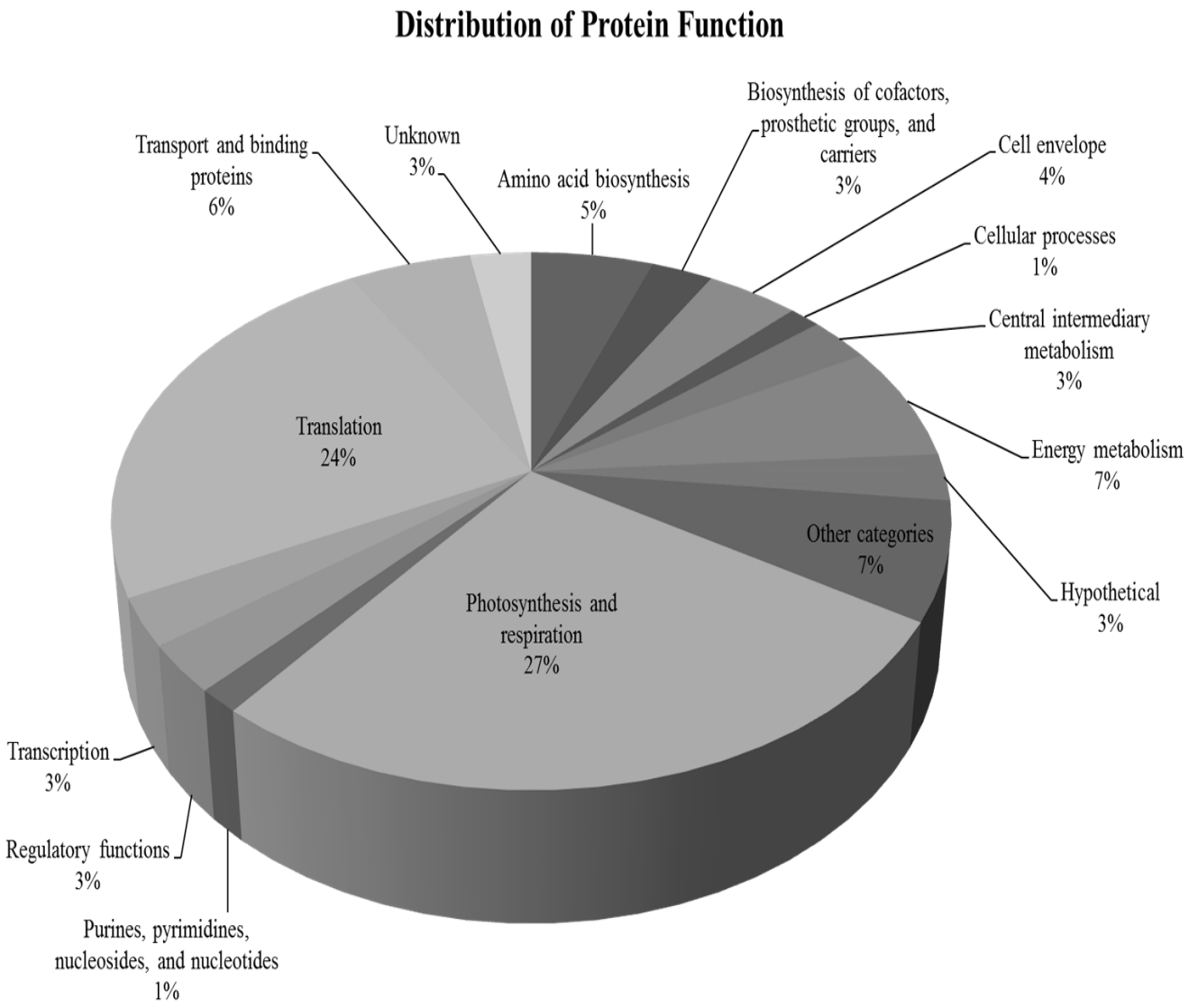


Figure 14. Protein Distribution of PG-regulated proteins *Synechocystis* PCC6803 *pgsA* strain. Utilizing genome information on this cyanobacterial strain (as part of CyanoBase **http://genome.kazusa.or.jp/cyanobase**, on online resource for cyanobacterial genomes).

REFERENCES

1. Gombos Z, Várkonyi Z, Hagio M, Iwaki M, Kovács L, Masamoto K, et al. Phosphatidylglycerol requirement for the function of electron acceptor plastoquinone QB in the photosystem II reaction center. Biochemistry (N Y). 2002;41(11):3796-802.

2. Wada H, Murata N. The essential role of phosphatidylglycerol in photosynthesis. Photosynth Res. 2007 May;92(2):205-15.

3. Bogos B, Ughy B, Domonkos I, Laczko-Dobos H, Komenda J, Abasova L, et al. Phosphatidylglycerol depletion affects photosystem II activity in synechococcus sp. PCC 7942 cells. Photosynth Res. 2010 Jan;103(1):19-30.

4. Domonkos I, Malec P, Sallai A, Kovacs L, Itoh K, Shen G, et al. Phosphatidylglycerol is essential for oligomerization of photosystem I reaction center. Plant Physiol. 2004 Apr;134(4):1471-8.

5. Domonkos I, Laczko-Dobos H, Gombos Z. Lipid-assisted protein-protein interactions that support photosynthetic and other cellular activities. Prog Lipid Res. 2008 Nov;47(6):422-35.

6. Mizusawa N, Wada H. The role of lipids in photosystem II. Biochim Biophys Acta. 2012 Jan;1817(1):194-208.

7. Norihiro Sato HW. Lipid biosynthesis and its regulation in cyanobacteria. In: Hajime Wada NM, editor. Lipids in Photosynthesis: Essential and Regulatory Functions. Dordrecht: Springer Netherlands; 2009. p. 157,157-177.

8. Sakurai I, Hagio M, Gombos Z, Tyystjarvi T, Paakkarinen V, Aro EM, et al. Requirement of phosphatidylglycerol for maintenance of photosynthetic machinery. Plant Physiol. 2003 Nov;133(3):1376-84.

9. Cravatt BF, Simon GM, Yates III JR. The biological impact of mass-spectrometry-based proteomics. Nature. 2007;450(7172):991-1000.

10. Zhu W, Smith JW, Huang CM. Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol. 2010;2010:840518.

11. Yeet Yeang H, Yusof F, Abdullah L. Protein purification for the lowry assay. acid precipitation of proteins in the presence of sodium dodecyl sulfate and other biological detergents. Anal Biochem. 1998;265(2):381-4.

12. Gan CS, Reardon KF, Wright PC. Comparison of protein and peptide prefractionation methods for the shotgun proteomic analysis of *Synechocystis* sp. PCC 6803. Proteomics. 2005 Jun;5(9):2468-78.

13. Guo J, Prokai-Tatrai K, Ngyuen V, Rauniyar N, Ughy B, Prokai L. Protein targets for carbonylation by 4-hydroxy-2-nonenal in rat liver mitochondria. J Proteomics. 2011 Jul 23.

14. Stevens SM,Jr, Rauniyar N, Prokai L. Rapid characterization of covalent modifications to rat brain mitochondrial proteins after ex vivo exposure to 4-hydroxy-2-nonenal by liquid chromatography-tandem mass spectrometry using data-dependent and neutral loss-driven MS3 acquisition. J Mass Spectrom. 2007 Dec;42(12):1599-605.

15. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. 2002 Oct 15;74(20):5383-92.

16. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. 2003 Sep 1;75(17):4646-58.

17. Prokai L, Stevens SM,Jr, Rauniyar N, Nguyen V. Rapid label-free identification of estrogen-induced differential protein expression in vivo from mouse brain and uterine tissue. J Proteome Res. 2009 Aug;8(8):3862-71.

18. Nakao M, Okamoto S, Kohara M, Fujishiro T, Fujisawa T, Sato S, et al. CyanoBase: The cyanobacteria genome database update 2010. Nucleic Acids Res. 2010;38(SUPPL.1):D379-81.

19. Ow SY, Wright PC. Current trends in high throughput proteomics in cyanobacteria. FEBS Lett. 2009;583(11):1744-52.

20. Barrios-Llerena ME, Chong PK, Gan CS, Snijders APL, Reardon KF, Wright PC. Shotgun proteomics of cyanobacteria - applications of experimental and data-mining techniques. Briefings in Functional Genomics and Proteomics. 2006;5(2):121-32.

21. Nagaprashantha LD, Vatsyayan R, Singhal J, Lelsani P, Prokai L, Awasthi S, et al. 2'-hydroxyflavanone inhibits proliferation, tumor vascularization and promotes normal differentiation in VHL-mutant renal cell carcinoma. Carcinogenesis. 2011 Apr;32(4):568-75.

22. Sato N, Suda K, Tsuzuki M. Responsibility of phosphatidylglycerol for biogenesis of the PSI complex. Biochimica et Biophysica Acta - Bioenergetics. 2004;1658(3):235-43.

23. Chitnis VP, Chitnis PR. PsaL subunit is required for the formation of photosystem I trimers in the cyanobacterium *Synechocystis* sp. PCC 6803. FEBS Lett. 1993;336(2):330-4.

24. Xu Q, Hoppe D, Chitnis VP, Odom WR, Guikema JA, Chitnis PR. Mutational analysis of photosystem I polypeptides in the cyanobacterium *Synechocystis* sp. PCC 6803: Targeted inactivation of psaI reveals the function of PsaI in the structural organization of PsaL. J Biol Chem. 1995;270(27):16243-50.

25. Kruip J, Chitnis PR, Lagoutte B, Rögner M, Boekema EJ. Structural organization of the major subunits in cyanobacterial photosystem 1. localization of subunits PsaC, -D, -E, -F, and -J. J Biol Chem. 1997;272(27):17061-9.

26. Thomton LE, Ohkawa H, Roose JL, Kashino Y, Keren N, Pakrasi HB. Homologs of plant PsbP and PsbQ proteins are necessary for regulation of photosystem II activity in the cyanobacterium synechocystis 6803 W inside box sign. Plant Cell. 2004;16(8):2164-75.

27. Ishikawa Y, Schroder WP, Funk C. Functional analysis of the PsbP-like protein (sll1418) in synechocystis sp. PCC 6803. Photosynth Res. 2005 Jun;84(1-3):257-62.

28. Kashino Y, Inoue-Kashino N, Roose JL, Pakrasi HB. Absence of the PsbQ protein results in destabilization of the PsbV protein and decreased oxygen evolution activity in cyanobacterial photosystem II. J Biol Chem. 2006 Jul 28;281(30):20834-41.

29. Roose JL, Kashino Y, Pakrasi HB. The PsbQ protein defines cyanobacterial photosystem II complexes with highest activity and stability. Proc Natl Acad Sci U S A. 2007 Feb 13;104(7):2548-53.

30. Li Y, Kimura Y, Yamauchi Y, Ohno T. Function of extrinsic proteins in stabilization of the photosynthetic oxygen-evolving complex. In: Satyen Saha, editor. Molecular Photochemistry - Various Aspects. InTech; 2012. p. 109,109-124.

31. Kaneko T, Tabata S. Complete genome structure of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. Plant and Cell Physiology. 1997;38(11):1171-6.

32. Wegener KM, Singh AK, Jacobs JM, Elvitigala T, Welsh EA, Keren N, et al. Global proteomics reveal an atypical strategy for carbon/nitrogen assimilation by a cyanobacterium under diverse environmental perturbations. Molecular and Cellular Proteomics. 2010;9(12):2678-89.

33. Sasso S, Pohnert G, Lohr M, Mittag M, Hertweck C. Microalgae in the postgenomic era: A blooming reservoir for new natural products. FEMS Microbiol Rev. 2012;36(4):761-85.

34. Tan X, Wang W, Lu X. Production of photosynthetic biofuels by genetically engineering cyanobacteria. Current Chemical Biology. 2012;6(1):26-31.

CHAPTER III

PROTEOMIC ANALYSIS OF THYLAKOID MEMBRANES OF PHOSPHATIDYLGLYCEROL-DEFICIENT *SYNECHOCYSTIS* SP. PCC6803 MUTANT STRAIN USING 2D-LC-MS/MS

The previous chapter demonstrated that the GeLC-MS approach was valuable for the global-scale proteomics analysis of the *Synechocystis* subproteome regulated by PG. Results indicated that PG-depletion modifies functions of both PSI and PSII proteins. These proteins are associated to the complexes that house photosynthetic reactions. In addition, photosynthetic complexes are located within thylakoid membranes of phototrophic organisms. This initiated the idea that PG-depletion may play a role in membranes functions. Therefore, this section will examine the role of PG-regulated proteins isolated from thylakoid membranes of *Synechocystis* sp. PCC6803 mutant cells using a 2D-LC-MS/MS approach for a better understanding of PG involvement in photosynthesis.

INTRODUCTION

Thylakoid membranes house the photosynthetic machinery in both cyanobacteria and chloroplasts of higher plants. Monogalactosyldiacylglycerol (MGDG),

digalactosyldiacylglycerol (DGDG), and sulfoquinovosyl diacylglycerol (SQDG) make up the majority of lipid components in the membranes of organisms such as cyanobacteria and higher plants (1). Similarly, a less abundant but still significant phospholipid, PG, has been demonstrated to be essential for the structure and function of these photosynthetic membranes (2-4). Numerous completed biochemical and molecular analyses demonstrate that PG is indeed involved in photosynthesis (1, 3, 5).

As noted in the previous chapter, several of the identifications from the PG-depleted samples contained photosynthesis and respiration proteins of the photosystem reaction centers (PSI and PSII). Photosystems are the sights where photosynthesis is carried out. Likewise, numerous studies have shown that PG affects photosynthetic processes and contributes to the efficiency of electron transfer (2, 6, 7). In addition, PG content is significantly higher in PSII complexes than in the actual thylakoid membrane (3). Therefore, depletion in PG may lead to malfunction of the photosystem complexes (e.g., electron transfer).

Mass spectrometry is an essential tool that can be used to understand membrane protein topology, membrane protein-protein interactions, and signaling networks that originate from the membrane surface. To characterize the proteins of the membrane, a survey of the complete proteome is required. Previously, we used a gel-based technique for the analysis of the cyanobacterial proteome on a global-scale. Gel-based methods are a valuable analytical tool to determine global protein expression profiles of complex samples. In addition, coupling liquid-chromatography (LC) to gel-based approaches can increase resolution of extracted peptides or improve the mass accuracy of the mass detection (8). Although the gel-based technique revealed information about the proteome of the *Synechocystis* sp. PCC6803 strain under PG-regulation, it has been acknowledged that this methodology still contains some limitations. Such boundaries

include limited dynamic range and difficulties resolving low abundance proteins such as membrane proteins (8). In addition, with the complexity of these proteomic samples, the separation of proteins prior to mass spectrometric analysis remains crucial (9). Consequently, these types of experiments will have to include additional biochemical enrichment procedures and separation techniques to optimize the isolation of membrane components. Likewise, shotgun proteomics has proven appropriate for these circumstances (10).

Alternative approaches to gel-based methods include shotgun proteomics. This analytical tool first requires digestion of the proteins by proteases to create peptide mixtures, followed by liquid chromatography-mass spectrometry (LC-MS), and sequence database searching. However, just like gel-based methods, there remains a challenge in isolating membrane proteins. Therefore, prior to mass spectrometry, many investigators have incorporated the membrane enrichment strategies reviewed previously (11). The most commonly used shotgun technique is the multidimensional protein identification technology (MudPIT) which combines strong cation exchange (SCX) and reversed-phase (RP) separation. This technique has consequently become routine for identifying proteins.

We have revealed that several photosystem (PSI and PSII) proteins have been affected by PG depletion. Furthermore, photosynthetic processes occur in thylakoid membranes of cyanobacteria. However, only a handful of membrane proteins were identified with our initial gel-based approach. Likewise, in global-scale proteomics studies, membrane proteins are usually underrepresented (12). As a result, further studies of the thylakoid membranes in cyanobacteria are necessary. In this chapter, our shotgun proteomics approach utilized an enrichment technique, to isolate thylakoid membranes, in concert with 2D-LC-MS/MS to identify PG-regulated proteins and their involvement in the photosynthetic complexes.

EXPERIMENTAL STRATEGY

Organism and Culture Conditions

The Δ*pgsA* mutant cells of *Synechocystis* sp. PCC 6803 (13) were grown photoautotrophically as previously described (14). We followed the protocol as described by Bogos *et al.* (15) for growth conditions of PG supplementation (PG+) and PG depletion (PG-). Samples were subsequently analyzed by 2D-LC-MS/MS.

Thylakoid Membrane Preparation

Synechocystis PCC6803 *pgsA* mutant cells were harvested after 21-days with or without exogenous PG, respectively. Cells were disrupted using glass beads in a Bead Beater homogenizer and thylakoid membranes were then isolated by a sucrose gradient centrifugation (16). Membrane samples (total protein 2 mg) were dissolved in lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 2% ASB-14, 1% DM, 20% glycerol, 200 mM KCl, 100 mM Na2HPO4 pH 7.6). An equal volume of digestion solution (8 M urea, 200 mM ammonium bicarbonate, and 20mM methylamine) was added and mixed. The proteins were then reduced with 20 mM tris(hydroxypropyl)phosphine (THP) at 50 °C for 15 min and alkylated with 40mM iodoacetamide in the dark at room temperature for 15 min. Lys-C was added (substrate:enzyme, 100:1) and incubated at 37 °C overnight. The mixture was then diluted 4-fold with 100 mM ammonium bicarbonate. For digestion procedures, trypsin (Trypsin Gold-Mass Spec Grade, Promega) was added [50:1 (w/w) and incubated overnight at 37 °C. Tryptic digests were then desalted on Discovery DSC18 cartridges (Sigma-Aldrich) and lyophilized to dryness in a vacuum concentrator.

Liquid chromatography–tandem mass spectrometry

The trypsin-digested samples were reconstituted in SCX buffer A (10 mM KH_2PO_4, 25% acetonitrile, pH 3) and fractionated by strong cation exchange (SCX) chromatography on a 4.6x200 mm, 5 µm, 300 Å PolySulfoethyl A column. Peptides were eluted with a gradient of buffer B (10 mM KH_2PO_4, 25% acetonitrile, 500 mM KCl) at a flow rate of 1 ml/min (17). The collected 1 ml fractions were pooled to get ten fractions/sample (PG+ and PG-, respectively). Fractions were further lyophilized, then dissolved in 200 µl of 0.1% formic acid and 30% acetonitrile, and transferred to Eppendorf tubes for lyophilization.

SCX fractions were reconstituted in for subsequent online reversed-phase high performance liquid chromatography (RP-HPLC) analysis on a hybrid linear quadrupole ion trap–Fourier transform ion cyclotron resonance (7-T) mass spectrometer (LTQ-FT, Thermo Finnigan, San Jose, CA). The analysis of PG- and PG+ cells were achieved by triplicate injections of both samples (SCX fractions of PG- and PG+, respectively). Data analysis was performed with Xcalibur (version 2.2) and Tune Plus (version 2.2) data acquisition software previously reported (18).

Database Search

MS/MS data generated by data-dependent acquisition via the LTQ-FT were extracted by BioWorks version 3.3 and searched against a composite *Synechocystis* protein sequence database of cyanobacteria (*Synechocystis*) using Mascot version 2.2 (Matrix Science, Boston, MA) search algorithm. Proteins were validated by Scaffold (version 3.0, Proteome Software). Quantifications were performed by extracted peptide intensity (MS) features generated with Progenesis LC-MS version 3.1 (Nonlinear Dynamics, Durham, NC).

Data compilation, relative quantification, and analysis of signaling networks

Data processing was performed using two software tools, Scaffold (version 3.0) and Progenesis LC-MS (version 3.1). Protein identifications were validated by Scaffold software including the following criteria: initial peptide identifications were accepted if they could be established at greater than 95% probability as specified by Peptide Prophet Algorithm (19), and protein identifications were accepted if they could be established at greater than 99% probability and contained at least two identified unique peptides.

Quantifications were performed by extracted peptide intensities (Progenesis LC-MS, Nonlinear Dynamics). Progenesis LC-MS software generates intensity maps from the raw files acquired from the mass spectrometer. Data is then aligned based on LC retention time of each analysis, and a single aggregate run containing all MS data with representing peptide ions is created. Subsequent feature outline maps are generated for detection and quantification of peptide ions from individual analyses. Carrying out peptide quantifications, the software normalizes peak intensities and filters data to retain only ions with positive charges (z) of two and three. Peak lists are then exported to query against the *Synechocystis* protein sequence database using Mascot (see *Database Search* subsection above). By means of analysis of variance (ANOVA, $p<0.05$), proteins were accepted as differentially expressed requiring at least a two-fold change in expression, and passing validation by Peptide Prophet and Protein Prophet using protein identifications from Scaffold.

To consider further biological relevance, identified proteins are submitted to analysis of genome information using CyanoBase (http://genome.kazusa.or.jp/cyanobase). This database is a comprehensive and freely accessible web database for cyanobacteria (20). Annotations including

function, biological processes, and transmembrane data can be retrieved from the database on the GeneView page.

RESULTS AND DISCUSSION

Global-scale Membrane Proteomics Analysis

It has been a challenging task for many researchers to study membrane proteomics (12). Hence the lack of fully characterized membrane proteomes. Membrane proteins are essential in biological systems, and they exist in a variety of biologically significant structural and functional roles. These include but are not limited to transport, cell communication, and signal transduction. Such physiological properties and responsibilities have made membrane proteins a subject of analysis in many research studies. Studying the biology of these critical components can provide an understanding of the molecular mechanisms regulating homeostasis of a system.

In regards to photosynthesis, once the first studies revealed that the chlorophyll molecules of photosynthesis were associated with proteins (21, 22), this initiated investigation of pigment-protein structure and function (23). Studies discovered several of these proteins were associated with photosynthetic electron transport. Cyanobacteria are oxygenic phototrophs containing chlorophyll and other pigments (24). In conjunction with the sequenced genome of *Synechocystis* sp. PCC6803 (25), there has been an increasing interest to use this microorganism in studying photosynthetic properties and related processes (24).

Significant advancement in proteomic techniques has improved the quality and quantity of data generated from *Synechocystis* organisms. Specifically, earlier studies that incorporated global-scale proteomic approaches used wild-type species and high-salt concentrated

environments to identify thylakoid membranes of *Synechocystis* sp. PCC6803 (26-29). These experiments used several different separation techniques as well as mass spectrometry analyses to obtain membrane information. As a result of their success, it remains reasonable to expand the collaboration between proteomics and cyanobacterial research to improve both the technical and biological significance (30) of the cyanobacterial proteome under mutant conditions.

Using the *Synechocystis* PCC6803 *pgsA* mutant, proteins were extracted from isolated thylakoid membranes and subjected to 2D-LC-MS/MS analyses permitting the identification of PG-regulated proteins by label-free approaches. Overall, 843 proteins were identified upon Mascot database searching. Using Scaffold, we established peptide identifications at >95% probability as specified by the Peptide Prophet algorithm and protein identifications were accepted if they could be established at >99% probability with at least 2 identified unique peptides. These criteria permitted 287 confirmed protein identifications in the thylakoid membrane fraction of the *Synechocystis* PCC6803 *pgsA* mutant strain (Figure 1).

Membrane proteins usually become underrepresented in LC-MS proteomics experiments because of their low abundance and poor solubility (12). In this case, to capture the essence of membrane proteins of a complex sample these two analytical challenges must be addressed. In this study, although we used an enrichment method to isolate membrane proteins, we still discovered a large amount of the proteins to be soluble (refer to Table 1). Of the 287 identified proteins, only 61 were actual membrane proteins (not shown).

Utilizing genome information on *Synechocystis* from the online resource CyanoBase we further translated the large dataset to prospective annotations associated with the constituents recognized in the PG-depleted conditions. The 287 identifications, including soluble and membrane proteins, were linked to protein functions in photosynthesis and respiration,

translation, cellular processes, and energy metabolism (Figure 2). In addition, 21% of these connotations were labeled as “hypothetical”. This indicates that the membrane proteome of this mutant is not completely characterized and will require additional studies to elucidate their functional annotations.

On the other hand, the interest of this study was dedicated to thylakoid membranes hence the enrichment procedures. Therefore, the 61 membrane proteins isolated from the cyanobacterial proteome were further subjected to analysis using CyanoBase interpretations. As a result, a majority of the PG-depleted membrane proteins were found to be functional cellular processes, photosynthesis, and respiration (Figure 3). Thirty six percent of this membrane proteome was linked with hypothetical functions. At this point, we have observed there is a need for enrichment since “soluble contaminants” exist in membrane proteins. Furthermore, due to the challenges that exist in isolating only membrane constituents, we still have many unanswered questions about the proteins identified in this proteomics approach.

Regulation of the Protein Distribution of Thylakoid Membrane Proteins

Mass spectrometry has proven beneficial for monitoring global protein expression levels and its undercurrents (23). For evaluation of differential protein expression levels in PG-depleted cyanobacteria mutants, Progenesis LC-MS was used to process the same set of raw data files utilized in Scaffold for protein identifications. Quantitative analyses were acquired from alignment and normalization of peak intensities, and data filtering to retain only peaks with charge state of two to three. Peak lists were exported to query against the cyanobacteria protein database using Mascot. By combining validated protein identifications and quantitation based on extracted peptide intensities, 55 proteins were revealed to be differentially expressed following

analysis of variance (ANOVA, $p<0.05$) requiring at least a 2-fold change in expression and excluding hypothetical or unknown proteins (Table 1).

Main protein complexes involved in photosynthesis are embedded in the membrane, including the photosystems and phycobilisome pigments (31). Photosystems are functional and structural units of protein complexes involved in photosynthesis. In this study, we focused specifically on thylakoid membranes, which are the site of light-dependent reactions of photosynthesis. Within the membrane, we identified six differentially expressed proteins. These included: Phycobilisome rod-core linker polypeptide (sll1471), Photosystem II manganese-stabilizing polypeptide (sll0427), Serine protease (slr0535), Protease FtsH (sll1463), Albino/OxaI homolog of YidC/OxaI family, inner membrane protein (slr1471), and a Putative homolog of the PSII assembly-stabilizing plant factor HCF136 (slr2034). As mentioned earlier, there are challenges in isolating pure membrane proteins from samples. Enrichment techniques as well as fractionation techniques have been required for sample preparation. In this study, although the thylakoid membrane was isolated and fractionated prior to MS analysis, several soluble proteins were identified. However, since this study was focused on the membrane of this organism, the soluble protein IDs have been excluded for comment on their protein expression levels during PG-depletion.

There were six PG-regulated membrane identifications. Amongst the six differentially expressed proteins, four (sll1463, sll0427, sll1471, slr0535) were shown to be down-regulated whereas the other two (slr1471, slr2034) were up-regulated. Using CyanoBase for Gene Ontology (GO) annotations, membrane proteins influenced by PG were found to function in photosynthesis and respiration, translation, and cellular processes (Table 2).

The up-regulated Albino3/OxaI homolog of YidC/OxaI family, inner membrane protein was the only protein categorized as having hypothetical functions (Figure 4). Since this is a hypothetical protein in *Synechocystis*, its role is not known. However, in higher plant studies it has been indicated that *alb3* is vital for thylakoid biosynthesis, namely in folding and complex assembly (32-34). The additional up-regulated protein belongs to Putative homolog of the PSII assembly-stabilizing plant factor HCF136. Consistent with GO annotations, this protein is associated with photosynthesis and respiration in PSII. Similarly, in Arabidopsis, the HCF136 homolog is located in the thylakoid lumen and functions as a PSII stability or assembly factor (35).

The down-regulated proteins slr0535 and sll1463 are largely involved in protein turnover. Slr0535 degrades and slows down the removal and attack of proteins (an example of the identified down-regulated peptide can be found in Figure 5). Furthermore, sll1463 removes dysfunctional or destroyed proteins. It has been observed, that during exponential growth, there is a considerably higher amount of ribosomes, protein synthesis factors, and chaperone/degradation proteins to maintain cell proliferation (36). In this study, since PG was depleted, this caused cells to grow at a slower rate, hence the decrease in the degradation proteins observed under these conditions.

The last two down-regulated proteins are in relation to photosynthesis. First, the sll0427 protein that encodes the *psbO* gene is involved in the oxygen-evolving complex of PSII. From previous studies, we know that oxygen evolution during PG depletion is decreased (37). The protein cluster is one of the most sensitive parts of PSII and can be influenced by many factors. Therefore, it may become sensitive during PG-depletion. Moreover, it may not be as surprising to see that sll0427 is down-regulated because as PG-depletion became longer, we further

observed oxygen-evolution was also decreasing. This might have been because the complex was no longer there.

The last down-regulated protein associated with photosynthesis is not part of the larger phycobilisome complex, but instead is a linker polypeptide. The c*pcG2* is believed to be bound to the membrane and transfers energy to PSI (38-41). Furthermore, the linker polypeptides are sensitive to proteases, meaning they can be degraded quickly if they are not assembled in protein complexes. From this, we have proposed a few theories with respect to the down-regulation of sll1471. From earlier studies, PG-depletion does not allow the PSI trimers to assemble properly, resulting in only PSI monomers (39). We thus hypothesize that if *cpcG2* really does bind to PSI trimers, then the monomers are unable to bind or protect this protein, and therefore they cannot be saved from degradation (i.e. less *cpcG2* in PG-depleted cells). In addition, this could also depend on the explanation that PG-depletion slows down the photosynthetic processes. PSII becomes dysfunctional, there is no oxygen evolution, and PSI activity is decreased. Consequently, there is less energy transfer needed for PSI because the phycobilisomes are light harvesting complexes. In other words, *cpcG2* is down-regulated because the protein is not bound to the membrane or surrounded by other proteins (not protected). In addition, the protein is not needed because PSI no longer requires energy transfer.

Overall, our thylakoid-focused proteomic approach has demonstrated the biological importance and diverse role of this phospholipid in the processes carried out in this pivotal cellular compartment of cyanobacteria. Up-regulated proteins were associated with stability and formation whereas down-regulated identifications were involved in degradation. Moreover, photosystems are the most accessible part of the photosynthetic complex. Therefore, they may be more susceptible to degradation.

CONCLUSION

Despite the recent analytical advancements in conjunction with the dedication to improve the coverage of membrane proteins (11, 12, 42) there remain challenges in addressing membrane proteins in proteomics analyses. Membrane proteomics requires the combination of enrichment and fractionation protocols alongside mass spectrometry in order to enhance membrane proteome coverage. Cyanobacteria are the simplest organisms performing photosynthesis. This process occurs within membranes, which house the lipid constituents that have been shown to play significant roles in these complexes. PG-regulated proteins and their involvement in the photosynthetic machinery could be studied by performing proteomic analyses of thylakoid membranes. Our thylakoid-focused proteomic approach has demonstrated the biological importance and diverse role of this phospholipid in the processes carried out in this pivotal cellular compartment of cyanobacteria including photosynthesis, translation, and cellular processes. It becomes evident that studies involving this genetically modified strain that is incapable of producing PG but capable of growing on PG supplemented media are expected to shed light on hitherto unexplored lipid signaling mechanisms of the process, which could be further investigated in higher plants.

Table 3. Differentially expressed proteins in cyanobacteria after phosphatidylglycerol-regulation identified by MS/MS (Mascot, validation by Scaffold) and extracted peptide intensity (MS) features generated in Progenesis LC-MS.

Protein Name[±]	Accession Number	Number of Unique Peptides	Sequence Coverage (%)	PG-depleted Average Normalized Abundance ± Standard Deviation	PG-supplemented Average Normalized Abundance ± Standard Deviation	ANOVA (P<0.05)	Max Fold Change
A) Up-regulated in PG-depleted cells							
Proline oxidase; (*putA*)	sll1561	13	3.8	1.12E+06 ± 1.49E+04	5.54E+05 ± 3.56E+04	0.008	2.0
Glycyl-tRNA synthetase alpha chain; (*glyQ*)	slr0638	3	9.4	3.05E+04 ± 3.59E+03	1.49E+04 ± 9.63E+02	0.03	2.0
Signal recognition particle protein; (*ffh*)	slr1531	4	5.4	5.69E+04± 4.20E+03	2.66E+04 ± 2.12E+02	0.02	2.1
Putative homolog of PSII assembly-stabilizing plant factor HCF136; (*ycf48*) *	slr2034	4	9.4	1.56E+05 ± 1.41E+04	6.81E+04 ± 6.12E+03	0.02	2.3
Global nitrogen regulator; (*ntcA*)	sll1423	4	17.0	7.04E+04 ± 1.16E+04	2.84E+04 ± 2.48E+02	0.03	2.5
Water-soluble carotenoid protein; (*ocp*)	slr1963	10	14.0	4.09E+05 ± 1.41E+04	1.64E+05 ± 7.32E+03	0.004	2.5
Albino3/OxaI homolog of YidC/OxaI family, inner membrane protein; (*alb3*) *	slr1471	4	6.2	1.99E+05 ± 1.03E+03	7.72E+04 ± 5.45E+03	0.006	2.6
N-acetyl-gamma-glutamyl-phosphate reductase; (*argC*)	sll0080	3	8.0	1.54E+05 ± 7.63E+03	5.65E+04 ± 1.71E+02	0.002	2.7
Phycobilisome core component; (*apcF*)	slr1459	5	33.0	3.72E+05 ± 4.16E+04	1.35E+05 ± 1.65E+04	0.03	2.8
Adenylate kinase; (*adk*)	sll1815	5	17.0	1.30E+05 ± 1.35E+04	4.39E+04 ± 8.57E+03	0.04	3.0

Twitching motility protein PilT; (*pilT1*)	slr0161	9	19.0	3.11E+05 ± 3.98E+04	1.05E+05 ± 2.19E+02	0.01	3.0
Phosphoglycerate kinase; (*pgk*)	slr0394	11	12.0	1.34E+06 ± 1.59E+03	4.49E+05 ± 1.39E+03	< 0.001	3.0
GTP-binding protein	slr1974	7	8.0	6.14E+04 ± 4.23E+03	2.02E+04 ± 1.24E+03	0.007	3.0
Phycobilisome small rod linker polypeptide; (*cpcD*)	ssl3093	4	45.0	3.32E+05 ± 2.14E+04	1.09E+05 ± 2.63E+03	0.004	3.0
Allophycocyanin-B; (*apcD*)	sll0928	5	16.0	6.67E+05 ± 8.35E+04	1.93E+05 ± 4.27E+03	0.01	3.5
Cytochrome b6-f complex iron-sulfur subunit (Rieske iron sulfur protein); (*petC*)	sll1316	4	17.0	1.70E+05 ± 2.91E+04	4.69E+04 ± 6.00E+03	0.03	3.6
Phosphoribosylformimino-5-amino-phophorylbosil-4-imidazolecarboxamideisomerase; (*hisA*)	slr0652	8	11.0	1.59E+05 ± 2.70E+04	4.06E+04 ± 2.54E+03	0.02	3.9
ATP synthase beta subunit; (*atpD*)	slr1329	17	25.0	6.03E+06 ± 2.30E+05	1.41E+06 ± 1.91E+05	0.009	4.3
Phycobilisome rod-core linker polypeptide; (*cpcG1*)	slr2051	11	33.0	3.34E+06 ± 4.27E+05	7.76E+05 ± 3.75E+04	0.009	4.3
50S ribosomal protein L12; (*rplL*)	sll1746	4	19.0	1.36E+05 ± 3.43E+03	3.15E+04 ± 1.43E+03	0.001	4.3
Adenylosuccinate lyase; (*purB*)	sll0421	3	10.0	3.45E+04 ± 3.74E+03	6.91E+03 ± 1.58E+03	0.02	5.0
Phycobilisome small core linker polypeptide; (*apcC*)	ssr3383	4	51.0	1.53E+06 ± 2.97E+05	2.90E+05 ± 2.93E+04	0.02	5.3
Allophycocyanin alpha subunit; (*apcA*)	slr2067	6	23.0	1.39E+06 ± 6.92E+04	2.47E+05 ± 7.57E+03	0.001	5.6

Stationary-phase survival protein SurE homolog	sll1459	2	11.0	3.60E+03 ± 8.28E+02	6.08E+02 ± 9.41E+00	0.02	5.9
Cytochrome b6-f complex subunit; (*petP*)	ssr2998	3	29.0	9.85E+04 ± 5.52E+03	1.59E+04 ± 9.72E+02	0.002	6.2
Photosystem II 13 kDa protein homolog; (*psb28-2*)	slr1739	4	28.0	1.48E+05 ± 6.54E+03	2.32E+04 ± 1.89E+03	0.002	6.4
Similar to D-3-phosphoglycerate dehydrogenase	slr2123	4	11.0	1.47E+05 ± 1.16E+04	2.18E+04 ± 1.64E+03	0.003	6.7
Photosystem II component; (*psb29*)	sll1414	2	8.7	1.13E+04 ± 7.52E+02	1.28E+03 ± 7.39E+00	0.001	8.8
Fructose-1,6-/sedoheptulose-1,7-bisphosphatase; (*fbpl*)	slr2094	7	12.0	4.09E+05 ± 1.13E+05	2.75E+04 ± 6.36E+03	0.02	14.9
Carbamoyl-phosphate synthase, pyrimidine-specific, large chain; (*carB*)	sll0370	2	2.1	1.46E+04 ± 9.94E+01	8.83E+02 ± 1.19E+02	0.002	16.5
Ornithine carbamoyltransferase; (*argF*)	sll0902	5	6.2	1.98E+04 ± 3.75E+03	1.13E+03 ± 1.69E+02	0.007	17.5
Succinate-semialdehyde dehydrogenase (NADP+); (*gabD*)	slr0370	2	5.5	8.24E+02 ± 3.76E+02	3.43E+01 ± 1.13E+01	0.04	24.0

[±] Gene symbols are in parentheses

* Indicated as a membrane protein via crosscheck in CyanoBase

Continued from Table 1

Protein Name	Accession Number	Number of Unique Peptides	Sequence Coverage (%)	SCX 4 Average Normalized Abundance ± Standard Deviation	SCX 3 Average Normalized Abundance ± Standard Deviation	ANOVA P < 0.05	Max Fold Change
B) Down-regulated in PG-depleted cells							
upp, uraP; uracil phosphoribosyltransferase; K00761 uracil phosphoribosyltransferase [EC:2.4.2.9]	syn:sll1035	2	12.0	1.10E+03 ± 4.65E+02	2.50E+04 ± 6.55E+02	0.02	22.7
sbpA; ABC-type sulfate transport system *substrate-binding* protein; K02048 sulfate transport system substrate-binding protein	syn:slr1452	7	18.0	2.18E+04 ± 1.60E+03	2.19E+05 ± 4.79E+03	0.001	10.1
ftsH; cell division protein *FtsH*; K03798 cell division protease FtsH [EC:3.4.24.-]	syn:sll1463	14	8.8	5.43E+05 ± 9.38E+04	4.40E+06 ± 3.58E+04	0.007	8.1
pilT2; twitching mobility *protein*; K02669 twitching motility protein PilT	syn:sll1533	7	12.0	8.40E+04 ± 9.11E+03	5.86E+05 ± 5.43E+03	0.003	7.0
pyrH, smbA; uridine monophosphate kinase; K09903 uridylate kinase [EC:2.7.4.22]	syn:sll0144	2	10.0	3.12E+03 ± 1.26E+03	2.05E+04 ± 1.98E+02	0.04	6.6
futA1, sufA; ABC-type iron transport *system* substrate-binding protein; K02012 iron(III) transport system substrate-binding	syn:slr1295	9	15.0	2.62E+05 ± 6.74E+04	1.54E+06 ± 3.79E+05	0.04	5.9

protein							
ndhI; NADH dehydrogenase I chain *I*; K05580 NADH dehydrogenase I subunit I [EC:1.6.5.3]	syn:sll0520	5	13.0	1.21E+05 ± 1.75E+03	6.91E+05 ± 5.96E+04	0.003	5.7
epsB; exopolysaccharide export *protein*	syn:sll0923	2	4.0	2.70E+03 ± 8.12E+02	1.46E+04 ± 1.09E+02	0.03	5.4
probable extracellular solute-binding protein	syn:slr1962	4	7.8	3.57E+03 ± 2.00E+02	1.55E+04 ± 1.62E+02	0.001	4.4
lrtA; light repressed protein A *homolog*; K05808 putative sigma-54 modulation protein	syn:sll0947	5	20.0	3.46E+05 ± 7.38E+04	1.26E+06 ± 5.32E+04	0.03	3.6
taxAY3, pilL-C, cheA, hik43; *a* part of two-component hybrid histidine kinase, homologous to chemotaxis protein CheA; K02487 type IV pili sensor histidine kinase/response regulator	syn:slr0322	3	2.7	1.18E+04 ± 8.84E+02	4.06E+04 ± 3.00E+03	0.007	3.5
clpC; ATP-dependent Clp protease ATPase *subunit*; K03696 ATP-dependent Clp protease ATP-binding subunit ClpC	syn:sll0020	19	9.0	8.46E+05 ± 8.18E+04	2.79E+06 ± 4.93E+04	0.007	3.3
rplA, rpl1; 50S ribosomal protein L1; K02863 large subunit ribosomal protein L1	syn:sll1744	5	17.0	7.02E+03 ± 1.03E+03	2.09E+04 ± 1.62E+03	0.02	3.0
pyrD; dihydroorotate *oxidase*; K00226 dihydroorotate oxidase [EC:1.3.3.1]	syn:slr1418	3	9.2	1.06E+04 ± 1.43E+03	3.04E+04 ± 2.78E+03	0.02	2.9

ftsZ; cell division protein *FtsZ*; K03531 cell division protein FtsZ	syn:sll1633	3	9.1	1.46E+05 ± 6.14E+03	3.99E+05 ± 9.24E+03	0.002	2.7
ccmK2; carbon dioxide concentrating mechanism protein *CcmK*, putative carboxysome assembly protein; K08696 carbon dioxide concentrating mechanism protein CcmK	syn:sll1028	4	26.0	1.03E+05 ± 2.82E+03	2.74E+05 ± 1.48E+03	< 0.001	2.7
serine proteinase; K01362	syn:slr0535	2	4.1	7.23E+03 ± 8.72E+02	1.88E+04 ± 1.42E+03	0.02	2.6
tufA, tuf; elongation factor EF-Tu; K02358 elongation factor EF-Tu [EC:3.6.5.3]	syn:sll1099	16	26.0	2.57E+06 ± 4.53E+04	6.57E+06 ± 7.11E+05	0.01	2.6
cpcG2; phycobilisome rod-core linker *polypeptide* Lrc, C-phycocyanin associated; K02290 phycobilisome rod-core linker protein	syn:sll1471	6	9.6	3.39E+04 ± 3.76E+03	8.64E+04 ± 6.22E+03	0.02	2.5
ccmM; carbon dioxide concentrating mechanism protein CcmM, putative carboxysome structural *protein*; K08698 carbon dioxide concentrating mechanism protein CcmM	syn:sll1031	13	9.9	6.63E+05 ± 2.99E+04	1.65E+06 ± 2.28E+05	0.03	2.5
sucC; succinyl-CoA synthetase beta *chain*; K01903 succinyl-CoA synthetase beta chain [EC:6.2.1.5]	syn:sll1023	2	6.2	3.14E+03 ± 3.34E+02	7.64E+03 ± 7.73E+02	0.03	2.4
cupA; involved in	syn:sll1734	10	10.0	6.06E+05 ± 1.85E+04	1.44E+06 ± 6.63E+04	0.004	2.4

inducible high-affinity CO2 uptake. psbO; photosystem II manganese-stabilizing *polypeptide*; K02716 photosystem II PsbO protein	syn:sll0427	7	15.0	3.73E+05 ± 5.05E+04	8.33E+05 ± 1.94E+04	0.03	2.2

[±] Gene symbols are in parentheses * Indicated as a membrane protein via crosscheck in CyanoBase

Table 4. Gene Ontology (GO) annotations associated with differentially expressed thylakoid membrane proteins in PG-depleted *Synechocystis* PCC6803.

Protein Name	Accession Number	Function	Biological Process	Cellular Component	Molecular Function	Transmembrane Region
Albino3/OxaI homolog of YidC/OxaI family, inner membrane protein; (*alb3*)	slr1471	Hypothetical: Hypothetical	Protein insertion into membrane (GO:0051205)	Integral to membrane (GO:0016021)	-	5 transmembrane helices
Putative homolog of plant HCF136, which is essential for stability or assembly of photosystem II; (*ycf48*)	slr2034	Photosynthesis and respiration: Photosystem II	-	-	-	1 transmembrane helix
Cell division protein FtsH; (*ftsH*)	sll1463	Cellular processes: Cell division	Protein catabolic process (GO:0030163); proteolysis (GO:0006508)	Integral to membrane (GO:0016021); membrane (GO:0016020)	ATP binding (GO:0005524); metalloendo-peptidase activity (GO:0004222); metallo-peptidase activity (GO:0008237); nucleoside-triphosphatase activity (GO:0017111); nuceleotide binding (GO:0000166); zinc ion binding (GO:0008270)	2 transmembrane helices

Photosystem II manganese-stabilizing polypeptide; (*psbO*)	sll0427	Photosynthesis and respiration: Photosystem II	Photosynthesis (0015979); photosystem II stabilization (GO:0042549)	Extrinsic to membrane (GO0019898); oxygen evolving complex (GO:0009654)	Calcium ion binding (GO:0005509)	1 transmembrane helix
Phycobilisome rod-core linker polypeptide; (*cpcG2*)	sll1471	Photosynthesis and respiration: Phycobilisome	Photosynthesis (0015979)	Phycobilisome (GO0030089)	-	1 transmembrane helix
Serine protease	slr0535	Translation: Degradation of proteins, peptides, and glycopeptides	Protein amino acid phosphorylation (GO:0006468); Proteolysis (GO:0006508)	-	ATP binding (GO:0005524); protein kinase activity (GO:0004672); subtilase activity (GO:0004289)	6 transmembrane helices

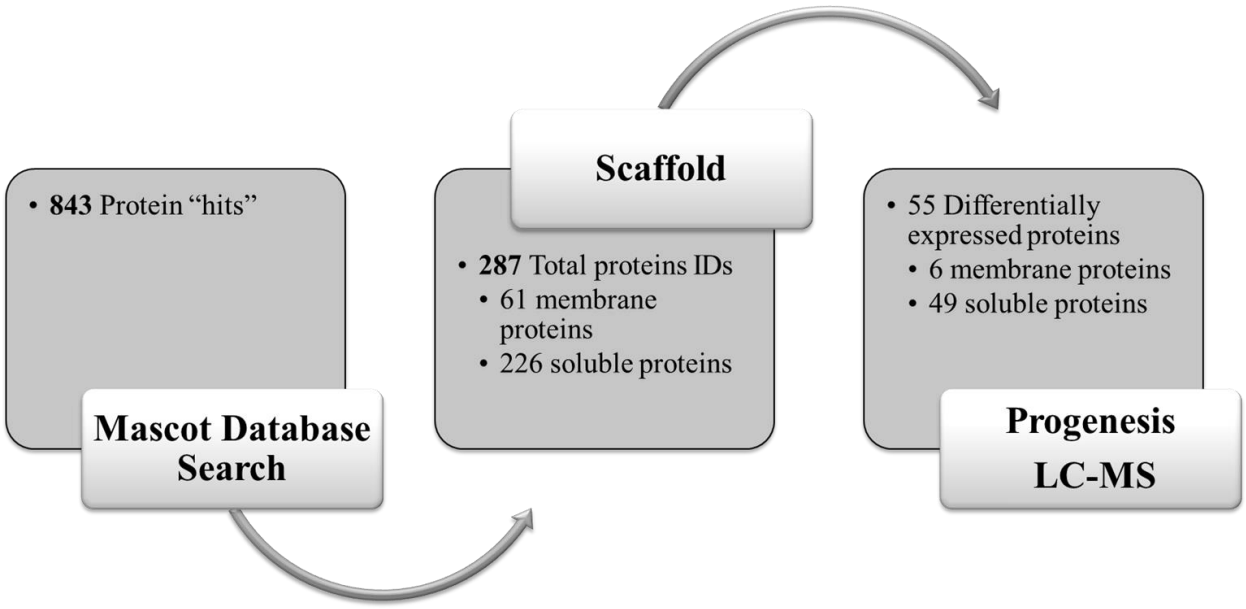


Figure 15. Experimental scheme for data processing using proteomics software. Mascot database searching identified 843 protein hits. Using Scaffold software, peptide identifications were established at >95% probability and protein identifications were established at >99% probability with at least 2 identified unique peptide, thus confirming 287 protein identifications in the thylakoid membrane fractions of the *Synechocystis* mutant strain. For quantitative analyses, Progenesis LC-MS was used to process the same set of raw data, revealing 68 proteins to be differentially expressed following analysis of variance (ANOVA, $p<0.05$) and requiring at least a 2-fold change in expression.

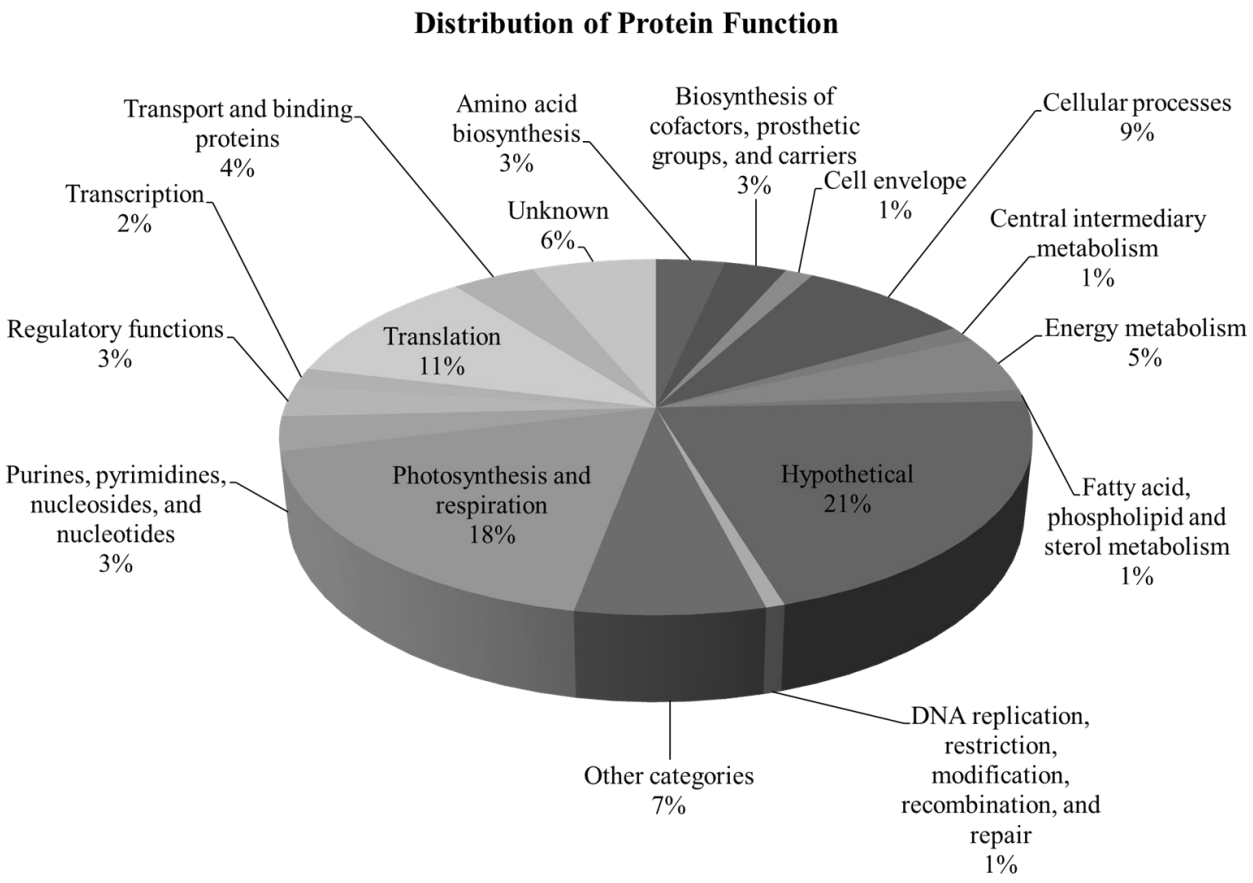


Figure 16. Distribution of the protein functions for the proteins (total) identified in PG-depleted *Synechocystis* sp. PCC6803. Scaffold analysis identified 287 total proteins including 61 membrane and 226 soluble proteins. Considering the biological relevance of these IDs, analysis of their associated annotations was retrieved using CyanoBase (an online resource for cyanobacterial genomes). A majority of the proteins identified in PG-regulated *Synechocystis* mutants were associated with photosynthesis and respiration, translation, cellular processes, and other processes.

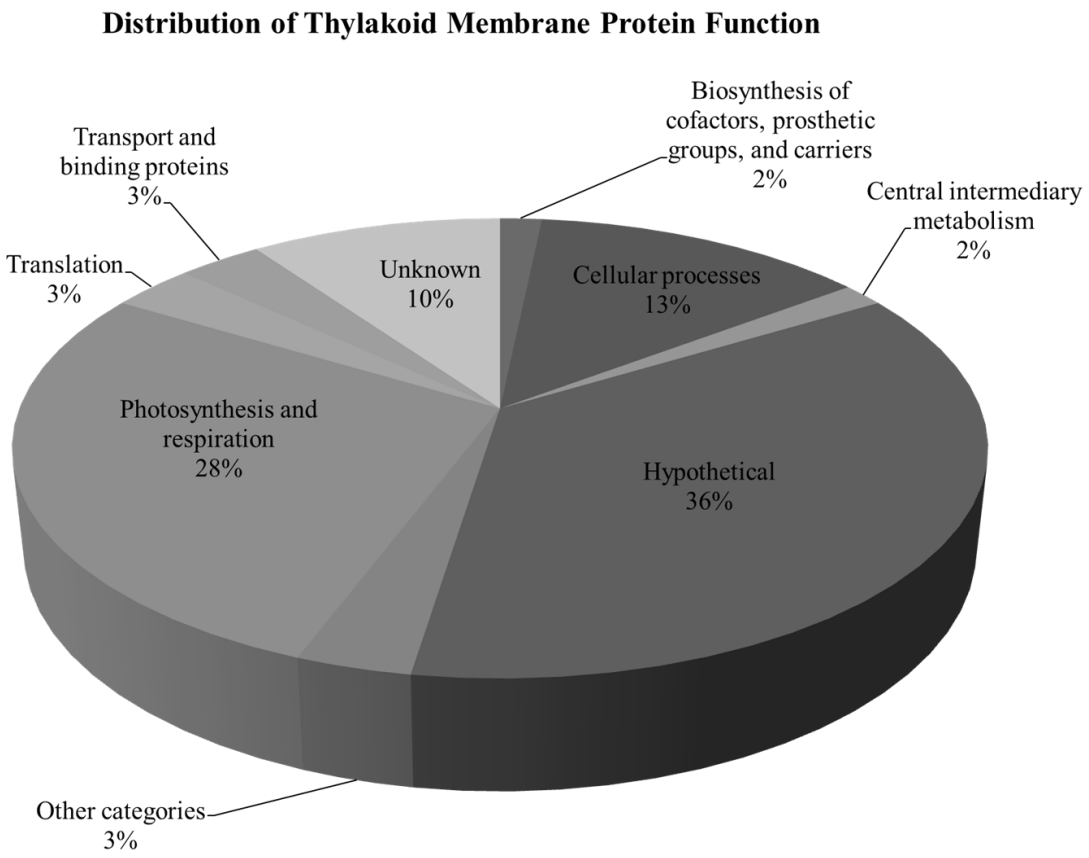


Figure 17. Distribution of the protein functions associated with thylakoid membrane proteins identified in PG-depleted *Synechocystis* sp. PCC6803. Isolation of thylakoid membranes from cyanobacteria was required in order to elucidate further details regarding the cellular processes being modulated in these mutant species. CyanoBase interpretation of annotations associated with the 61 membrane proteins identified in Scaffold was considered valuable in regards to the membrane constituents. Several annotations were categorized in photosynthesis and respiration, cellular processes, and hypothetical processes.

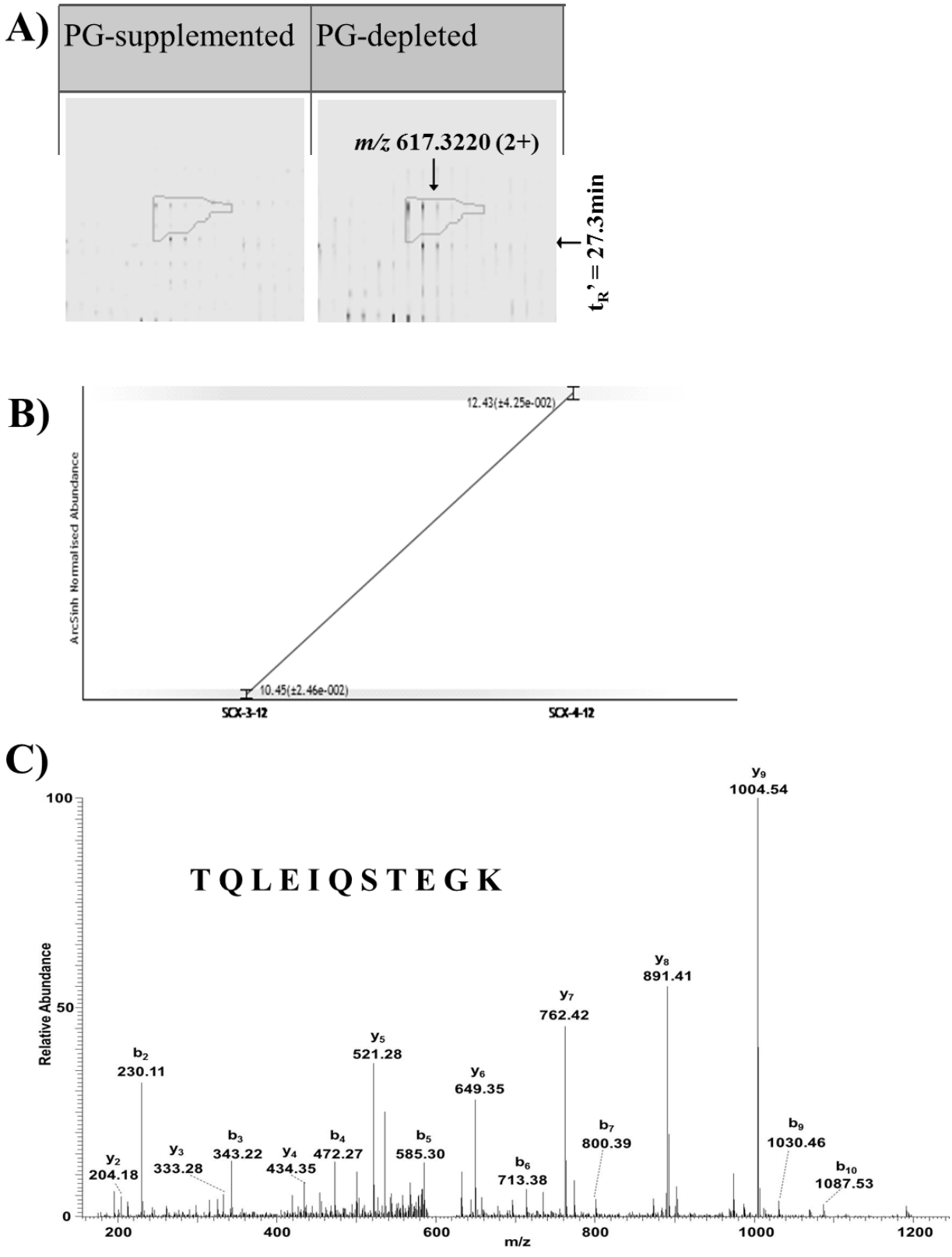

A)
PG-supplemented
PG-depleted
m/z 617.3220 (2+)
t_R' = 27.3min
B)
12.43(±4.25e-002)
ArcSinh Normalised Abundance
10.45(±2.46e-002)
SCX-3-12
SCX-4-12
C)
T Q L E I Q S T E G K
Relative Abundance
m/z
y9 1004.54
y8 891.41
y7 762.42
y5 521.28
y6 649.35
b2 230.11
b7 800.39
b9 1030.46
b10 1087.53
b3 343.22
b4 472.27
b5 585.30
b6 713.38
y2 204.18
y3 333.28
y4 434.35

Figure 18. Example of an up-regulated peptide identified in PG-depleted mutant. Based on extracted peptide intensities obtained from MS scans (FTICR), a feature showing difference between *Synechocystis* PCC 6803 *pgsA*- mutant cells harvested at 21 days with or without exogenous PG was obtained at *m/z* 617.3230 (2+) with an aligned retention time of 27.3 min (A). Average normalized abundances shown here tag this protein as being down regulated upon PG-depletion (B). MS/MS spectra for the sequence identified by Mascot database search TQLEIQSTEGK, was associated with *Albino3/OxaI homolog of YidC/OxaI family, inner membrane protein, alb3* (slr1471) with a score of 63.62 (C).

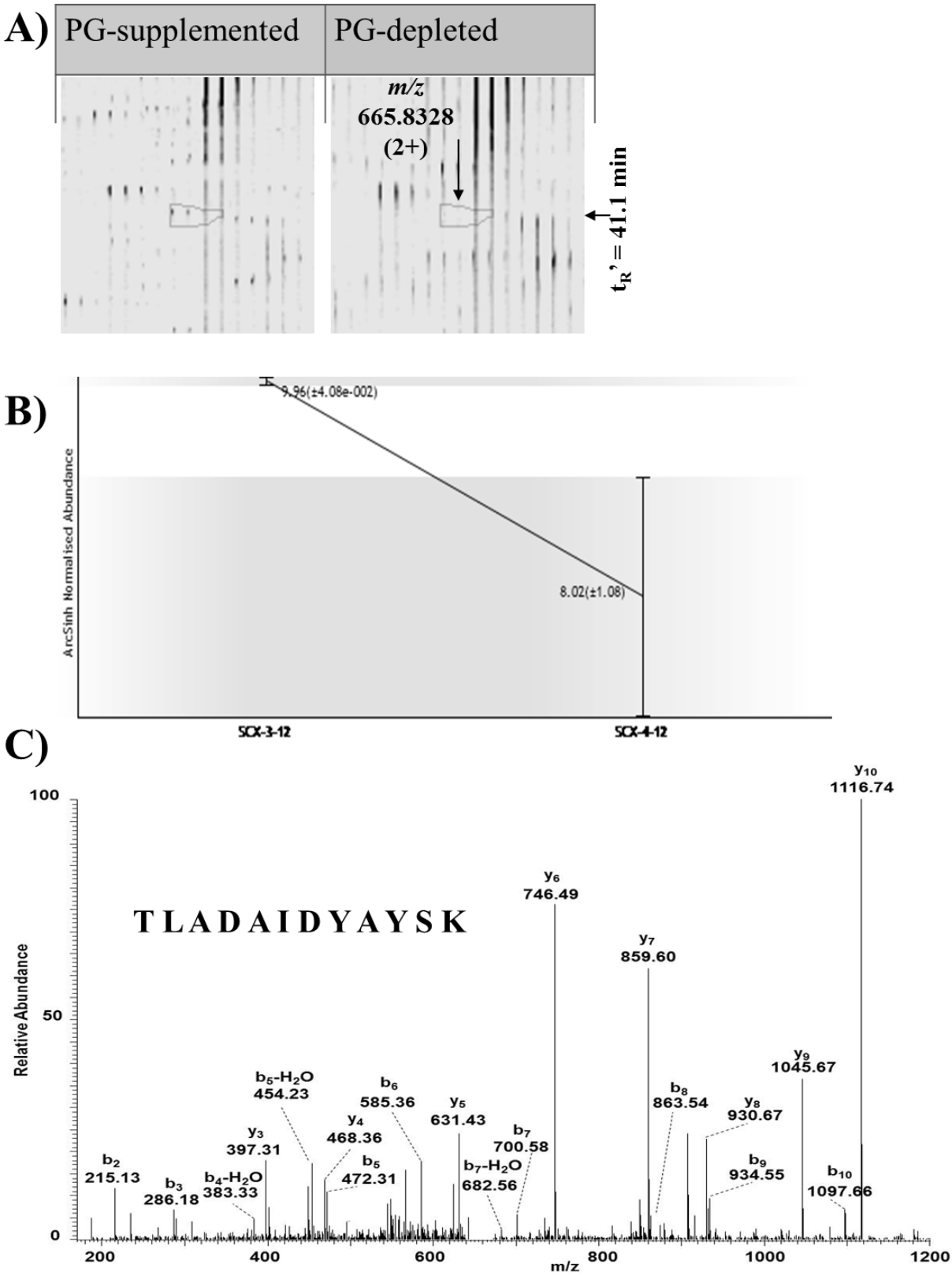

A)
PG-supplemented
PG-depleted
m/z
665.8328
(2+)
t_R' = 41.1 min
B)
9.96(±4.08e-002)
8.02(±1.08)
ArcSinh Normalised Abundance
SCX-3-12
SCX-4-12
C)
T L A D A I D Y A Y S K
Relative Abundance
m/z
y10 1116.74
y6 746.49
y7 859.60
y9 1045.67
b8 863.54
y8 930.67
b9 934.55
b10 1097.66
b5-H2O 454.23
b6 585.36
y5 631.43
y4 468.36
y3 397.31
b7 700.58
b5 472.31
b7-H2O 682.56
b2 215.13
b3 286.18
b4-H2O 383.33

Figure 19. Example of a down-regulated peptide identified in PG-depleted mutant. Based on extracted peptide intensities obtained from MS scans (FTICR), a feature showing difference between *Synechocystis* PCC 6803 *pgsA*- mutant cells harvested at 21 days with or without exogenous PG was obtained at *m/z* 665.8328 (2+) with an aligned retention time of 41.1 min (A). Average normalized abundances shown here tag this protein as being down regulated upon PG-depletion (B). MS/MS spectra for the sequence identified by Mascot database search TLADAIDYAYSK, was associated with *Serine protease* (slr0535) with a score of 76.17 (C).

REFERENCES

1. Sato N, Hagio M, Wada H, Tsuzuki M. Requirement of phosphatidylglycerol for photosynthetic function in thylakoid membranes. Proc Natl Acad Sci U S A. 2000;97(19):10655-60.

2. Gombos Z, Várkonyi Z, Hagio M, Iwaki M, Kovács L, Masamoto K, et al. Phosphatidylglycerol requirement for the function of electron acceptor plastoquinone QB in the photosystem II reaction center. Biochemistry (N Y). 2002;41(11):3796-802.

3. Sakurai I, Shen JR, Leng J, Ohashi S, Kobayashi M, Wada H. Lipids in oxygen-evolving photosystem II complexes of cyanobacteria and higher plants. J Biochem. 2006 Aug;140(2):201-9.

4. Wada H, Murata N. The essential role of phosphatidylglycerol in photosynthesis. Photosynth Res. 2007 May;92(2):205-15.

5. Wu F, Yang Z, Kuang T. Impaired photosynthesis in phosphatidylglycerol-deficient mutant of cyanobacterium anabaena sp. PCC7120 with a disrupted gene encoding a putative phosphatidylglycerophosphatase. Plant Physiol. 2006;141(4):1274-83.

6. Sakurai I, Mizusawa N, Wada H, Sato N. Digalactosyldiacylglycerol is required for stabilization of the oxygen-evolving complex in photosystem II. Plant Physiol. 2007 Dec;145(4):1361-70.

7. Domonkos I, Laczko-Dobos H, Gombos Z. Lipid-assisted protein-protein interactions that support photosynthetic and other cellular activities. Prog Lipid Res. 2008 Nov;47(6):422-35.

8. Wu CC, Yates III JR. The application of mass spectrometry to membrane proteomics. Nat Biotechnol. 2003;21(3):262-7.

9. Gilmore JM, Washburn MP. Advances in shotgun proteomics and the analysis of membrane proteomes. Journal of Proteomics. 2010;73(11):2078-91.

10. Savas JN, Stein BD, Wu CC, Yates JR. Mass spectrometry accelerates membrane protein analysis. Trends Biochem Sci. 2011;36(7):388-96.

11. Speers AE, Wu CC. Proteomics of integral membrane proteins - theory and application. Chem Rev. 2007;107(8):3687-714.

12. Vuckovic D, Dagley LF, Purcell AW, Emili A. Membrane proteomics by high performance liquid chromatography-tandem mass spectrometry: Analytical approaches and challenges. Proteomics. 2013;13(3-4):404-23.

13. Hagio M, Gombos Z, Varkonyi Z, Masamoto K, Sato N, Tsuzuki M, et al. Direct evidence for requirement of phosphatidylglycerol in photosystem II of photosynthesis. Plant Physiol. 2000 Oct;124(2):795-804.

14. Domonkos I, Malec P, Sallai A, Kovacs L, Itoh K, Shen G, et al. Phosphatidylglycerol is essential for oligomerization of photosystem I reaction center. Plant Physiol. 2004 Apr;134(4):1471-8.

15. Bogos B, Ughy B, Domonkos I, Laczko-Dobos H, Komenda J, Abasova L, et al. Phosphatidylglycerol depletion affects photosystem II activity in synechococcus sp. PCC 7942 cells. Photosynth Res. 2010 Jan;103(1):19-30.

16. Domonkos I, Malec P, Laczko-Dobos H, Sozer O, Klodawska K, Wada H, et al. Phosphatidylglycerol depletion induces an increase in myxoxanthophyll biosynthetic activity in *Synechocystis* PCC6803 cells. Plant Cell Physiol. 2009 Feb;50(2):374-82.

17. Gan CS, Reardon KF, Wright PC. Comparison of protein and peptide prefractionation methods for the shotgun proteomic analysis of *Synechocystis* sp. PCC 6803. Proteomics. 2005 Jun;5(9):2468-78.

18. Stevens SM,Jr, Rauniyar N, Prokai L. Rapid characterization of covalent modifications to rat brain mitochondrial proteins after ex vivo exposure to 4-hydroxy-2-nonenal by liquid chromatography-tandem mass spectrometry using data-dependent and neutral loss-driven MS3 acquisition. J Mass Spectrom. 2007 Dec;42(12):1599-605.

19. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. 2002 Oct 15;74(20):5383-92.

20. Nakao M, Okamoto S, Kohara M, Fujishiro T, Fujisawa T, Sato S, et al. CyanoBase: The cyanobacteria genome database update 2010. Nucleic Acids Res. 2010;38(SUPPL.1):D379-81.

21. Ogawa T, Obata F, Shibata K. Two pigment proteins in spinach chloroplasts. BBA - Biophysics Including Photosynthesis. 1966;112(2):223-34.

22. Thornber JP, Smith CA, Bailey JL. Partial characterization of two chlorophyll-protein complexes isolated from spinach-beet chloroplasts. Biochem J. 1966;100:14.

23. Whitelegge JP. Thylakoid membrane proteomics. Photosynthesis Res. 2003;78(3 SPEC. ISS.):265-77.

24. Burja AM, Dhamwichukorn S, Wright PC. Cyanobacterial postgenomic research and systems biology. Trends Biotechnol. 2003;21(11):504-11.

25. Kaneko T, Tabata S. Complete genome structure of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. Plant and Cell Physiology. 1997;38(11):1171-6.

26. Agarwal R, Matros A, Melzer M, Mock H-, Sainis JK. Heterogeneity in thylakoid membrane proteome of *Synechocystis* 6803. Journal of Proteomics. 2010;73(5):976-91.

27. Agarwal R, Maralihalli G, Sudarsan V, Choudhury SD, Vatsa RK, Pal H, et al. Differential distribution of pigment-protein complexes in the thylakoid membranes of *Synechocystis* 6803. J Bioenerg Biomembr. 2012;44(4):399-409.

28. Srivastava R, Pisareva T, Norling B. Proteomic studies of the thylakoid membrane of *Synechocystis* sp. PCC 6803. Proteomics. 2005;5(18):4905-16.

29. Pandhal J, Noirel J, Wright PC, Biggs CA. A systems biology approach to investigate the response of *Synechocystis* sp. PCC6803 to a high salt environment. Saline Systems. 2009;5(1).

30. Ow SY, Wright PC. Current trends in high throughput proteomics in cyanobacteria. FEBS Lett. 2009;583(11):1744-52.

31. Barrios-Llerena ME, Chong PK, Gan CS, Snijders APL, Reardon KF, Wright PC. Shotgun proteomics of cyanobacteria - applications of experimental and data-mining techniques. Briefings in Functional Genomics and Proteomics. 2006;5(2):121-32.

32. Moore M, Harrison MS, Peterson EC, Henry R. Chloroplast Oxa1p homolog albino3 is required for post-translational integration of the light harvesting chlorophyll-binding protein into thylakoid membranes. J Biol Chem. 2000;275(3):1529-32.

33. Ossenbühl F, Göhre V, Meurer J, Krieger-Liszkay A, Rochaix J-, Eichacker LA. Efficient assembly of photosystem II in chlamydomonas reinhardtii requires Alb3.1p, a homolog of arabidopsis ALBINO3. Plant Cell. 2004;16(7):1790-800.

34. Bellafiore S, Ferris P, Naver H, Göhre V, Rochaix J-. Loss of Albino3 leads to the specific depletion of the light-harvesting system. Plant Cell. 2002;14(9):2303-14.

35. Covshoff S, Majeran W, Liu P, Kolkman JM, Van Wijk KJ, Brutnell TP. Deregulation of maize C4 photosynthetic development in a mesophyll cell-defective mutant. Plant Physiol. 2008;146(4):1469-81.

36. Mrázek J, Bhaya D, Grossman AR, Karlin S. Highly expressed and alien genes of the *Synechocystis* genome. Nucleic Acids Res. 2001;29(7):1590-601.

37. Itoh S, Kozuki T, Nishida K, Fukushima Y, Yamakawa H, Domonkos I, et al. Two functional sites of phosphatidylglycerol for regulation of reaction of plastoquinone Q B in photosystem II. Biochimica et Biophysica Acta - Bioenergetics. 2012;1817(2):287-97.

38. Kondo K, Xiao XG, Katayama M, Ikeuchi M. Distinct roles of CpcG1 and CpcG2 in phycobilisome assembly in the cyanobacterium *Synechocystis* sp. PCC 6803. Photosynthesis Res. 2005;84(1-3):269-73.

39. Kondo K, Ochiai Y, Katayama M, Ikeuchi M. The membrane-associated CpcG2-phycobilisome in *Synechocystis*: A new photosystem I antenna. Plant Physiol. 2007;144(2):1200-10.

40. Kondo K, Mullineaux CW, Ikeuchi M. Distinct roles of CpcG1-phycobilisome and CpcG2-phycobilisome in state transitions in a cyanobacterium *Synechocystis* sp. PCC 6803. Photosynthesis Res. 2009;99(3):217-25.

41. Deng G, Liu F, Liu X, Zhao J. Significant energy transfer from CpcG2-phycobilisomes to photosystem i in the cyanobacterium synechococcus sp. PCC 7002 in the absence of ApcD-dependent state transitions. FEBS Lett. 2012;586(16):2342-5.

42. Rabilloud T. Membrane proteins and proteomics: Love is possible, but so difficult. Electrophoresis. 2009;30(SUPPL. 1):S174-80.

CHAPTER IV

DETECTION OF *IN VIVO* CARBONYLATION BY 4-HYDROXY-2-NONENAL DURING EARLY DEVELOPMENT OF ZEBRAFISH EMBRYOS THROUGH DISCOVERY-DRIVEN PROTEOMICS: SEEKING POTENTIAL IMPLICATIONS FOR HUMAN DISEASES USING 2D LC-MS/MS

The previous sections of this dissertation focused on the broader understanding of lipid influences in a simple microorganism (i.e. cyanobacteria) utilizing a discovery-driven proteomics approach. We successfully surveyed the cyanobacterial proteome during PG-regulation and revealed cellular processes impacted by this constituent. The efficiency of this strategy was evaluated and tested in this microorganism and deemed applicable to the investigation of insights into the underlying biochemistry of lipid electrophile stress in a higher-order zebrafish model.

INTRODUCTION

Protein Carbonylation, an Indicator of Oxidative Stress

Under normal physiological conditions, there is a constant generation of oxidants (i.e. reactive oxygen species, ROS). Disturbance in this normal redox state, known as oxidative stress, can cause production of peroxides and free radicals, which can ultimately cause damage

to components of the cell such as proteins, lipids, and DNA (1-7). ROS, e.g., H_2O_2, have high propensity to produce free radicals under specific conditions. Free radicals have unpaired electron(s) and, in order to reach a more stable energy level, they easily pick up electron(s) or hydrogen atoms from another molecule. In turn, this converts those molecules into free radicals, thus, setting up a chain reaction. Once this process begins, it can cascade, resulting in tissue and/or cell damage and ultimately lead to cell death (8). In addition to this cascade of events, several inflammatory and age-associated disease states, including macular degeneration and muscular dystrophy, cancer, diabetes, sepsis, cardiovascular disease and neurological disorders, have been linked to this oxidative stress damage (3, 5-7, 9-13).

As high levels of oxidative stress generate ROS, these molecules further react with membrane phospholipids to generate lipid electrophiles (14). Polyunsaturated fatty acyl chains found in membranes and lipoproteins have been shown to be particularly susceptible to these electrophiles (5, 7, 10, 11). This process, known as lipid peroxidation (5, 7, 10) generates highly reactive aldehyde end products. These include keto aldehydes, 2-alkenals, and 4-hydroxy-2-alkenals (15-18). The α,β-unsaturated hydroxyalkenal, 4-hydroxy-2-nonenal (HNE) is a byproduct of the peroxidation of membrane-derived ω-6 polyunsaturated fatty acids such as arachidonic and linoleic acids (5, 7, 10, 19). HNE is also a highly reactive α,β-unsaturated electrophilic aldehyde that contains cytotoxic properties, and easily attacks the nucleophile function groups on the side-chain of His, Lys, and Cys residues in proteins, preferentially via Michael addition (3, 5, 6, 10-12, 20-23). The result of Michael addition is protein carbonylation (Fig. 1). Protein carbonylation can also be formed via metal-catalyzed oxidized protein carbonyl products containing glutamic and aminoadipic semialdehydes (24-26). Additionally, the accumulation of HNE-modified proteins has been associated with various pathological processes

(5, 27). With an emphasis directed at HNE as a mediator of oxidative stress, extensive research over the last few decades has revealed that oxidative-induced protein modifications play a role in structural alterations and/or protein dysfunction (5, 14). Considering that reactive carbonyls can overwhelm the antioxidant defense systems and thus damage several components of cells, we can use HNE as a potential biomarker of oxidative stress damage (28, 29).

Application and Methodologies to Uncover Carbonylated Proteins

Identification of carbonylated proteins within cells and tissue has been studied using several techniques of derivitization of proteins with carbonyl groups. Some of these techniques include 2,4-dinitrophenylhydrazine (DNPH), biotin hydrazide, and affinity chromatography to select derivatized groups (5, 7, 19, 30). Regardless of these established techniques and the growing list of proteins modified by lipid electrophiles like HNE, it remains a challenge to identify specific sites of amino acid modification and their relation to physiological responses (7). Effective techniques to map these site-specific posttranslational modifications (PTMs) include a combination of protein chemistry, bioaffinity strategies, and liquid chromatography-tandem mass spectrometry (LC-MS/MS) techniques (31-34).

It remains to be fully understood which specific biomarkers play a role in the initiation and/or progression of associated diseases (7, 19). Discovering details about specific sites of modification enables follow-up studies on the functional consequences of HNE-induced carbonyl stress involving these protein targets. *In vitro* studies have suggested that various types of this experimentally induced oxidative stress may mimic the effect of age and correlate with the extent of protein carbonylation (35). Several pathological disorders attributed to protein carbonylation such as cellular aging, inflammation, and neurodegeneration have been defined through the study

of individual protein targets (36). For instance, our lab has previously shown that incorporated enrichment methods prior to mass spectrometric analysis significantly enhanced the quality of the MS/MS spectra to detect peptide carbonyls from HNE-treated rat liver mitochondria (37).

As proteomics is becoming more popular, the technology for it is also improving (38). Mass spectrometry-based proteomics has proven to be a beneficial tool for identifying as well as quantifying proteins (39-41). In addition, incorporating bioinformatic analyses will further decipher the large datasets and can illustrate the biological processes occurring in the chosen model system (42). Investigating the behavior and relationships of certain components in a particular biological system integrates discovery- and hypothesis-driven science. This allows for the complexities of the system to become exposed, and thus, permits future-directed studies.

Discovery-driven Methodology to Elucidate Carbonyl Stress on a Systems Level

In this study, we chose zebrafish as our *in vivo* model. This vertebrate model has progressively been recognized as a valuable model organism for drug discovery, target validation, and toxicological studies (43-48). Drug delivery into zebrafish larvae is straightforward. Once the embryo has hatched and the chorion (outer membrane surrounding embryo) has been removed, they are able to absorb small molecules through their skin and gills (~3-4 days post fertilization) (47). Since this vertebrate model permits a simple rout to deliver drugs, application of agents that can cause oxidative stress to specific tissues or organs will allow experimentation to assess the *in vivo* damage. In addition, statistically significant numbers of zebrafish can be utilized for each treatment condition, and, because of the size of zebrafish embryos, small amounts of drugs can be used for each dosing (47). These innate qualities of the

zebrafish embryos represents an attractive vertebrate model for studies targeted at understanding toxic and signaling mechanisms of molecules such as HNE.

As HNE continues to reveal itself as a significant modulator of cell function and signaling, there remains a considerable amount to learn about HNE and its cellular effect at the protein level. The scope of this chapter begins with the identification of protein carbonylation of protein targets via the reactive aldehyde HNE during the early development of zebrafish embryos. Verification of HNE-induced protein carbonylation distributed throughout the zebrafish embryos was followed by detection of specific HNE-modified protein targets. The application of enrichment strategies coupled with shotgun-based mass spectrometric analysis revealed site-specific HNE modifications among the crucial protein targets recognized in this zebrafish model. Concurrently, we observed that HNE was playing a role in regulating the proteome. By defining the biological consequences using HNE-induced carbonyl stress, we highlighted potential signaling pathways and processes, as well as isolated the precise role and/or contribution of the posttranslational modifications (PTMs) targeted in the zebrafish proteome. The goal of this systems approach was to further evaluate and establish how HNE interrelates with various associated dysfunctions and pathophysiological disorders.

EXPERIMENTAL STRATEGY

Chemicals

HNE was purchased from Cayman Chemical Company (Ann Arbor, MI). HPLC grade solvents were all obtained from Fisher Scientific (Atlanta, GA). Sequencing grade trypsin was

from Applied Biosystems (Foster City, CA). All other chemicals were acquired through Sigma-Aldrich (St. Louis, MO), unless otherwise stated.

Origin and Maintenance of Zebrafish

Animals were treated in compliance with the guidelines set by the National Institutes of Health and the Institutional Animal Care and Use Committee at the University of North Texas Health Science Center. Wild-type, reproductively mature zebrafish were ordered from Aquatica Biotech (Sun City Center, FL). Animals were maintained in natural mating trios (one male and two females) in our aquatic habitat under standard laboratory conditions (49) at a constant temperature of approximately 28.5 °C in a 14:10-hour light:dark cycle, and fed two-three times daily with commercial tropical fish flake food supplemented by brine shrimp (Petsolutions, Beavercreek, OH). Directly after spawning, fish eggs were collected and cleaned with fish water. Embryos were maintained and treated in embryo medium approximately 28.5 °C. At 24 hours post fertilization (hpf), zebrafish were sorted for viability. Since embryos received nourishment from an attached yolk ball, no additional maintenance was required (49).

Exposure of HNE to Zebrafish Embryos *in vivo*

At 72 hpf, 100 embryos were treated in 2 mL embryo medium containing 62.5 µL HNE at approximately 28.5 °C. Control embryos were treated analogously without HNE exposure. Control and treated groups were then processed simultaneously by washing with embryo medium and resuspending them in lysis buffer (8M urea, 1% CHAPS, cOmplete Mini Protease Inhibitor Cocktail Tablet in 0.1M PBS). The lysate was then incubated on ice for 30 min, followed by 10 freeze/thaw cycles to ensure sufficient lysis of the cells. Samples were

centrifuged at 1,000g for 5 min at 4 °C. The supernatant was collected, and protein concentration was measured via the BCA Protein Assay Kit (Piercenet, Rockford, IL). The lysates were stored at -80 °C until the appropriate experiments were performed (in-solution and in-gel protocols).

Immunohistochemistry

The following procedure was modified from the AbCam Zebrafish Mount Immunohistochemistry protocol (50). Briefly, at 72 hpf, zebrafish embryos were treated with 0.2 mM HNE in embryo medium for 45 min. Following treatment, both control and treated embryos were fixed with 4% formaldehyde for one hour. Subsequently, formaldehyde was washed off four times with PBS 1% Triton. Embryos were permeabilized with ice cold acetone/PBS for 8 min and washed four times with PBS 1% Triton. Permeabilized embryos were incubated twice for one hour at room temperature in blocking solution (PBS 1% Triton + 10% lamb serum) then transferred to 2mL Eppendorf tubes and incubated for 2 days on gentle rotation at 4 °C in the primary antibody (1:30 Anti-HNE, R&D Systems Catalog# MAB3249). After incubation, embryos were washed three times for one hour at room temperature in blocking solution and then incubated for two days on gentle rotation at 4 °C, in the dark, in the fluorescent secondary antibody (1:200 Alexa Fluor 594 Anti-Mouse, Invitrogen Grand Island, NY). Lastly, embryos were washed three times for 10 min in PBS 1% Triton and mounted on slides for microscope analysis.

Gel Electrophoresis and Western Blot

HNE-treated embryonic protein samples as well as controls (20 µg per well), were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using

12% polyacrylamide gels. Precision Plus Protein™ marker (Bio-Rad, Hercules, CA) was also run along with samples. One gel was processed for Western blot analysis, while the other was visualized with Coomassie Blue G-250 (Bio-Rad, Hercules, CA) per standard protocol as previously described (37) and utilized for in-gel digestion experiments. Primary antibodies to probe the blots with were mouse anti-HNE antibody (R&D Systems, Minneapolis, MN), anti-alpha tubulin monoclonal mouse antibody (Thermo Scientific, Rockford, IL), monoclonal anti-actin mouse antibody (Sigma, St. Louis, MO), or anti-myosin heavy chain antibody (Millipore, Temecula, CA) diluted to 1:500, 1:750, 1:2400, and 1:3000, respectively. The primary antibody was removed, then the blots were incubated with horseradish peroxidase-conjugated anti-mouse secondary antibody (Invitrogen, Carlsbad, CA), at a dilution of 1:3000. Blots were washed and subsequently visualized with SuperSignal West Pico chemiluminescent substrate (Thermo Scientific, Rockford, IL), and developed by LabWorks Image Acquisition and Analysis software version 4.6 (Ultraviolet Products, Cambridge, UK).

In-gel Digestion

Thirteen gel bands of interest were cut horizontally with a clean scalpel, sliced into small pieces, and placed into Eppendorf tubes followed by washing to remove Coomassie stain. Following the previously reported protocol (37), gel pieces were dried and proteins were reduced, alkylated, and digested with trypsin overnight. Tryptic peptides were then extracted from the gels and dried by Speedvac (Thermo Scientific Savant, San Jose, CA) before LC–MS/MS analysis.

Solid-phase Hydrazine Enrichment of HNE-modified Peptides

In-solution HNE-treated and control digests were reconstituted in reaction buffer, following procedures previously reported, to capture HNE-modified peptides (51). Enriched HNE-modified peptides were reconstituted in 20 µL of 5% (v/v) acetonitrile in water containing 0.1% (v/v) acetic acid, and aliquots of 5 µL were used for LC-MS/MS analyses.

Data-dependent LC-MS/MS and NL-driven MS^3 Acquisitions

The SPH eluate fractions as well as in-gel digests were analyzed using a hybrid linear ion trap−Fourier transform ion-cyclotron resonance (7-Tesla) mass spectrometer (LTQ-FT, Thermo Finnigan, San Jose, CA) equipped with a nano-electrospray ionization (ESI) source and operated with Xcalibur (version 2.2) and Tune Plus (version 2.2) data acquisition software. Online reversed-phase high performance liquid chromatography (RP-HPLC) was performed with an Eksigent nano-LC-2D (Eksigent, Dublin, CA) system as previously reported (52).

In addition to MS/MS acquisitions, a NL-driven MS^3 data-dependent acquisition method was employed to facilitate the detection and identification of HNE-modified peptides (52) from in-solution and SPH enriched samples. Isolation and subsequent fragmentation of ions exhibiting a m/z 78, 52 or 39 difference (representing NL of HNE from C2, C3, or C4 precursor ions respectively) from the precursor ion occurred if the NL fragment ions passed specified selection criteria (i.e., among the three most intense ions in the MS/MS spectra).

Database Compilation, Label-free Quantification, and Signaling Pathway Analysis

MS/MS data generated by conventional and NL-driven MS^3 data dependent acquisition via the LTQ-FT were extracted by BioWorks version 3.3 and searched against a composite

International Protein Index (IPI) zebrafish protein sequence database containing both forward and randomized sequences using Mascot version 2.2 (Matrix Science, Boston, MA) search algorithm. Mascot was searched with a fragment ion mass tolerance of 0.80 Da and a parent ion tolerance of 25.0 ppm assuming the digestion enzyme trypsin with the possibility of one missed cleavage. Carbamidomethylation of cysteine was specified as a fixed modification while oxidation of methionine, HNE Michael adduction formation on His, Cys, and Lys, as well as HNE Schiff base adducts on Lys were specified as variable modifications in the database search.

The software program Scaffold (versions Scaffold_2.0.0 and version Scaffold_3.5.1, Proteome Software Inc, Portland, OR) was employed to validate tandem MS-based peptide and protein identifications. Peptide information was accepted if it could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm (53). Protein identifications, where protein probabilities were assigned by the Protein Prophet (54), were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides. In order to discard false positives and accept false negatives, if any, manual validation of MS^n data of each HNE-modified peptide was performed by comparing the masses of b- and y-ions obtained during CID-MS/MS of peptides with the fragment ion masses generated from MS-product module of ProteinProspector (http://prospector.ucsf.edu).

Scaffold 3 readily extracts MS/MS total ion current (TIC) value for each identified peptide and calculates the total TIC for each protein. To test for significant changes in protein expression between treatments, the Student's t-test was performed on the normalized TIC values and accepted at $P<0.05$ requiring at least 1.5-fold change.

Additionally, Ingenuity Pathway Analysis (IPA) 8.8 (Ingenuity Systems, Redwood City, CA) was utilized to derive potential protein interaction networks among the identified

carbonylated proteins in the zebrafish embryos. IPA accepts the upload and analysis of human, mouse, rat, and canine identifiers. Similarly, identifiers generated from EntrezGene are also accepted and mapped according to their HomoloGene to the ortholog information in IPA's data bank (content will be specific to human, mouse, and rat) (55). As a result, accession numbers generated for identified carbonylated proteins in the zebrafish embryo samples from Scaffold software (IPI accession numbers) were cross-referenced with their associated EntrezGene ID numbers to further map out networks for zebrafish carbonylated proteins.

RESULTS

Detection of HNE-carbonylation *in vivo* in Zebrafish Embryos

Here, we began with an exploratory survey of the total effects of carbonyl accumulation, induced by HNE, of cellular constituents (i.e., proteins) *in vivo* using zebrafish embryos. Immunohistochemistry permitted the detection of the overall distribution of HNE-protein adducts within this vertebrate model. Localization of lipoxidation product immunoreactivities was confined to the cranial cavity in control embryos (Fig. 2A) whereas immunoreactivities in HNE-treated zebrafish embryos were extended throughout the dorsal and ventral regions (Fig. 2B). Lipoxidation products are oxidation-specific lipid- and/or amino acid-protein adducts (56), therefore the anti-HNE antibody used in this study was specific to HNE-protein adducts and does not cross react with other structures.

Identification of HNE-carbonylated Protein Targets from Zebrafish Embryos

To confirm detection of carbonylated targets in zebrafish embryos, proteins extracted from HNE-treated samples were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotted with anti-HNE antibody. Both Coomassie-stained SDS-PAGE gels and immunoblots revealed specific bands in regards to HNE treated groups (Fig. 3). To identify the entire zebrafish proteome during HNE-regulation, all bands generated from the Coomasie-stained SDS-PAGE gels were cut horizontally (Fig. 3A) and subjected to in-gel digestion. Proteins were identified by data-dependent reversed phase LC-MS/MS analysis and MS/MS spectra were searched against a zebrafish database using the Mascot search algorithm. Results were accepted if peptide identifications were at >95% probability along with protein identifications at >99% probability with at least two unique peptides. These criteria allowed for 311 identifications (not shown) from the zebrafish embryos' proteome. However, limitations of these experimental conditions did not allow for identification of specific HNE protein targets and/or localization of carbonylation sites.

Localization of HNE-modification of Identified Protein Targets in Zebrafish Embryos

Previously, Rauniyar *et al.* (51) indicated the importance of enrichment of HNE-modified peptides for carbonyl-directed identifications by LC-MS/MS analysis. Enrichment by solid-phase hydrazide (SPH) chemistry along with data-dependent reversed-phase (RP) LC-MS/MS, allowed for the identification of 11 proteins (Table 1) susceptible to HNE attack in the zebrafish embryos. MS/MS spectra that were extracted and searched against the zebrafish protein database using the Mascot search algorithm were mainly associated with muscle and housekeeping proteins (Table 1). Among the muscle-associated proteins were actin alpha cardiac muscle 1 (*Actc1*), slow

myosin heavy chain 1 (*Smyhc1l*), and tubulin alpha 8 like 4 (*Tuba8l4*). Furthermore, identification of posttranslational modifications (PTMs) implicated His residues as corresponding targets for modification (Table 1). Manual validation of the spectra was incorporated to approve HNE-modifications. Accordingly, examination of the MS/MS spectrum of *Actc1* revealed the doubly charged precursor ion at *m/z* 1056.58 and displaying the most intense peak at *m/z* 978.6, which is characteristic of HNE (156 Da) neutral loss (Fig. 4A). Conversely, due to limited product ions, identification by conventional data dependent acquisition methods indeed required additional confirmation. Therefore, samples were further validated by an additional neutral loss (NL)-dependent MS^3 (NL-MS^3) acquisition method.

Validation of HNE-modification of Identified Protein Targets in Zebrafish Embryos

Incorporating the NL-driven MS^3 acquisition method provided further confirmation of His sites of modification for fast myosin heavy chain 4 (*Myhc4*) and skeletal alpha actin (*Acta1*) proteins (Table 1). To confirm results, MS^3 spectra were manually evaluated. For example, *Acta1* revealed the neutral loss ion at *m/z* 978.6 and extensive fragmentation permitted sequence identification after database searching to HNE-modified actin tryptic fragment, VAPEEH*PTLLTEAPLNPK, with the HNE-modification at His (Fig. 4B). Both MS/MS and MS^3 acquisitions specified HNE-targets on His residues, indicating these modified proteins occurred under Michael addition reaction. This further correlates with previous research using this proteomic approach that reported HNE to cause modification to His residues (37, 57).

Identifying actin, tubulin, and myosin, via database searching, suggested a need for supplemental confirmation of their presence in HNE-treated groups by Western Blot analysis. As a result, using antibodies specific for actin, tubulin, and myosin, we were able to confirm their

presence within zebrafish embryos (Fig. 5). As seen in Fig. 5A, myosin heavy chain clearly depicts there is up-regulation of the protein in the HNE-treated group compared to control. Additionally, comparing control embryo samples to HNE-treated samples, we also see the actin protein levels increased in HNE-treated embryos (Fig. 5C).

HNE Exposure Drives Differential Protein Regulation within Zebrafish Embryos

Samples treated with HNE yielded a broad smear on the immunoblots. However, bands located around 20±10kDa, 50±10kDa, and 200±10kDa had noticeably stronger intensities. As a result, we further anticipated that differential protein regulation might be occurring due to HNE treatment. For example, referring to immunoblots of myosin and actin (Fig. 5A and C, respectively), there were noticeably different intensities between the control and HNE-treated samples. Consequently, the raw data generated from previously excised SDS-PAGE gel bands (ranging from 10±5kDa to 200±5kDa) were further used to confirm if protein regulation was occurring using a label-free proteomics approach. Briefly, extracted total ion currents (TICs) from identified MS/MS spectra, were used to identify differently expressed proteins due to HNE exposure. Out of the 311 identified proteins, TIC permitted a substantial confirmation of 111 differentially expressed proteins (Table 2). Moreover, a majority of these HNE-regulated proteins were observed to be up-regulated in this model organism.

In order to gather biological conclusions from the label-free quantitative data, the 111 differentially expressed proteins were submitted to Ingenuity Pathway Analysis (IPA) to extract signaling networks and biological processes associated with these proteins. From the 111 differentially expressed proteins, 96 were successfully mapped to 11 networks (Fig. 6). The top five overlapping network functions included: 1. Cancer, Hematological Disease, Protein Trafficking; 2. Cancer, DNA Replication, Recombination, and Repair, Energy Production; 3.

Cell Death and Survival, Cancer, Hematological Disease; 4. Nucleic Acid Metabolism, Small Molecule Biochemistry, DNA Replication, Recombination, Repair; and 5. Lipid Metabolism, Small Molecule Biochemistry, Molecular Transport. In addition, IPA integrates the top biological functions associated with the mapped proteins. Specifically, this was comprised of 1. Diseases and Disorders: cancer, neurological disorders, and skeletal and muscular disorders; 2. Molecular and Cellular Functions: molecular transport, protein synthesis, nucleic acid metabolism; and 3. Physiological System Development and Function: tissue development, nervous system development and function, endocrine system development and function. Furthermore, top canonical pathways included intra- and intercellular processes involved in eIF2 signaling, glycolysis, regulation of eIF4 and p70S6K signaling, remodeling of epithelial adherens junctions, and 14-3-3-mediated signaling.

DISCUSSION

The accumulation of reactive aldehydes resulting from oxidative stress and lipid peroxidation have been recognized as an underlying factor in the progression of numerous clinical conditions including cancer, neurological disorders, and cardiovascular disease (13, 58). The development of therapeutic approaches as well as the refining of disease pathologies remains the primary foci when determining the biochemical factors that elicit cellular responses resulting from protein carbonylation. Application of *in vitro* methods has been beneficial in detecting targets of protein carbonylation but there remain limitations in detecting their relevance to physiologic insults (7, 57). Conversely, *in vivo* approaches preserve essential qualities aimed at understanding the physiological impact of ROS (7, 59, 60). Together, *in vitro* and *in vivo*

methods have incorporated techniques for isolating and characterizing carbonylated proteins. It has been recognized that these techniques require a high level of specificity and sensitivity (58). Namely, gel-based and gel-free strategies, along with methods based on carbonyl trapping, have been implemented for separating and concentrating carbonylated proteins from complex solutions (13, 58). Although these techniques have been beneficial for targeting carbonyl moieties, understanding molecular mechanisms of a disease along with identifying specific targets of impairment to oxidants, is narrowly focused (61). Nevertheless, combining proteomics, mass spectrometry, and enrichment methodologies can further contribute to a better understanding of oxidant modulation. Hence, our approach concentrated on a discovery-driven proteomics approach to elucidate the effects of HNE in vivo during the early development of zebrafish embryos. This will allow for a systems analysis of the effects of HNE, which will contribute to establishing a relationship between signaling and cellular processes that are implicated in pathophysiological disorders.

Zebrafish have become a well-established model organism for studying vertebrate embryogenesis (62) as well as developmental regulatory networks (63). In agreement with these studies, our model has been demonstrated to be an indispensable source for providing information regarding oxidative stress *in vivo*. Initially we revealed that exposure of HNE in zebrafish embryos produced an overall distribution of immunoreactivities of lipoxidation products ranging throughout their dorsal and ventral cavities, thus mimicking an environment of oxidative stress *in vivo* (Fig. 2). It has been recognized that lipid peroxidation results in damage. In addition, it has further been suggested that the effects of oxidants can provoke cellular properties depending on the species present, their concentration, and their reactivity with protein targets (60). Likewise, many studies have indicated that HNE, at low concentrations, can be

found in human tissue and plasma under physiological conditions, and is able to modulate the expression of many genes involved in the control of signal transduction, cell proliferation, and differentiation (60, 64-66).

As we generated protein carbonylation *in vivo* in zebrafish embryos, we noticed differential regulation of protein expression levels due to HNE treatment. Several of the differentially expressed proteins identified in our model organism were up-regulated after exposure to HNE. The literature suggests that there is limited work on identifying critical protein targets underlying HNE's effects. Interestingly, the majority of differentially expressed proteins identified here were associated with housekeeping genes related to structure, gene regulation, and metabolism. In a "stressed" environment, the organism's defense mechanism is to control protein expression needed to repair and/or guard the genome. Therefore, uncovering differential protein expression levels induced by lipid oxidants (e.g., HNE) are beneficial in demonstrating protein stimulation and/or repression in response to oxidant stress. These proteins were interrelated to pathophysiological disorders including tissue development, hematological disorders, nervous system, and endocrine system development and function. Such processes have been associated with diseases including cancer, diabetes, neurodegeneration, and skeletal and muscular disorders.

We have shown in our study that the muscle proteins actin and myosin are modified *in vivo* by HNE exposure in zebrafish embryos. Accumulation of HNE can lead to protein carbonylation and regulation of other important proteins in response pathways in many diseases as we have shown in IPA analyses. Specifically, in the muscles, this accumulation of carbonylated proteins can lead to a disruption in how the muscle is able to process glucose and accept insulin, leading to an insulin resistance. The *in vitro* increase in HNE-modified proteins

and disruption of both glucose uptake and insulin signaling (67), is linked to cellular insulin resistance. In addition, there is a moderate correlation with HNE-modified muscle proteins and the severity of insulin resistance in clinical patients (68). A significant top canonical pathway that was identified in our IPA analysis was glycolysis (p-value = 6.66E-06). This validates our findings that *in vivo* HNE-induced carbonylation of muscle proteins affects metabolic pathways, which can ultimately lead to insulin resistance and type-2 diabetes mellitus. This area is limited, and more work needs to be done to be able to fully elucidate the role HNE has on insulin resistance and diabetes.

Oxidative stress has been linked to chronic inflammation, which may mediate a multitude of diseases including cancer, diabetes, cardiovascular disease, and neurological disorders (69). There are two stages of inflammation, acute and chronic. Acute inflammation recruits immune cells to the site of damage, causing a respiratory burst from the increased uptake of oxygen, thus producing an increased release and accumulation of ROS (69, 70). In chronic inflammation, soluble mediators such as arachidonic acid metabolites are recruited to the site of damage, producing more ROS. In turn, these different mechanisms of increased ROS are also key mediators in activating signal transduction cascades, and also cause changes in transcription factors (e.g. hypoxia-inducible factor-1α, HIF-1A; nuclear factor kappa B, NF-κB) (71). In our study, the HNE-regulated dataset was subjected to IPA analysis. Accordingly, it generated 11 networks mapping out processes regulated and influenced by HNE treatment. Coincidentally, network 2 (Cancer, DNA Replication, Recombination, and Repair, Energy Production) included several transcription regulators such as HIF1A, Enolase 1 (ENO1), C-myc myelocytomatosis viral oncogene homolog (MYC), Prohibitin (PHB), as well as the NFκB complex (Fig. 7). Once these transcription regulators become activated, they begin to mediate cellular stress responses

(69). For example, HIF1A functions as a regulator of cellular and systematic homeostatic responses to hypoxia, and thus activates many genes involved in energy metabolism as well as apoptosis. It has also been shown to play an essential role in activating transcription responses of ENO1, further establishing evidence that HIF1A mediates transcriptional activation of genes encoding glycolytic enzymes (72). Accordingly, IPA demonstrated that HIF1A directly interacted with the up-regulated ENO1 in network 2 (Fig 7) of the HNE-regulated zebrafish embryo proteins. Moreover, IPA gene view feature reported several literature citations illustrating that expression of ENO1 and HIF1A are associated with a variety of cancers including breast cancer and renal cell carcinoma respectively (73, 74). HNE-regulated protein networks created in IPA have further supplemented the field with data contributing to the theory that constant inflammation/oxidative stress progresses into a vicious cycle, potentially damaging healthy constituents of cells, and thus leads to cancer progression.

An enrichment approach permitted us to uncover a few proteins susceptible to PTM by HNE. Production of HNE and/or modification caused by this reactive aldehyde may not be the principal cause of certain diseases but identification of primary sites of modification due to HNE could provide details to disease progression (13). Additionally, PTMs of proteins can contribute to a variety of functions including protein-protein interactions, modulation of signaling cascades, cell division, and protein turnover and localization (75). Most PTMs are irreversible processes. Therefore, their contribution towards protein dysfunction may lead towards the progression of many diseases. For the first time in the zebrafish model, we have detected 11 proteins comprising of His residues as the primary sites of modification by HNE. Interestingly, of the 11 HNE-modified proteins, only TUBB4B was also shown to be regulated by this reactive aldehyde. TUBB4B is a member of the globular proteins, which make up microtubules. Neely *et*

al. (76) first indicated that HNE causes inhibition of neurite growth as well as disrupts neuronal microtubule organization and modifies cellular tubulin in concentrations relevant to Alzheimer's disease patients. Microtubules have been known to participate in physiological functions such as cell proliferation and differentiation. Similarly, microtubule fractions of rat brains were exposed to HNE, causing modification to tubulin moieties, and thus, provoking impairment to its ability to polymerize, resulting in the disappearance of microtubule structure in these cells (77). Both of these studies included results conducted *in vitro*. Our study has provided some insight into what is occurring on the systems level *in vivo.* In addition, the HNE-modified TUBB4B identified in our study was shown to be up-regulated. Furthermore, to determine if the modified proteins (excluding vitellogenin proteins) would be interrelated with or mapped into networks/pathways with the HNE-regulated proteins that were identified, we subjected the combined set for IPA analysis. The software created a slightly modified set of 11 networks that mapped both significant proteins as well as the modified proteins (Suppl. Fig. 1). Nevertheless, the new analysis showed TUBB4B to be up-regulated in network 5 (Fig. 8) (Cancer, Respiratory Disease, Inflammatory Response) which similarly included all proteins and transcriptional regulators seen in network 2, but had some additional IDs as compared to the original HNE-regulated dataset (Fig. 7). In agreement with previous results, our study further implicates a cytoskeletal impairment occurring due to HNE exposure.

CONCLUSION

To establish an *in vivo* effect of HNE, a zebrafish embryonic model was utilized. After a short-term exposure to HNE, it was recognized that the induced oxidative stress overwhelmed

the system not only through regulation of protein expression levels but also through carbonyl-modification to protein targets on specific amino acid residues. This discovery-driven proteomics-based approach constructed a large dataset of proteins along with bioinformatics information, supplementing these proteins with numerous interconnected networks and associated pathways in regards to the effects of HNE exposure. For this reason, interpretation of each network, pathway, and/or protein interaction goes beyond the scope of this study. Nevertheless, the survey of this large dataset reveals interactions affected by HNE at the protein level in the zebrafish model. Ultimately, these discoveries have provided an opportunity to generate new hypotheses about the underlying biochemistry of lipid electrophile stress on a systems level.

Table 5. A list of HNE-modified tryptic peptides extracted from zebrafish embryos and identified after SPH enrichment followed by standard data-dependent and neutral loss (NL)-driven MS^3 acquisition on a hybrid linear-ion trap-FTICR (LTQFT) mass spectrometer.

Protein Name	IPI Accession Number*	Molecule Symbol[+]	HNE-modified Peptide Sequence	Identified by NL-driven MS^3
Myosin, heavy chain 4, skeletal muscle	IPI00773363 (*799300*)	MYH4	(R)DLEESTLQ**H**EATAAALR(K)	
Vitellogenin 1[±]	IPI00508594 (*559475*)	-	(K)GCVEVHS**H**NAAFIR(N)	**
Novel protein similar to vitellogenin 1[±]	IPI00500668 (*559931*)	-	(K)**H**LVENNVAMVHDDAPLK(F)	
Actin, alpha, skeletal muscle 1	IPI00628880 (*407658*)	ACTA1	(R)VAPEE**H**PTLLTEAPLNPK(A)	
Myosin, heavy chain, cardiac muscle	IPI00902207 (*321552*)	MYH7	(R)DLEEATLQ**H**EATAATLR(K)	
Tubulin, alpha 3e	IPI00492989 (*393154*)	TUBA3E	(R)I**H**FPLATYAPVISAEK(A)	
Actin, alpha, cardiac muscle 1	IPI00504207 (*58114*)	ACTC1	(R)VAPEE**H**PTLLTEAPLNPK(A)	***
Myosin, heavy chain, skeletal muscle	IPI00497758 (*334274*)	MYH4	(R)DLEESTLQ**H**EATAAALR(K)	***
Tubulin, beta 4B, class IVb	IPI00656505 (*641421*)	TUBB4B	(R)EIV**H**LQAGQCGNQIGAK(F)	
Desmin	IPI00774314 (*30148*)	-	(R)TFGSGLGSSIFAG**H**GSSGSSGSSR(L)	
Myosin, light chain, cardiac muscle	IPI00509545 (*393488*)	-	(R)**H**VLATLGEK(M)	

[*] Entrez Gene IDs are given in parentheses

[±]Proteins excluded from Ingenuity Pathway Analysis

[+]Gene Symbols from Ingenuity Pathway Analysis

[**]Proteins identified in MS2 data-dependent and NL-driven MS3 acquisition

[***]Proteins identified only by NL-driven MS3 acquisition

Table 6. Example of differentially expressed proteins between HNE-treated and control zebrafish embryos identified and quantified by MS/MS Total Ion Current (TIC) features generated in Scaffold 3 Software. Minimum protein: 99%, minimum peptide: 95% confidence scores with a minimum of 2 peptides. Student *t*-test were accepted at $p < 0.05$ with at least a 1.5-fold change.

Protein Name	Accession Number	Molecular Weight (kDa)	Student *t*-test (p<0.05)	Control Average Total TIC ± Standard Error	HNE-Treated Average Total TIC ± Standard Error	Fold Change
A) Up-regulated in HNE-treated Zebrafish Embryos						
Histone H2B ½	IPI00488618	14	0.02	2.94E+03 ±7.04E+02	1.92E+04 ± 4.39E+03	6.5
Histone H2A	IPI00486495	14	0.009	3.19E+03 ± 9.56E+02	1.68E+04 ± 2.66E+03	5.3
Peptidyl-prolyl cis-trans isomerase	IPI00500119	24	0.04	5.45E+02 ± 3.29E+02	1.90E+03 ± 2.80E+02	3.5
Ribosomal protein L6	IPI00481915	31	0.008	1.22E+04 ± 2.97E+03	2.84E+04 ± 1.29E+03	2.3
Solute carrier family 25 (mitochondrial carrier; oxoglutarate carrier), member 11	IPI00508542	34	< 0.001	4.87E+02 ± 4.85E+02	7.48E+03 ± 2.83E+02	15.4
Acidic leucine-rich nuclear phosphoprotein 32 family member A	IPI00491296	29	0.03	6.99E+02 ± 6.97E+02	4.43E+03 ± 8.62E+02	6.3
L-lactate dehydrogenase B-A chain	IPI00495855	36	0.001	1.13E+04 ± 1.45E+03	2.91E+04 ± 1.52E+03	2.6
Ribosomal protein, large, P0	IPI01024243	34	< 0.001	6.10E+03 ± 9.17E+02	2.21E+04 ± 1.55E+03	3.6
Ribosomal protein l5b	IPI00492829	38	0.02	8.71E+03 ± 1.64E+03	2.53E+04 ± 4.21E+03	2.9
Malate dehydrogenase	IPI00495240	36	0.008	4.43E+03 ± 1.51E+03	1.37E+04 ± 1.08E+03	3.1
Fructose-bisphosphate aldolase	IPI00509500	40	0.01	3.13E+03 ± 2.22E+02	1.19E+04 ± 1.91E+03	3.8
ELAV (embryonic lethal, abnormal vision, Drosophila)-like 3 (Hu antigen C)	IPI00629354	38	0.02	9.74E+02 ± 7.14E+02	8.22E+03 ± 1.85E+03	8.4
Guanine nucleotide binding protein (G protein), beta polypeptide 1b	IPI00483806	37	0.006	1.34E+03 ± 2.81E+02	3.03E+03 ± 1.51E+02	2.3
Adenylyl cyclase-associated protein	IPI00483581	49	0.009	8.52E+03 ± 3.01E+03	2.44E+04 ± 1.33E+03	2.9
KH domain containing, RNA binding, signal transduction associated 1a	IPI00504880	40	0.008	8.80E+03 ± 2.27E+03	2.11E+04 ± 1.00E+03	2.4

Nucleosome assembly protein 1, like 1	IPI00513317	44	0.02	9.39E+03 ± 1.36E+03	2.19E+04 ± 2.97E+03	2.3
ATP synthase subunit alpha	IPI00491975	60	0.002	1.26E+05 ± 1.25E+04	2.64E+05 ± 1.43E+04	2.1
Type II basic cytokeratin	IPI00834037	54	0.002	1.53E+04 ± 1.61E+03	3.78E+05 ± 5.21E+04	24.7
Tubulin, beta 2c	IPI00500040	50	0.01	5.54E+04 ± 1.03E+04	1.09E+05 ± 6.26E+03	2.0
Keratin 5	IPI00484294	59	< 0.001	4.82E+00 ± 8.12E-01	3.19E+01 ± 1.86E-02	6.6
Glutamate dehydrogenase 1b	IPI00483853	60	0.02	1.47E+04 ± 1.88E+03	3.65E+04 ± 5.85E+03	2.5
Srl protein (Fragment)	IPI00504982	56	0.002	1.15E+04 ± 1.70E+03	2.97E+04 ± 1.78E+03	2.6
Elongation factor 1-alpha	IPI00512240	50	0.05	1.87E+04 ± 1.98E+03	2.85E+04 ± 2.83E+03	1.5
Chaperonin containing TCP1, subunit 2	IPI00505928	58	0.02	6.59E+03 ± 1.76E+03	1.95E+04 ± 2.80E+03	3.0
Aldehyde dehydrogenase 2 family (mitochondrial)a	IPI00500752	57	0.02	1.21E+03 ± 1.21E+03	1.43E+04 ± 3.51E+03	11.8
Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	IPI00512371	57	0.002	8.83E+02 ± 8.81E+02	1.61E+04 ± 1.84E+03	18.3
T-complex protein 1 subunit delta	IPI00487122	57	0.001	3.81E+03 ± 5.63E+02	1.04E+04 ± 5.98E+02	2.7
Fascin homolog 1, actin-bundling protein a (Strongylocentrotus purpuratus)	IPI00798928	55	0.003	4.07E+03 ± 3.10E+02	7.40E+03 ± 3.92E+03	1.8
Proteasome (prosome, macropain) 26S subunit, non-atpase, 3	IPI00507046	58	0.03	2.89E+03 ± 2.24E+02	1.26E+04 ± 2.79E+03	4.4
Tubulin, beta 4A class iva	IPI00996437	53	0.04	3.58E+04 ± 1.94E+04	9.60E+04 ± 5.00E+03	2.7
Retinoblastoma binding protein 4, like	IPI00484227	50	0.003	1.99E+03 ± 1.01E+03	9.73E+03 ± 6.35E+02	4.9
Tubulin, beta 4A class iva	IPI00494039	50	0.04	3.83E+04 ± 2.14E+04	1.07E+05 ± 5.48E+03	2.8
Eukaryotic translation elongation factor 2b	IPI00511483	96	0.03	1.22E+03 ± 1.22E+03	1.00E+04 ± 2.20E+03	8.2
Pyruvate kinase	IPI00494547	67	0.05	3.15E+03 ± 1.91E+03	1.03E+04 ± 1.73E+03	3.3
Type 1 cytokeratin, enveloping layer	IPI00504087	47	0.03	1.40E+03 ± 1.40E+03	6.72E+03 ± 8.87E+02	4.8
40S ribosomal protein S28	IPI00503732	8	0.03	3.22E+03 ± 1.83E+03	1.42E+04 ± 2.97E+03	4.4
Fatty acid binding protein 7, brain, a	IPI00504420	15	0.002	8.03E+02 ± 8.02E+02	1.33E+04 ± 1.50E+03	16.6
Parvalbumin-2	IPI00486605	12	< 0.001	1.17E+04 ± 6.87E+03	3.30E+05 ± 3.40E+04	28.2
Parvalbumin 1	IPI00494456	11	< 0.001	4.13E+04 ± 2.76E+03	2.02E+05 ± 8.57E+03	4.9
Histone H3.2-like	IPI00962132	17	0.02	8.59E+02 ± 4.35E+02	3.52E+04 ± 9.50E+03	41.0

Ras-related protein Rab-11B	IPI00489756	24	0.02	4.00E+02 ± 3.99E+02	3.02E+04 ± 8.30E+03	75.5
ribosomal protein S9	IPI00487430	22	0.01	1.22E+03 ± 1.22E+03	1.25E+05 ± 2.89E+04	102.6
Ribosomal protein S5	IPI00508758	23	0.007	1.20E+04 ± 3.29E+03	1.07E+05 ± 1.85E+04	8.9
Ribosomal protein L18	IPI00493625	21	0.006	4.21E+02 ± 4.20E+02	1.02E+05 ± 1.91E+04	242.5
Myosin, light chain 1, alkai, skeletal, fast	IPI00499941	21	0.005	3.47E+02 ± 3.46E+02	7.11E+04 ± 1.24E+04	205.0
High-mobility group box 1b	IPI00854267	19	0.001	5.44E+02 ± 5.43E+02	2.52E+04 ± 2.82E+03	46.3
Adenylate kinase 1	IPI00501538	21	< 0.001	8.62E.03 ± 1.41E+03	1.50E+05 ± 1.23E+04	17.4
Apolipoprotein A-I	IPI00994508	30	0.002	1.15E+04 ± 1.11E+03	7.37E+04 ± 8.29E+03	6.4
Apolipoprotein A-I	IPI00495830	30	< 0.001	1.57E+04 ± 4.78E+03	4.17E+05 ±2.68E+04	26.5
Ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	IPI00512853	24	0.03	2.45E+04 ± 4.09E+03	6.67E+04 ± 1.21E+04	2.7
Triosephosphate isomerase B	IPI00484161	27	0.01	5.25E+03 ± 6.22E+02	2.52E+04 ± 4.23E+03	4.8
GTP-binding nuclear protein Ran	IPI00505092	24	0.007	1.05E+04 ± 2.01E+03	6.98E+04 ± 1.13E+04	6.7
Triosephosphate isomerase A	IPI00498329	27	0.002	4.40E+03 ± 2.12E+03	2.19E+04 ± 1.39E+03	5.0
Ribosomal protein L15	IPI00499892	24	0.001	3.36E+03 ± 1.12E+03	1.02E+05 ± 1.20E+04	30.4
Adenylate kinase 2, mitochondrial	IPI00505523	27	0.003	7.49E+02 ± 3.78E+02	2.83E+04 ± 4.26E+03	37.8
Crystallin, beta B1, like	IPI00607401	28	< 0.001	5.12E+02 ± 5.11E+02	1.79E+04 ± 9.90E+02	34.9
60S ribosomal protein L19	IPI00501378	23	0.007	7.20E+02 ± 3.60E+02	2.01E+04 ± 3.75E+03	27.9
Proteosome, 26S subunit	IPI00505432	31	< 0.001	5.03E+02 ± 5.02E+02	1.13E+04 ± 9.64E+02	22.4
Prohibitin	IPI00972490	34	0.002	2.23E+03 ± 2.23E+03	4.66E+04 ± 5.52E+03	20.9
60S ribosomal protein L13	IPI00500626	24	< 0.001	2.15E+04 ± 2.67E+03	3.96E+05 ± 3.13E+04	18.4
Glutathione S-transferase M	IPI00482761	26	0.04	8.72E+02 ± 8.71E+02	1.49E+04 ± 4.76E+03	17.1
Etfb protein	IPI00492814	28	0.009	1.31E+03 ± 4.57E+02	2.15E+04 ± 4.17E+03	16.4
Ribosomal protein S2	IPI00493658	30	< 0.001	4.14E+03 ± 8.02E+02	5.92E+04 ± 4.68E+03	14.3
Apolipoprotein Ea	IPI00607437	32	0.002	1.87E+03 ± 1.25E+03	2.60E+04 ± 3.32E+03	13.9
Protein nipsnap homolog 2	IPI00490338	34	< 0.001	3.90E+03 ± 5.07E+02	4.93E+04 ± 4.70E+03	12.6
Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide b	IPI00501838	28	< 0.001	1.13E+04 ± 2.76E+03	1.22E+05 ± 7.30E+03	10.8
14-3-3 protein beta/alpha-A	IPI00509301	28	< 0.001	5.42E+04 ± 1.19E+04	4.79E+05 ± 2.68E+04	8.8
Ywhai protein	IPI00511497	28	< 0.001	1.23E+04 ± 2.14E+03	1.01E+05 ± 5.56E+03	8.2

14-3-3 protein beta/alpha-B	IPI00498898	27	< 0.001	3.25E+04 ± 9.22E+03	2.57E+05 ± 1.82E+04	7.9
Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide	IPI00500460	28	< 0.001	1.72E+04 ± 2.53E+03	1.32E+05 ± 1.02E+04	7.7
40S ribosomal protein S4, X isoform	IPI00493144	30	0.002	1.01E+04 ± 5.80E+02	7.70E+04 ± 8.73E+03	7.7
Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide like	IPI00509935	28	< 0.001	2.83E+04 ± 3.98E+03	2.03E+05 ± 1.93E+04	7.2
Ribosomal protein L7	IPI00500824	28	< 0.001	3.95E+04 ± 5.65E+03	2.81E+05 ± 2.34E+04	7.1
Crystallin, beta B1	IPI00861740	27	0.004	7.22E+03 ± 1.65E+03	4.31E+04 ± 5.64E+03	6.0
Hydroxysteroid (17-beta) dehydrogenase 10	IPI00505229	27	< 0.001	1.12E+04 ± 1.73E+03	4.22E+04 ± 2.93E+03	3.8
Proteasome subunit alpha type	IPI00490313	26	0.04	8.43E+03 ± 1.51E+03	1.64E+04 ± 2.29E+03	1.9
Solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 4	IPI00511653	33	0.003	1.74E+03 ± 2.97E+02	4.50E+04 ± 6.40E+03	25.9
Keratin 23	IPI00833704	92	0.02	1.49E+06 ± 1.45E+05	2.34E+06 ± 3.19E+05	1.6
Enolase	IPI00483215	47	0.05	1.24E+04 ± 2.08E+02	4.54E+04 ± 2.02E+04	3.6
GDP dissociation inhibitor 2	IPI00498442	51	0.05	8.87E+03 ± 3.51E+03	3.72E+04 ± 1.64E+04	4.2
Glycine amidinotransferase, mitochondrial	IPI00504541	48	0.04	6.32E+03 ± 3.16E+03	3.42E+04 ± 1.56E+04	5.4
Actin, beta 2	IPI00656493	42	0.03	4.32E+05 ±4.46E+04	7.52E+05 ± 1.57E+05	1.7
Heterogeneous nuclear ribonucleoprotein A/B	IPI00491050	37	0.03	8.14E+04 ± 3.62E+03	1.40E+05 ± 3.06E+04	1.7
60S acidic ribosomal protein P0	IPI00497680	35	0.02	1.56E+04 ± 7.22E+02	2.81E+04 ± 5.52E+03	1.8
Proliferation-associated 2G4, a	IPI00489868	43	0.02	7.69E+03 ± 3.91E+03	3.60E+04 ± 1.08E+04	4.7
Type I cytokeratin, enveloping layer, like	IPI00636729	47	0.02	1.51E+06 ± 4.87E+04	2.28E+06 ± 3.31E+05	1.5
Keratin 23 (histone deacetylase inducible)	IPI00758877	47	0.02	3.09E+05 ± 1.59E+05	1.25E+06 ± 2.82E+05	4.0
Keratin 23 (histone deacetylase	IPI00886613	50	0.01	1.31E+05 ± 2.70E+04	2.75E+05 ± 3.42E+04	2.1

inducible)						
Actin, cytoplasmic 1	IPI00482295	42	0.008	6.14E+05 ± 5.52E+04	1.03E+06 ± 1.14E+05	1.7
Keratin 17	IPI00498656	46	0.005	7.70E+05 ± 2.23E+04	1.57E+06 ± 2.44E+05	2.0
Acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain	IPI00554392	46	0.003	2.16E+03 ± 2.16E+03	3.26E+04 ± 7.54E+03	15.1
Beta-enolase	IPI00490877	47	0.002	9.52E+03 ± 5.97E+02	9.25E+04 ± 1.99E+04	9.7
Ribosomal protein L3	IPI00511925	46	0.002	2.83E+04 ± 5.43E+02	1.01E+05 ± 1.73E+04	3.6
Isocitrate dehydrogenase	IPI00500416	50	< 0.001	2.67E+04 ± 5.11E+03	9.59E+04 ± 4.07E+03	3.6
Ybx1 protein	IPI00837236	35	< 0.001	7.88E+02 ± 4.12E+02	1.93E+04 ± 2.13E+03	24.5
Transketolase-like protein 2	IPI00498510	68	0.04	1.08E+05 ± 2.08E+04	2.72E+05 ± 4.87E+04	2.5
5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	IPI00501402	64	0.003	4.63E+04 ± 1.21E+04	1.47E+05 ± 1.05E+04	3.2
Tubulin, alpha 2	IPI00488901	50	< 0.001	5.49E+02 ± 5.48E+02	7.01E+04 ± 6.00E+03	127.7
Heat shock protein 90, beta (grp94), member 1	IPI00506057	91	0.05	2.44E+04 ± 7.24E+02	3.87E+04 ± 5.03E+03	1.6
Atpase, Ca++ transporting, cardiac muscle, fast twitch 1	IPI00505794	110	0.04	8.34E+05 ± 9.73E+04	1.15E+06 ± 2.81E+04	1.4
Heat shock protein HSP 90-beta	IPI00491598	83	0.03	1.32E+05 ± 3.08E+04	2.43E+05 ± 7.72E+03	1.8
Actinin alpha 3a	IPI00509187	104	0.02	6.74E+03 ± 5.23E+03	2.88E+04 ± 3.34E+03	4.3
Solute carrier family 25 alpha, member 5	IPI00492110	33	0.02	4.71E+03 ± 2.84E+03	1.75E+04 ± 1.96E+03	3.7
Protein disulfide isomerase associated 4	IPI00507572	73	0.02	6.61E+03 ± 4.08E+03	6.57E+04 ± 1.48E+04	9.9
Myomesin 2 (Fragment)	IPI00960438	43	0.006	9.49E+03 ± 9.49E+03	6.97E+04 ± 5.95E+03	7.3

Continued from Table 2

Protein Name	Accession Number	Molecular Weight (kDa)	Student *t*-test ($p<0.05$)	Control Average Total TIC ± Standard Error	HNE-Treated Average Total TIC ± Standard Error	Fold Change
B) Down-regulated in HNE-treated Zebrafish Embryos						
Ribosomal protein L22	IPI00491519	16	0.04	4.07E+03 ± 8.29E+02	1.08E+03 ± 5.63E+02	0.3
Ribosomal protein S13	IPI00494738	17	0.03	8.26E+02 ± 1.92E+02	1.07E+02 ± 1.05E+02	0.1
Cofilin 2, like	IPI00511345	19	0.02	5.75E+02 ± 1.18E+02	5.44E+01 ± 5.24E+01	0.09
ATP synthase subunit beta	IPI00809369	56	0.04	7.12E+04 ± 9.46E+03	3.98E+04 ± 3.70E+03	0.6
ATP synthase subunit beta	IPI00897805	56	0.007	1.16E+05 ± 9.12E+03	6.66E+04 ± 3.00E+03	0.6
Ribosomal protein S19	IPI00489287	16	0.03	3.06E+03 ± 4.58E+02	9.19E+02 ± 4.58E+02	0.3

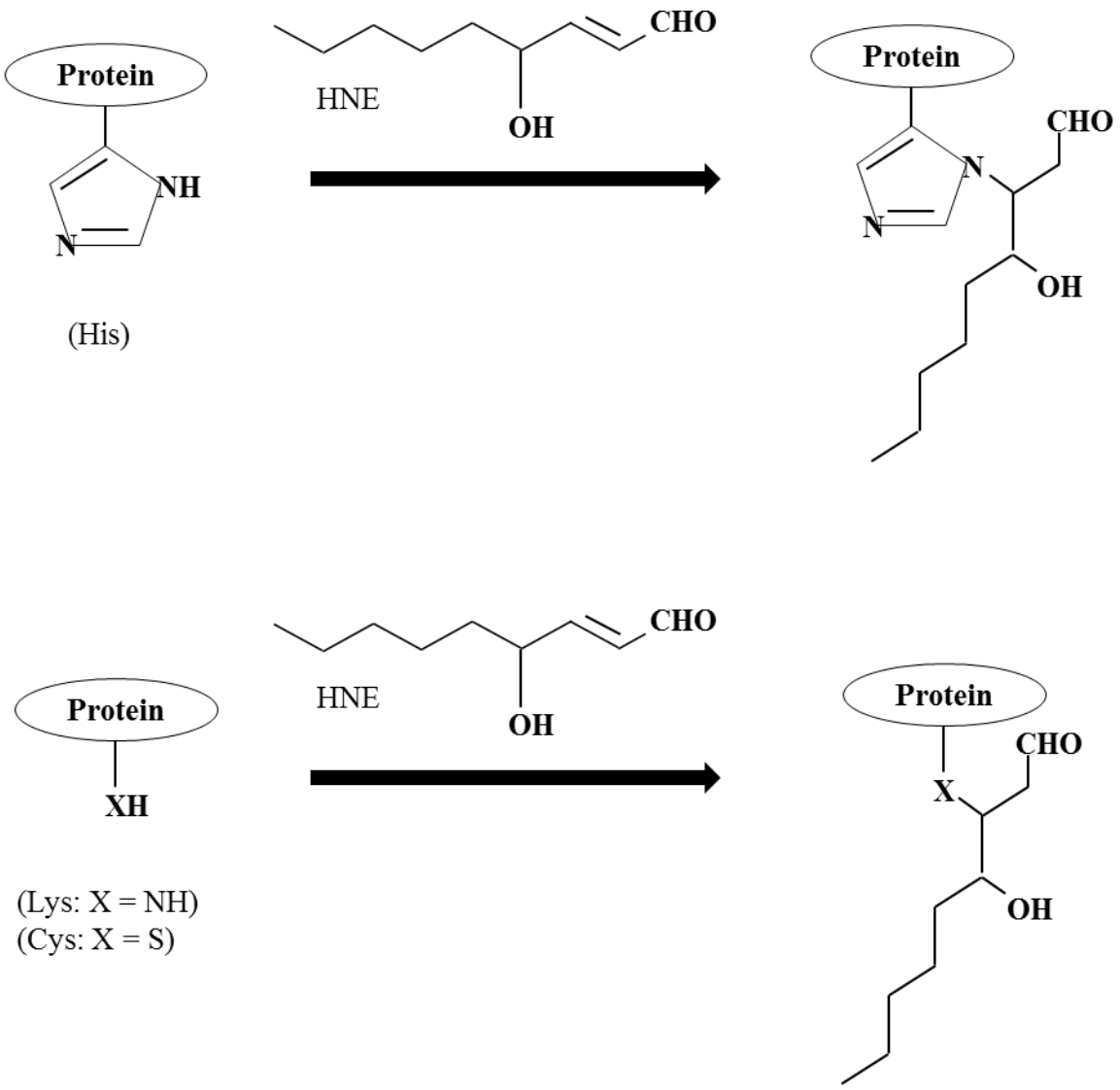


Figure 20. Schematic illustration of HNE-modification on an amino acid residue (His, Cys or Lys) of a protein via Michael addition. Adapted with permission from reference 47. Copyright 2007 John Wiley & Sons, Ltd.

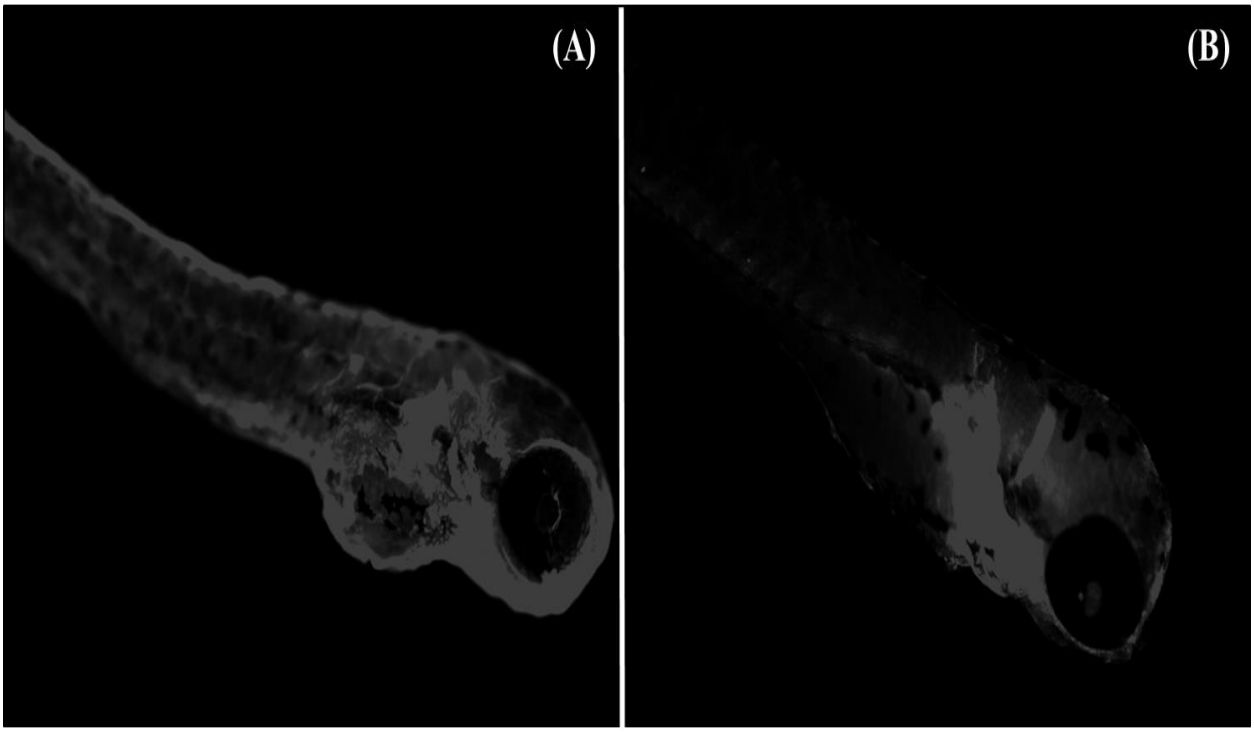


Figure 21. Immunohistochemical detection of lipoxidation product, 4-hydroxy-2-nonenal (HNE)-protein adducts, in 0.2 mM HNE (A) or control (B). HNE-protein adducts antibodies (recognizing the occurrence of HNE-induced PTMs) is a representative case of carbonyl-induced stress in zebrafish embryos.

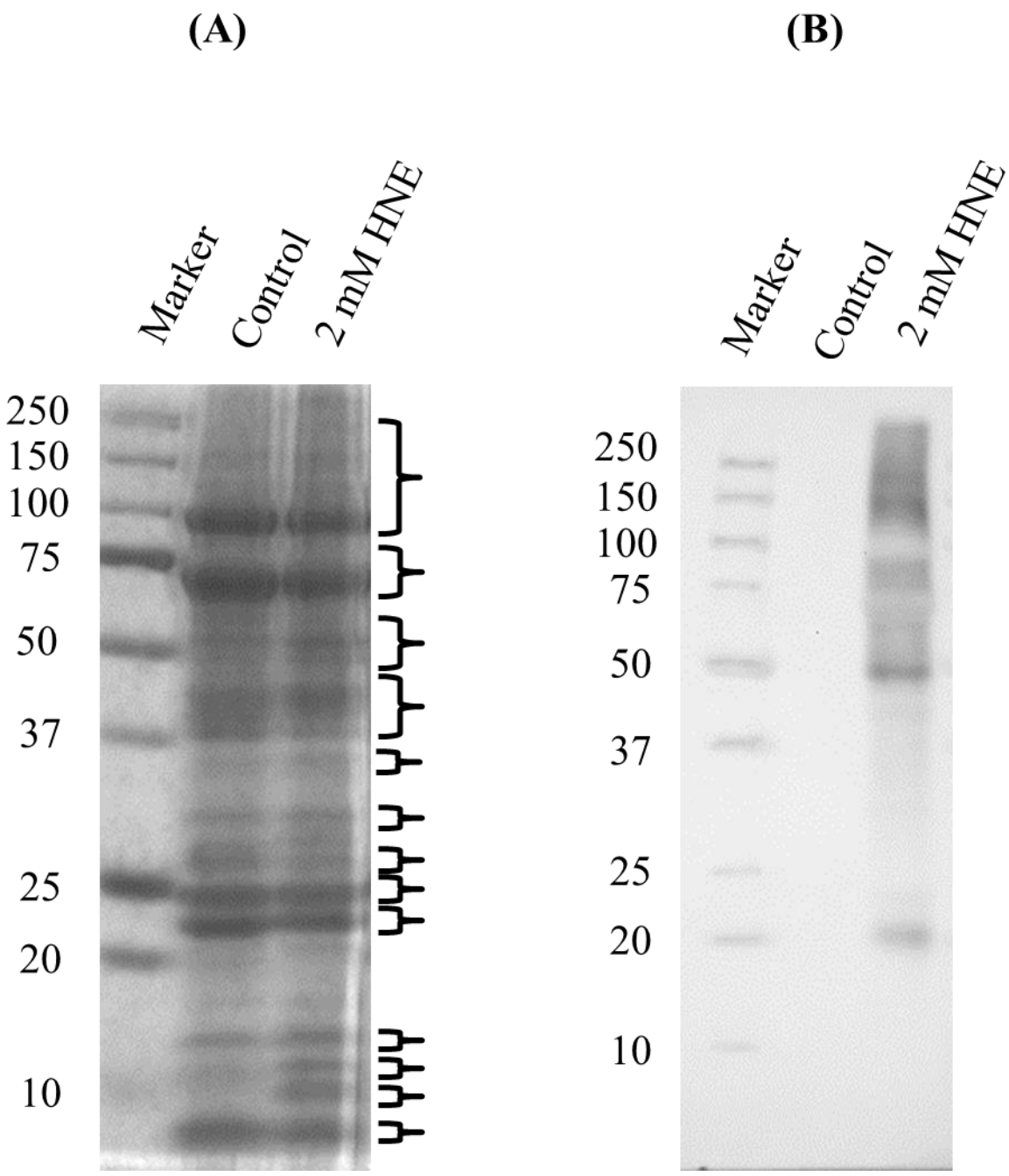


Figure 22. SDS-PAGE and immunoblots of 2mM HNE-treated proteins from 3-days post fertilization (dpf) zebrafish embryos. Coomassie-stained SDS-PAGE (A) and anti-HNE (B) immunodetection from the SDS-PAGE. Right braces indicate bands cut for in-gel digestion and protein identification.

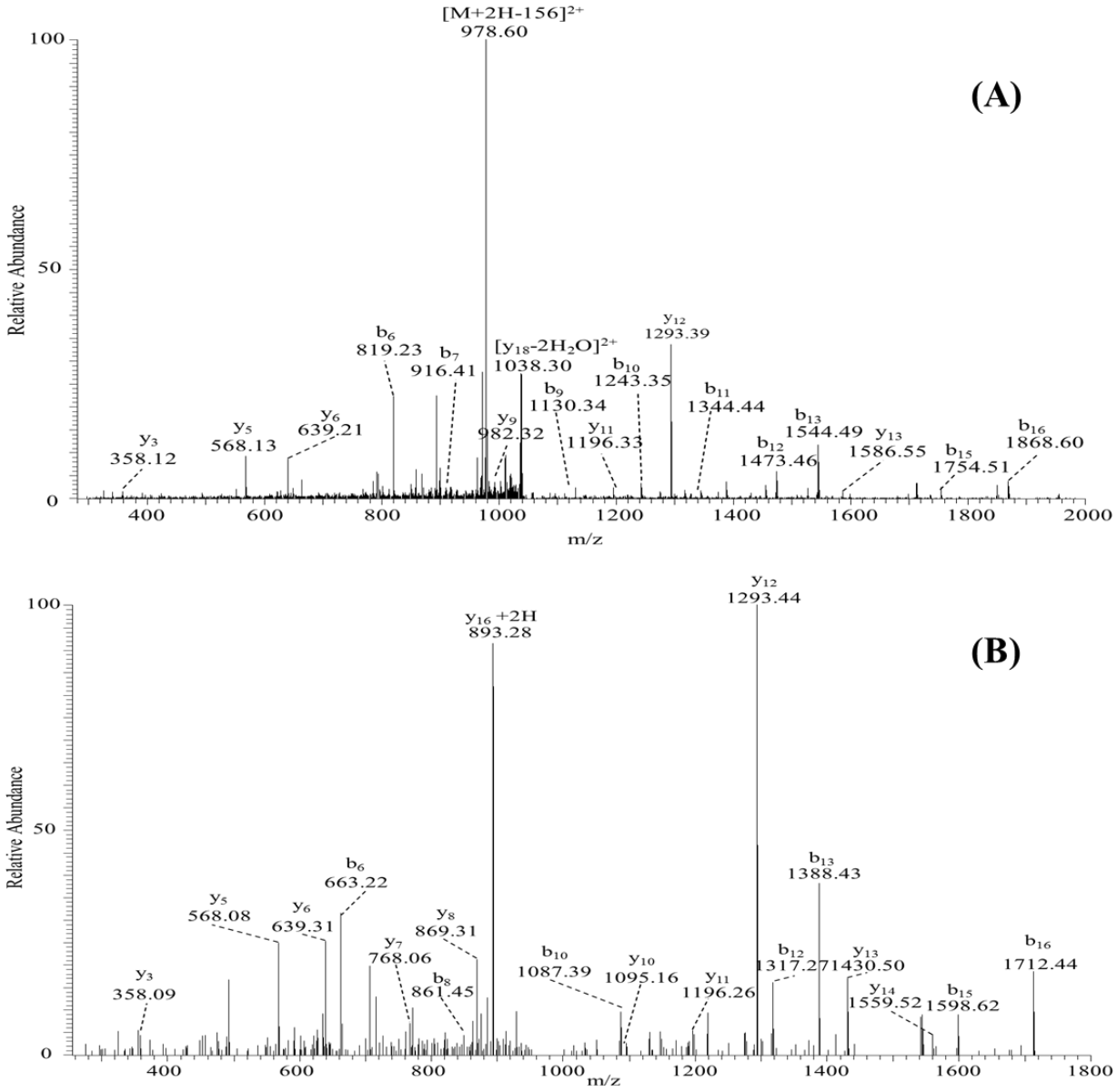


Figure 23. MS/MS spectrum for $[M + 2H]^{2+}$ ion of VAPEE<u>H</u>PTLLTEAPLNPK from *actin, alpha, cardiac muscle 1;* ACTC1 (*m/z* 1056.60) (A) HNE-modified amino acid residue is underlined. MS/MS/MS spectrum of the neutral loss ion *m/z* 978.60 obtained by NL-MS^3 method for ACTC1 (B).

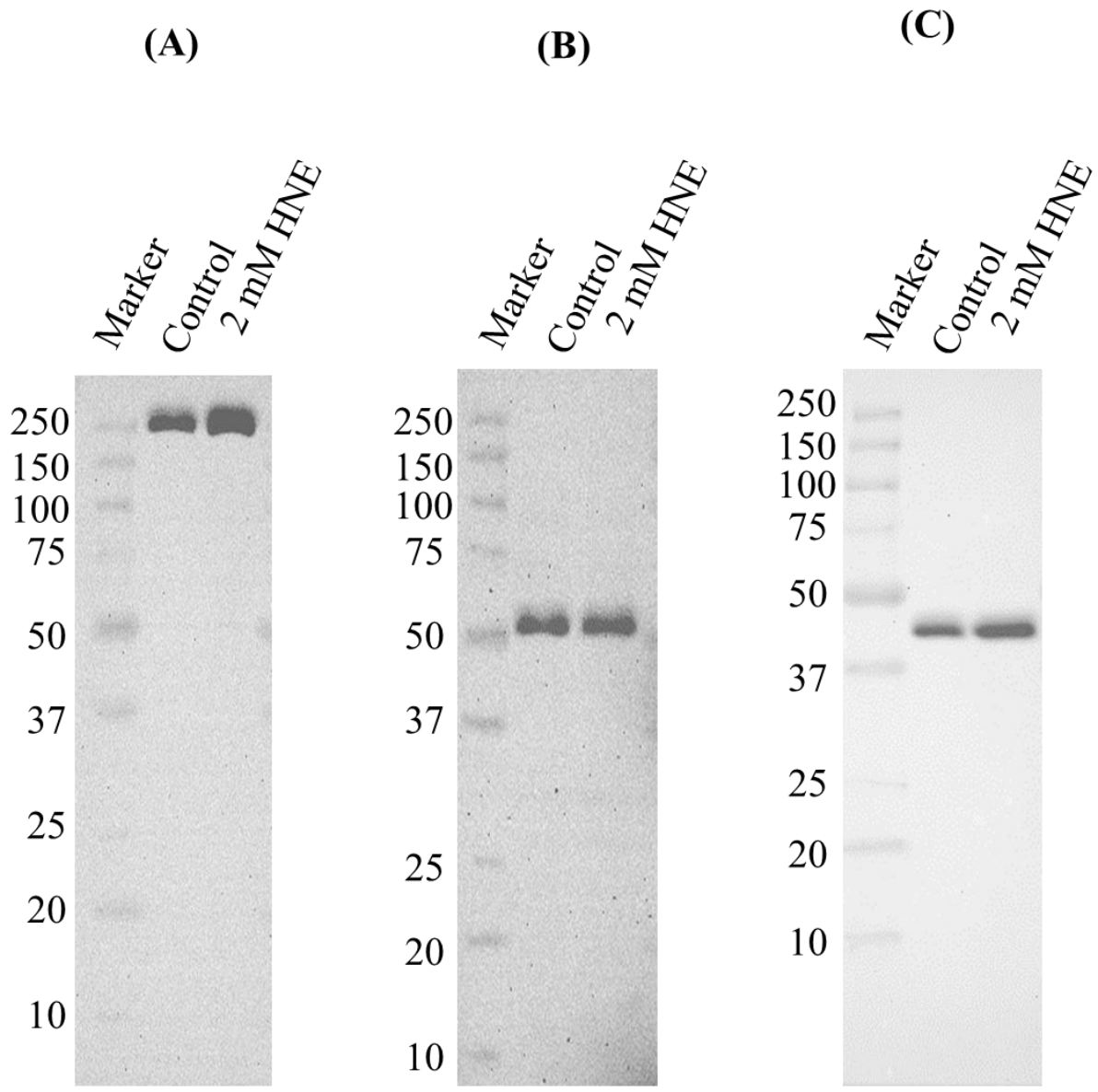


Figure 24. Immunoblots of 2mM HNE-treated proteins from 3-days post fertilization (dpf) zebrafish embryos. Anti-myosin heavy chain (~223 kDa) (A), anti-alpha tubulin (~50 kDa) (B), and anti-actin (~42 kDa) (C) immunodetection from the SDS-PAGE.

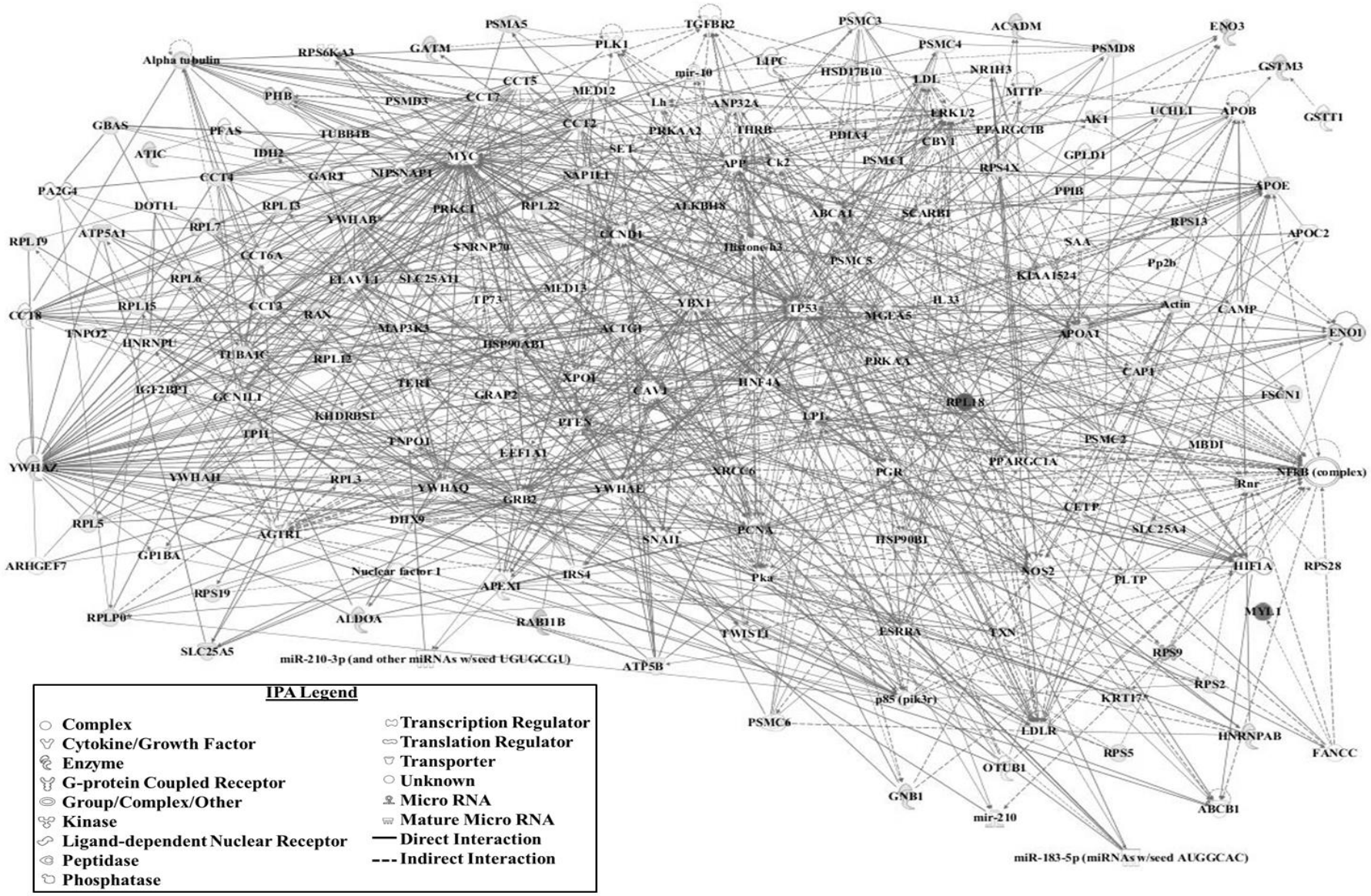

PSMA5
TGFBR2
PSMC3
ACADM
ENO3
Alpha tubulin
RPS6KA3
GATM
PLK1
PSMC4
PSMD8
GSTM3
L1PC
HSD17B10
NR1H3
CCT5
mir-10
LDL
PHB
PSMD3
CCT7
MED12
Lh
ANP32A
MTTP
ERK1/2
UCHL1
APOB
GSTT1
GBAS
PFAS
TUBB4B
CCT2
PRKAA2
THRB
PDIA4
AK1
PPARGC1B
CBY1
ATIC
IDH2
MYC
SET
APP
Ck2
PSMC1
GPLD1
RPS4X
CCT4
GART
NIPSNAP1
NAP1L1
PA2G4
PPIB
APOE
DOT1L
RPL13
PRKCI
RPL22
ALKBH8
ABCA1
SCARB1
RPS13
RPL7
YWHAB
ATP5A1
CCND1
APOC2
RPL19
SNRNP70
Histone h3
SAA
CCT6A
PSMC5
Pp2b
RPL6
ELAVL1
SLC25A11
KIAA1524
MED13
TP73
YBX1
TP53
IL33
Actin
CAMP
CCT8
RPL15
CCT3
RAN
MGEA5
ENO1
TNPO2
MAP3K3
ACTG1
APOA1
HNRNPU
HSP90AB1
TUBA1C
RPLI2
PRKACA
CAP1
IGF2BP1
TERT
XPOI
HNF4A
GCNL1
GRAP2
CAV1
FSCN1
RPL18
KHDRBS1
PTEN
LPL
TPH1
TNPO1
PSMC2
MBDI
EEF1A1
YWHAZ
PPARGC1A
NFKB (complex)
YWHAH
RPL3
XRCC6
PGR
YWHAQ
YWHAE
Rnr
GRB2
CETP
RPL5
DHX9
PCNA
SLC25A4
AGTR1
SNAI1
HSP90B1
GP1BA
ARHGEF7
HIF1A
RPS28
Nuclear factor 1
NOS2
PLTP
IRS4
Pka
APEX1
RPS19
MYL1
RPLP0*
ALDOA
RAB1B
TWIST1
ESRRA
TXN
SLC25A5
RPS9
miR-210-3p (and other miRNAs w/seed UGUGCGU)
ATP5B
RPS2
p85 (pik3r)
KRT17*
PSMC6
LDLR
HNRNPAB
FANCC
RPS5
OTUB1
GNB1
ABCB1
mir-210
miR-183-5p (miRNAs w/seed AUGGCAC)
IPA Legend
Complex
Cytokine/Growth Factor
Enzyme
G-protein Coupled Receptor
Group/Complex/Other
Kinase
Ligand-dependent Nuclear Receptor
Peptidase
Phosphatase
Transcription Regulator
Translation Regulator
Transporter
Unknown
Micro RNA
Mature Micro RNA
Direct Interaction
Indirect Interaction

Figure 25. Networks constructed through Ingenuity Pathway Analysis based on HNE-regulated proteins. Networks included 4 down-regulated proteins and 74 up-regulated proteins. Top Functions involved: (1) Cancer, Hematological Disease, Protein Trafficking; (2) Cancer, DNA Replication, Recombination, and Repair, Energy Production; (3) Cell Death and Survival, Cancer, Hematological Disease; (4) Nucleic Acid Metabolism, Small Molecule Biochemistry, DNA Replication, Recombination, and Repair; (5) Lipid Metabolism, Small Molecule Biochemistry, Molecular Transport; (6) Amino Acid Metabolism, Developmental Disorder, Hereditary Disorder; (7) Cell Cycle, Cell Death and Survival, Cellular Assembly and Organization; (8) Cardiovascular System, Development and Function, Molecular Transport, Neurological Disease; (9) Molecular Transport, Nucleic Acid Metabolism, Small Molecule Biochemistry; (10) Drug Metabolism, Protein Synthesis, Cell Cycle; and (11) Cell Morphology, Cellular Assembly and Organization, Cancer. Single lines indicate protein-protein interactions from the network diagram and arrows specify proteins/compounds that regulate another protein. The intensity of green and red molecule colors represents the degree of down- or up-regulation, respectively. Solid or dashed lines show direct or indirect interactions respectively.

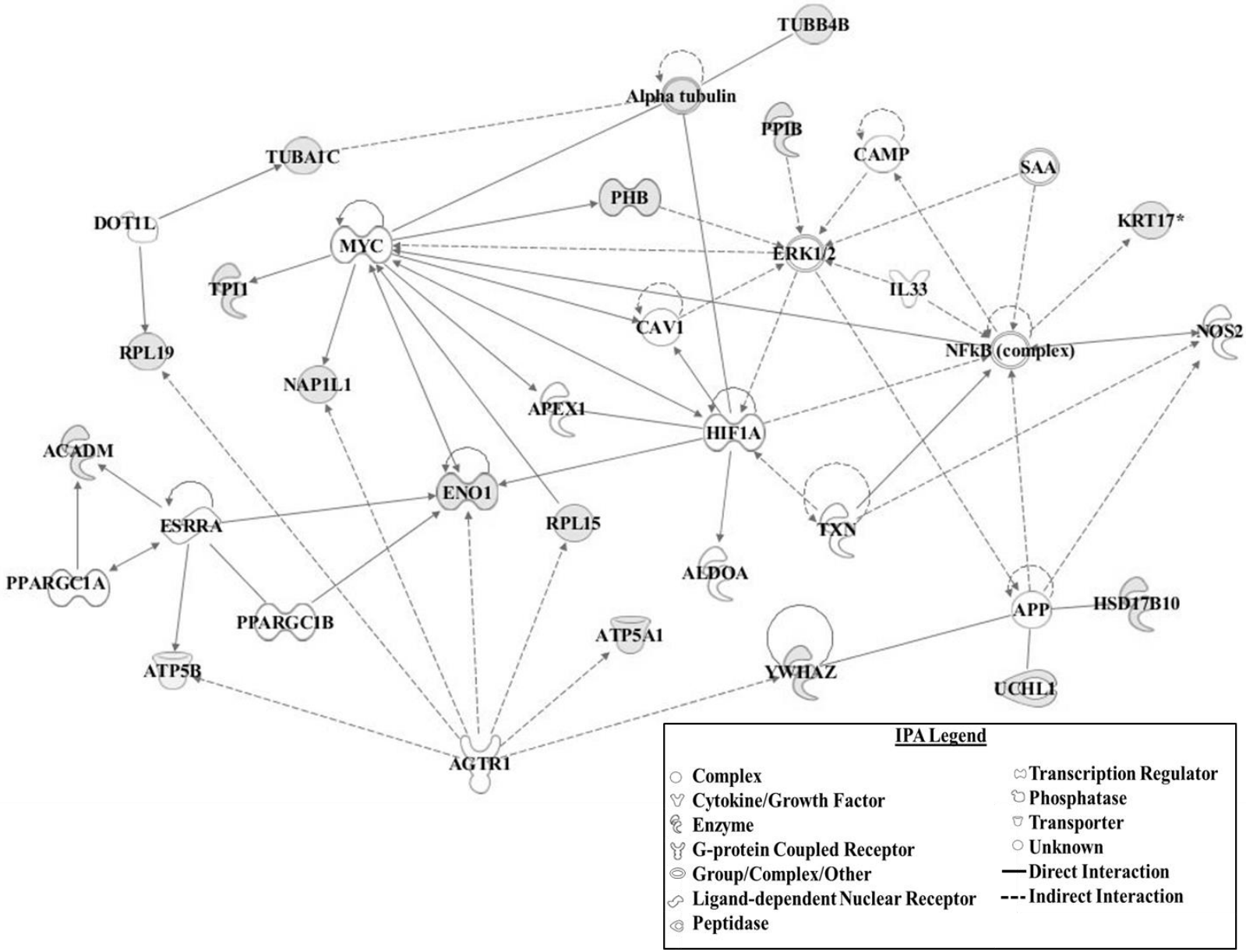

TUBB4B
Alpha tubulin
PPIB
CAMP
SAA
TUBA1C
DOT1L
PHB
KRT17*
MYC
ERK1/2
TPI1
IL33
CAV1
NOS2
RPL19
NFkB (complex)
NAP1L1
APEX1
HIF1A
ACADM
ENO1
ESRRA
RPL15
TXN
ALDOA
PPARGC1A
APP
HSD17B10
PPARGC1B
ATP5A1
ATP5B
YWHAZ
UCHL1
AGTR1
IPA Legend
Complex
Cytokine/Growth Factor
Enzyme
G-protein Coupled Receptor
Group/Complex/Other
Ligand-dependent Nuclear Receptor
Peptidase
Transcription Regulator
Phosphatase
Transporter
Unknown
Direct Interaction
Indirect Interaction

Figure 26. Based on the HNE-regulated proteins, Ingenuity Pathway Analysis constructed Network 2: Cancer, DNA Replication, Recombination, and Repair, Energy Production. Network 2 is the second highest scoring network in the HNE-regulated dataset, showing 35 molecules, 16 up-regulated (red) and 1 down-regulated (green). Single lines indicate protein-protein interactions and arrows indicate proteins/compounds that regulate another protein. Solid or dashed lines represent direct or indirect interactions respectively.

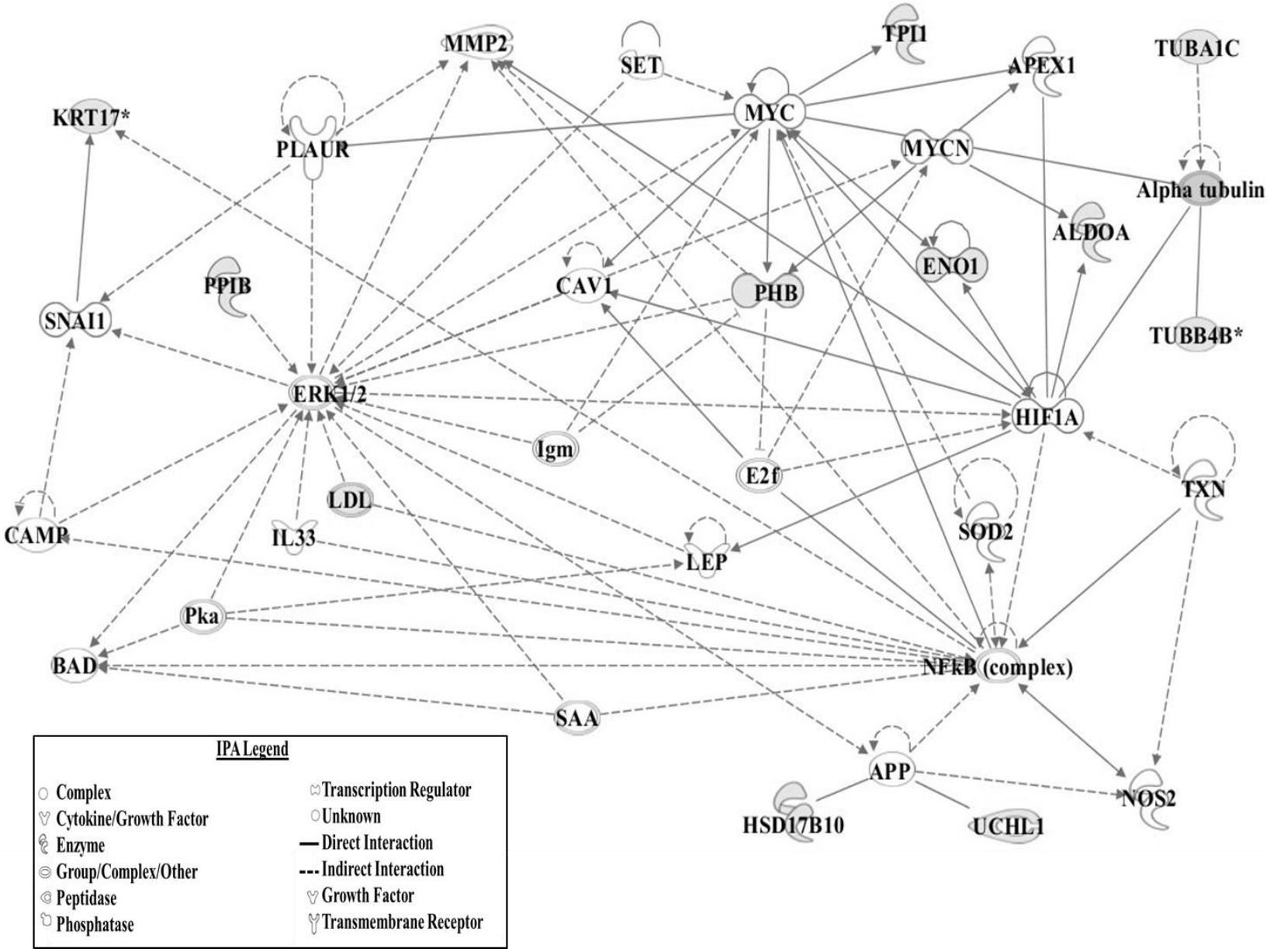
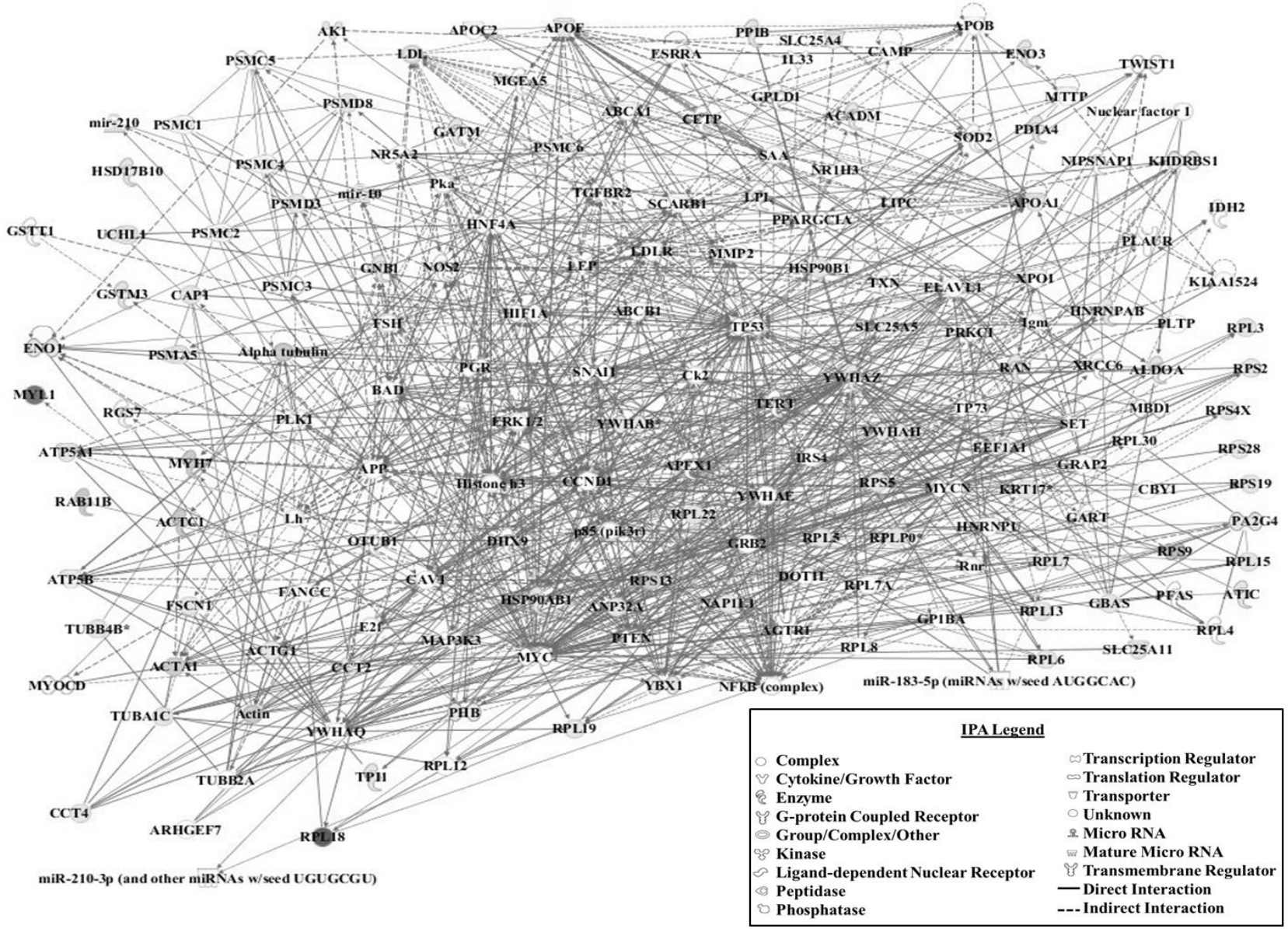

MMP2
SET
TPI1
TUBA1C
APEX1
KRT17*
MYC
PLAUR
MYCN
Alpha tubulin
ALDOA
ENO1
PPIB
CAV1
PHB
SNAI1
TUBB4B*
ERK1/2
HIF1A
Igm
E2f
TXN
LDL
CAMP
IL33
SOD2
LEP
Pka
BAD
NFkB (complex)
SAA
APP
NOS2
HSD17B10
UCHL1
IPA Legend
Complex
Cytokine/Growth Factor
Enzyme
Group/Complex/Other
Peptidase
Phosphatase
Transcription Regulator
Unknown
Direct Interaction
Indirect Interaction
Growth Factor
Transmembrane Receptor

Figure 27. Based on the HNE-regulated along with HNE-modified proteins, Ingenuity Pathway Analysis constructed Network 5: Cancer, Respiratory Disease, Inflammatory Response. Network 5 is the fifth top network in the HNE-modified/HNE-regulated dataset, showing 35 molecules, 10 that are up-regulated (red). Single lines indicate protein-protein interactions and arrows indicate proteins/compounds that regulate another protein. Solid or dashed lines represent direct or indirect interactions respectively.


AK1
APOC2
APOF
PPIB
SLC25A4
APOB
PSMC5
LDL
ESRRA
CAMP
ENO3
IL33
TWIST1
MGEA5
GPLD1
MTTP
PSMD8
ABCA1
CETP
ACADM
Nuclear factor 1
mir-210
PSMC1
GATM
SOD2
PDIA4
NR5A2
PSMC6
SAA
MPSNAP1
KHDRBS1
HSD17B10
PSMC4
NR1H3
Pka
mir-10
TGFBR2
SCARB1
LPL
LIPC
APOA1
IDH2
PSMD3
HNF4A
PPARGC1A
GSTT1
UCHL1
PSMC2
PLAUR
LDLR
MMP2
GNB1
NOS2
LEP
HSP90B1
TXN
EEAVL1
XPO1
KIAA1524
GSTM3
CAP1
PSMC3
HNRNPAB
HIF1A
ABCB1
TSH
TP53
SLC25A5
PRKCI
Tgm
PLTP
RPL3
ENO1
PSMA5
Alpha tubulin
PGR
SNAI1
Ck2
YWHAZ
RAN
XRCC6
ALDOA
RPS2
MYL1
BAD
TERT
TP73
MBD1
RPS4X
RGS7
PLK1
ERK1/2
YWHAB*
YWHAH
SET
RPL30
ATP5A1
MYH7
APP
APEX1
IRS4
EEF1A1
RPS28
GRAP2
Histone h3
CCND1
RPS5
MYCN
KRT17*
CBY1
RPS19
RAB11B
YWHAE
ACTC1
Lh
RPL22
GART
PA2G4
OTUB1
DHX9
p85 (pik3r)
GRB2
RPL5
RPLP0*
HNRNPU
RPS9
RPL15
Rnr
RPL7
ATP5B
CAV1
RPS13
DOT1L
RPL7A
FANCC
FSCN1
HSP90AB1
ANP32A
NAP1L1
GBAS
PFAS
ATIC
RPLI3
TUBB4B*
E2f
AGTRI
GP1BA
RPL4
MAP3K3
PTEN
ACTG1
RPL8
SLC25A11
MYC
ACTA1
RPL6
CCT2
MYOCD
YBX1
NFkB (complex)
miR-183-5p (miRNAs w/seed AUGGCAC)
TUBA1C
Actin
PHB
YWHAQ
RPL19
RPL12
TUBB2A
TPI1
CCT4
ARHGEF7
RPL18
miR-210-3p (and other miRNAs w/seed UGUGCGU)
IPA Legend
Complex
Cytokine/Growth Factor
Enzyme
G-protein Coupled Receptor
Group/Complex/Other
Kinase
Ligand-dependent Nuclear Receptor
Peptidase
Phosphatase
Transcription Regulator
Translation Regulator
Transporter
Unknown
Micro RNA
Mature Micro RNA
Transmembrane Regulator
Direct Interaction
Indirect Interaction

Supplementary Figure 1. Networks constructed through Ingenuity Pathway Analysis based on HNE-regulated along with HNE-modified proteins. Networks included 5 HNE-modified proteins, 4 down-regulated proteins, and 74 up-regulated proteins. Top Functions involved: (1) Cancer, Hematological Disease, Protein Synthesis; (2) Cancer, Hematological Disease, Protein Trafficking; (3) Lipid Metabolism, Small Molecule Biochemistry, Vitamin and Mineral Metabolism; (4) DNA Replication, Recombination, and Repair, Energy Production, Nucleic Acid Metabolism; (5) Cancer, Respiratory Disease, Inflammatory Response; (6) Amino Acid Metabolism, Developmental Disorder, Hereditary Disorder; (7) Cell Cycle, Cell Death and Survival, Cellular Assembly and Organization; (8) Cardiovascular System, Development and Function, Molecular Transport, Neurological Disease; (9) Molecular Transport, Nucleic Acid Metabolism, Small Molecule Biochemistry; (10) Drug Metabolism, Protein Synthesis, Cell Cycle; and (11) Cell Morphology, Cellular Assembly and Organization, Cancer. Single lines indicate protein-protein interactions from the network diagram and arrows specify proteins/compounds that regulate another protein. Solid or dashed lines show direct or indirect interactions respectively.

REFERENCES

1. Beckman KB, Ames BN. The free radical theory of aging matures. Physiol Rev. 1998 Apr;78(2):547-81.

2. Shacter E. Quantification and significance of protein oxidation in biological samples. Drug Metab Rev. 2000 Aug-Nov;32(3-4):307-26.

3. Dalle-Donne I, Giustarini D, Colombo R, Rossi R, Milzani A. Protein carbonylation in human diseases. Trends Mol Med. 2003 Apr;9(4):169-76.

4. Dalle-Donne I, Aldini G, Carini M, Colombo R, Rossi R, Milzani A. Protein carbonylation, cellular dysfunction, and disease progression. J Cell Mol Med. 2006 Apr-Jun;10(2):389-406.

5. Sayre LM, Lin D, Yuan Q, Zhu X, Tang X. Protein adducts generated from products of lipid oxidation: Focus on HNE and one. Drug Metab Rev. 2006;38(4):651-75.

6. Feng J, Xie H, Meany DL, Thompson LV, Arriaga EA, Griffin TJ. Quantitative proteomic profiling of muscle type-dependent and age-dependent protein carbonylation in rat skeletal muscle mitochondria. J Gerontol A Biol Sci Med Sci. 2008 Nov;63(11):1137-52.

7. Fritz KS, Petersen DR. An overview of the chemistry and biology of reactive aldehydes. Free Radic Biol Med. 2012 Jun 28.

8. Dizdaroglu M, Jaruga P, Birincioglu M, Rodriguez H. Free radical-induced damage to DNA: Mechanisms and measurement. Free Radic Biol Med. 2002 Jun 1;32(11):1102-15.

9. Nystrom T. Role of oxidative carbonylation in protein quality control and senescence. EMBO J. 2005 Apr 6;24(7):1311-7.

10. Dalle-Donne I, Carini M, Vistoli G, Gamberoni L, Giustarini D, Colombo R, et al. Actin Cys374 as a nucleophilic target of alpha,beta-unsaturated aldehydes. Free Radic Biol Med. 2007 Mar 1;42(5):583-98.

11. Yuan Q, Zhu X, Sayre LM. Chemical nature of stochastic generation of protein-based carbonyls: Metal-catalyzed oxidation versus modification by products of lipid oxidation. Chem Res Toxicol. 2007 Jan;20(1):129-39.

12. Grimsrud PA, Xie H, Griffin TJ, Bernlohr DA. Oxidative stress and covalent modification of protein with bioactive aldehydes. J Biol Chem. 2008 Aug 8;283(32):21837-41.

13. Fritz KS, Petersen DR. Exploring the biology of lipid peroxidation-derived protein carbonylation. Chem Res Toxicol. 2011 Sep 19;24(9):1411-9.

14. Ullery JC, Marnett LJ. Protein modification by oxidized phospholipids and hydrolytically released lipid electrophiles: Investigating cellular responses. Biochim Biophys Acta. 2012 Oct;1818(10):2424-35.

15. Van Kuijk FJ, Holte LL, Dratz EA. 4-hydroxyhexenal: A lipid peroxidation product derived from oxidized docosahexaenoic acid. Biochim Biophys Acta. 1990 Mar 12;1043(1):116-8.

16. Esterbauer H, Schaur RJ, Zollner H. Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes. Free Radic Biol Med. 1991;11(1):81-128.

17. Uchida K. Role of reactive aldehyde in cardiovascular diseases. Free Radic Biol Med. 2000 Jun 15;28(12):1685-96.

18. Ichihashi K, Osawa T, Toyokuni S, Uchida K. Endogenous formation of protein adducts with carcinogenic aldehydes: Implications for oxidative stress. J Biol Chem. 2001 Jun 29;276(26):23903-13.

19. Montine TJ, Neely MD, Quinn JF, Beal MF, Markesbery WR, Roberts LJ, et al. Lipid peroxidation in aging brain and alzheimer's disease. Free Radic Biol Med. 2002 Sep 1;33(5):620-6.

20. Uchida K, Stadtman ER. Modification of histidine residues in proteins by reaction with 4-hydroxynonenal. Proc Natl Acad Sci U S A. 1992 May 15;89(10):4544-8.

21. Coyle JT, Puttfarcken P. Oxidative stress, glutamate, and neurodegenerative disorders. Science. 1993 Oct 29;262(5134):689-95.

22. Nadkarni DV, Sayre LM. Structural definition of early lysine and histidine adduction chemistry of 4-hydroxynonenal. Chem Res Toxicol. 1995 Mar;8(2):284-91.

23. Barreiro E, Hussain SN. Protein carbonylation in skeletal muscles: Impact on function. Antioxid Redox Signal. 2010 Mar;12(3):417-29.

24. Requena JR, Chao CC, Levine RL, Stadtman ER. Glutamic and aminoadipic semialdehydes are the main carbonyl products of metal-catalyzed oxidation of proteins. Proc Natl Acad Sci U S A. 2001 Jan 2;98(1):69-74.

25. Levine RL, Stadtman ER. Oxidative modification of proteins during aging. Exp Gerontol. 2001 Sep;36(9):1495-502.

26. Stadtman ER, Levine RL. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. Amino Acids. 2003 Dec;25(3-4):207-18.

27. Aldini G, Dalle-Donne I, Colombo R, Maffei Facino R, Milzani A, Carini M. Lipoxidation-derived reactive carbonyl species as potential drug targets in preventing protein carbonylation and related cellular dysfunction. ChemMedChem. 2006 Oct;1(10):1045-58.

28. Berlett BS, Stadtman ER. Protein oxidation in aging, disease, and oxidative stress. J Biol Chem. 1997 Aug 15;272(33):20313-6.

29. Dalle-Donne I, Rossi R, Giustarini D, Milzani A, Colombo R. Protein carbonyl groups as biomarkers of oxidative stress. Clin Chim Acta. 2003 Mar;329(1-2):23-38.

30. Madian AG, Regnier FE. Proteomic identification of carbonylated proteins and their oxidation sites. J Proteome Res. 2010 Aug 6;9(8):3766-80.

31. Heal WP, Tate EW. Getting a chemical handle on protein post-translational modification. Org Biomol Chem. 2010 Feb 21;8(4):731-8.

32. Maier CS, Chavez J, Wang J, Wu J. Protein adducts of aldehydic lipid peroxidation products identification and characterization of protein adducts using an aldehyde/keto-reactive probe in combination with mass spectrometry. Methods Enzymol. 2010;473:305-30.

33. Rauniyar N, Prokai L. Isotope-coded dimethyl tagging for differential quantification of posttranslational protein carboynylation by 4-hydroxy-2-nonenal, an end-product of lipid peroxidation. J Mass Spectrom. 2011.

34. Han B, Hare M, Wickramasekara S, Fang Y, Maier CS. A comparative 'bottom up' proteomics strategy for the site-specific identification and quantification of protein modifications by electrophilic lipids. J Proteomics. 2012 Oct 22;75(18):5724-33.

35. Salmon AB, Richardson A, Perez VI. Update on the oxidative stress theory of aging: Does oxidative stress play a role in aging or healthy aging? Free Radic Biol Med. 2010 Mar 1;48(5):642-55.

36. Jacobs AT, Marnett LJ. Systems analysis of protein modification and cellular responses induced by electrophile stress. Acc Chem Res. 2010 May 18;43(5):673-83.

37. Guo J, Prokai-Tatrai K, Ngyuen V, Rauniyar N, Ughy B, Prokai L. Protein targets for carbonylation by 4-hydroxy-2-nonenal in rat liver mitochondria. J Proteomics. 2011 Jul 23.

38. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR,3rd. Protein analysis by shotgun/bottom-up proteomics. Chem Rev. 2013 Feb 26.

39. Eng JK, McCormack AL, Yates JR. 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.

40. Ideker T, Galitski T, Hood L. A new approach to decoding life: Systems biology [Internet]; 2001 [cited 24 March 2013].

41. Zhang W, Li F, Nie L. Integrating multiple 'omics' analysis for microbial biology: Application and methodologies. Microbiology. 2010;156(2):287-301.

42. McKenna NJ. Discovery-driven research and bioinformatics in nuclear receptor and coregulator signaling. Biochimica et Biophysica Acta - Molecular Basis of Disease. 2011;1812(8):808-17.

43. Grunwald DJ, Eisen JS. Headwaters of the zebrafish -- emergence of a new model vertebrate. Nat Rev Genet. 2002 Sep;3(9):717-24.

44. Shin JT, Fishman MC. From zebrafish to human: Modular medical models. Annu Rev Genomics Hum Genet. 2002;3:311-40.

45. Goldsmith P, Harris WA. The zebrafish as a tool for understanding the biology of visual disorders. Semin Cell Dev Biol. 2003 Feb;14(1):11-8.

46. Parng C. In vivo zebrafish assays for toxicity testing. Curr Opin Drug Discov Devel. 2005 Jan;8(1):100-6.

47. Parng C, Ton C, Lin YX, Roy NM, McGrath P. A zebrafish assay for identifying neuroprotectants in vivo. Neurotoxicol Teratol. 2006 Jul-Aug;28(4):509-16.

48. Wu M, Xu H, Shen Y, Qiu W, Yang M. Oxidative stress in zebrafish embryos induced by short-term exposure to bisphenol A, nonylphenol, and their mixture. Environ Toxicol Chem. 2011 Oct;30(10):2335-41.

49. Westerfield M. The zebrafish book. A guide for the laboratory use of zebrafish (danio rerio). 4th ed. Eugene, OR: University of Oregon Press; 2000.

50. Whole mount staining protocols: Zebrafish whole mount immunohistochemistry protocol. [Internet].: AbCam®; 2013; cited June 2012]. Available from: http://www.abcam.com/index.html?pageconfig=resource&rid=11330.

51. Rauniyar N, Stevens SM, Prokai-Tatrai K, Prokai L. Characterization of 4-hydroxy-2-nonenal-modified peptides by liquid chromatography-tandem mass spectrometry using data-dependent acquisition: Neutral loss-driven MS3 versus neutral loss-driven electron capture dissociation. Anal Chem. 2009 Jan 15;81(2):782-9.

52. Stevens SM,Jr, Rauniyar N, Prokai L. Rapid characterization of covalent modifications to rat brain mitochondrial proteins after ex vivo exposure to 4-hydroxy-2-nonenal by liquid chromatography-tandem mass spectrometry using data-dependent and neutral loss-driven MS3 acquisition. J Mass Spectrom. 2007 Dec;42(12):1599-605.

53. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. 2002 Oct 15;74(20):5383-92.

54. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. 2003 Sep 1;75(17):4646-58.

55. Data upload definitions [Internet].: IPA®; 2010; cited January2013]. Available from: http://ingenuity.force.com/ipa/IPATutorials?id=kA250000000TNBaCAO.

56. Suzuki D, Miyata T, Saotome N, Horie K, Inagi R, Yasuda Y, et al. Immunohistochemical evidence for an increased oxidative stress and carbonyl modification of proteins in diabetic glomerular lesions. J Am Soc Nephrol. 1999 Apr;10(4):822-32.

57. Roe MR, Xie H, Bandhakavi S, Griffin TJ. Proteomic mapping of 4-hydroxynonenal protein modification sites by solid-phase hydrazide chemistry and mass spectrometry. Anal Chem. 2007 May 15;79(10):3747-56.

58. Bachi A, Dalle-Donne I, Scaloni A. Redox proteomics: Chemical principles, methodological approaches and biological/biomedical promises. Chem Rev. 2013 Jan 9;113(1):596-698.

59. Wang Q, Zhao X, He S, Liu Y, An M, Ji J. Differential proteomics analysis of specific carbonylated proteins in the temporal cortex of aged rats: The deterioration of antioxidant system. Neurochem Res. 2010 Jan;35(1):13-21.

60. Higdon A, Diers AR, Oh JY, Landar A, Darley-Usmar VM. Cell signalling by reactive lipid species: New concepts and molecular mechanisms. Biochem J. 2012 Mar 15;442(3):453-64.

61. Butterfield DA, Dalle-Donne I. Redox proteomics. Antioxid Redox Signal. 2012 Dec 1;17(11):1487-9.

62. Alestrom P, Holter JL, Nourizadeh-Lillabadi R. Zebrafish in functional genomics and aquatic biomedicine. Trends Biotechnol. 2006 Jan;24(1):15-21.

63. Chan TM, Longabaugh W, Bolouri H, Chen HL, Tseng WF, Chao CH, et al. Developmental gene regulatory networks in the zebrafish embryo. Biochim Biophys Acta. 2009 Apr;1789(4):279-98.

64. Barrera G, Pizzimenti S, Dianzani MU. Lipid peroxidation: Control of cell proliferation, cell differentiation and cell death. Mol Aspects Med. 2008;29(1-2):1-8.

65. Pizzimenti S, Toaldo C, Pettazzoni P, Dianzani MU, Barrera G. The "two-faced" effects of reactive oxygen species and the lipid peroxidation product 4-hydroxynonenal in the hallmarks of cancer. Cancers. 2010;2(2):338-63.

66. Wong CM, Marcocci L, Liu L, Suzuki YJ. Cell signaling by protein carbonylation and decarbonylation. Antioxid Redox Signal. 2010 Mar;12(3):393-404.

67. Pillon NJ, Croze ML, Vella RE, Soulère L, Lagarde M, Soulage CO. The lipid peroxidation by-product 4-hydroxy-2-nonenal (4-HNE) induces insulin resistance in skeletal muscle through both carbonyl and oxidative stress. Endocrinology. 2012;153(5):2099-111.

68. Ingram KH, Hill H, Moellering DR, Hill BG, Lara-Castro C, Newcomer B, et al. Skeletal muscle lipid peroxidation and insulin resistance in humans. J Clin Endocrinol Metab. 2012;97(7):E1182-6.

69. Hussain SP, Hofseth LJ, Harris CC. Radical causes of cancer. Nature Reviews Cancer. 2003;3(4):276-85.

70. Coussens LM, Werb Z. Inflammation and cancer. Nature. 2002;420(6917):860-7.

71. Reuter S, Gupta SC, Chaturvedi MM, Aggarwal BB. Oxidative stress, inflammation, and cancer: How are they linked? Free Radical Biology and Medicine. 2010;49(11):1603-16.

72. Semenza GL, Jiang B-, Leung SW, Passantino R, Concordet J-, Maire P, et al. Hypoxia response elements in the aldolase A, enolase 1, and lactate dehydrogenase A gene promoters contain essential binding sites for hypoxia- inducible factor 1. J Biol Chem. 1996;271(51):32529-37.

73. Zhang D, Tai LK, Wong LL, Chiu LL, Sethi SK, Koay ES. Proteomic study reveals that proteins involved in metabolic and detoxification pathways are highly expressed in HER-2/neu-positive breast cancer. Mol Cell Proteomics. 2005 Nov;4(11):1686-96.

74. Pollard PJ, Briere JJ, Alam NA, Barwell J, Barclay E, Wortham NC, et al. Accumulation of krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations. Hum Mol Genet. 2005 Aug 1;14(15):2231-9.

75. Karve TM, Cheema AK. Small changes huge impact: The role of protein posttranslational modifications in cellular homeostasis and disease. J Amino Acids. 2011;2011:207691.

76. Neely MD, Sidell KR, Graham DG, Montine TJ. The lipid peroxidation product 4-hydroxynonenal inhibits neurite outgrowth, disrupts neuronal microtubules, and modifies cellular tubulin. J Neurochem. 1999;72(6):2323-33.

77. Kokubo J, Nagatani N, Hiroki K, Kuroiwa K, Watanabe N, Arai T. Mechanism of destruction of microtubule structures by 4-hydroxy-2-nonenal. Cell Struct Funct. 2008;33(1):51-9.

CHAPTER V

SUMMARY

The science community has come a long way in the advancement of sequenced genomes with regards to a variety of organisms (1, 2). Monitoring environmental conditions or changes in functions of organisms affords a better understanding of cellular responses upon modulation in genes, proteins, and metabolites. With the advent of 'omics' technology, research has been revolutionized in methodologies to study systems biology. The 'omics' era utilizes genomics, transcriptomics, proteomics, metabolomics, and so on (3-5). While there has been some progression in the aforementioned fields, lipids and their byproducts remain elusive in regards to their influences on biological and cellular processes. The primary objective of the studies conducted in this dissertation was to monitor the global protein expression and its dynamics during lipid stress in biological systems using discovery-driven bottom-up proteomics.

The progress of shotgun proteomics-based methods for the quantitative survey of complex proteomes is documented for the analysis of the *pgsA* mutant *Synechocystis* sp. PCC6803 strain in **Chapter II**. This genetically modified microorganism is unable to produce phosphatidylglycerol (PG). However, it is able to grow in the presence of exogenously added PG and can survive for an extended period after the phospholipid has been removed (6).

Cyanobacteria are characterized as oxygenic phototrophs capable of photosynthetic processes. Previous studies have reported PG to be the only phospholipid that is an indispensable constituent of photosynthetic complexes (7-9). Many of the underlying mechanisms of PG and its influence on the cyanobacterial proteome are unknown. Through this genetically modified model, we were able to study the importance of PG in photosynthetic and cellular processes using label-free quantitation alongside shotgun proteomics. We successfully identified key PG-regulated proteins on a global-scale in the cyanobacteria mutant strain for the first time. Mass spectrometry became an essential tool for this proteomics approach due to its ability to profile proteins in biological systems on a global-scale. Additionally, bioinformatic applications that analyzed the biological relevance of annotations in regards to associated protein IDs made it possible to reveal cellular processes influenced by this lipid regulation.

Analysis of several of the proteins conditions revealed a large dataset mostly consisting of soluble proteins alongside a handful of membrane proteins. Previous studies have suggested that PG plays a vital role for the structure and function of photosynthetic membranes (7, 10). However, in global-scale proteomics studies, membrane proteins are usually underrepresented (11). Approaches to membrane proteomics have been a challenge, and therefore supplementary steps are required to isolate and characterize these constituents. In **Chapter III**, a strategy based on membrane proteomics was designed to isolate thylakoid membrane proteins of the *Synechocystis* mutant. The efficiency of the shotgun proteomics method was tested to identify PG-regulated thylakoid membrane proteins on a global scale. We successfully identified key thylakoid membrane proteins associated with photosynthesis as well as cellular processes. In addition, label-free methods permitted the elucidation of protein expression levels for the related membrane proteins. Bioinformatic applications using cyanobacterial genome annotations

suggested that a majority of these membrane proteins were associated with photosynthesis, cellular processes, and translation under these lipid stress conditions. Our thylakoid-focused proteomic approach has demonstrated the biological importance and diverse role of this phospholipid in the processes carried out in this pivotal cellular compartment of cyanobacteria. Furthermore, it is important to note that studies involving this genetically modified strain are expected to reveal previously unexplored lipid signaling mechanisms of the process, which could be further investigated in higher plants.

Lipids are a diverse collection of biomolecules essential to all living species (12). In addition, phospholipids are located within cellular membranes and consist of an abundant assortment of unsaturated fatty acid constituents throughout their backbone structure. Fatty acids can be involved in the regulation of biological processes such as inflammatory response, neuronal signal transmission, and carbohydrate metabolism (13). Polyunsaturated fatty acids (PUFA) have been shown to be more vulnerable to peroxidation (14-16). Several reports have implicated that the byproducts of lipid peroxidation can cause damage to proteins (17). Dysfunctions in membrane proteins have been implicated as mediators of underlying mechanisms and/or causes of several human diseases (11, 18).

So far, our focus has relied on the broader role of phospholipids in a microorganism. However, the role of phospholipids is a dynamic process (19). We have seen this in the cyanobacteria model. The components of biomembranes are also variable and not fully established across different species. In addition to the above-mentioned work, we proposed to utilize the discovery-driven methods to further investigate lipid-derived byproduct 4-hydroxy-2-nonenal (HNE) and its consequences in disturbing the homeostasis of early developing zebrafish embryos in **Chapter IV**. HNE is a highly reactive α,β-unsaturated electrophilic aldehyde capable

of causing modification to amino acid residues (20-24). This type of protein modification can cause structural damage to proteins and can lead to protein malfunctions (25-27). Since this electrophile has been recognized as a broad biomarker of oxidative stress (28), it has become crucial to elucidate its consequences in association with cellular and biological processes of a system's proteome. We detected that lipid peroxidation was overwhelming the zebrafish embryos during HNE exposure. We also confirmed that HNE adducts linked to histidine residues on a handful of HNE-treated proteins, which coincides with previous studies of carbonyl modification (20, 21, 29). Utilizing label-free methods, we revealed that a large dataset of proteins that were regulated in response to the HNE treatment. The survey of this global-scale initiative was also valuable in terms of interrelating protein IDs to biological and cellular processes using bioinformatics software. As a result, HNE-regulated proteins along with HNE-modified proteins were implicated in several biological processes, as well as in the pathogenesis of diseases such as cancer, diabetes, and cytoskeletal dysfunction. These results provide an example of the potential functional consequences of lipid stress with respect to HNE-induced carbonyl stress.

FUTURE DIRECTIONS

We have implemented an extensive global-scale study utilizing a discovery-driven proteomics approach for the elucidation of lipid stress effects on the systems level in both cyanobacteria and zebrafish embryos. Exploring differential protein expression on large datasets, while screening the entire genome for proteins that interact with specific regulatory dynamics,

has provided insights into pivotal cellular and biological processes required to maintain homeostasis in both of these organisms.

At this point, any additional studies towards understanding the photosynthetic processes in regards to PG-regulated membrane proteins will require further critical review and interpretation by specialists of cyanobacteria. However, a few things to note in preparation for follow-up studies of photosynthetic processes and associated membrane proteins include:

(1) Membrane proteomics remains a challenge for many labs (11, 30). Although we were able to identify membrane constituents, our samples were also "contaminated" by large amounts of soluble proteins. This suggests that isolation and enrichment techniques are not perfect, and alternative approaches may be required to isolate membrane components.

(2) Since we used enrichment/isolation of the membrane proteins, soluble proteins identified in these datasets cannot necessarily be considered as true targets of the PG-depletion. Further studies should be performed in order to determine the roles these proteins play in photosynthetic processes.

(3) After bioinformatic analysis, several membrane proteins were categorized as hypothetical, indicating the need for more refined methods to isolate and characterize this part of the proteome.

We have successfully profiled the proteomes of cyanobacteria and zebrafish under their respective lipid conditions. We have highlighted a grander scheme of what is occurring systematically under lipid stress. Furthermore, after identifying the protein expression levels from quantitative analyses using control *versus* treated samples, we now have an opportunity to formulate new hypotheses of why and how these targeted species are being modulated under these conditions. Specifically, we have displayed several pathways and networks involved in

zebrafish development by way of HNE exposure. HNE-regulated and/or HNE-modified proteins involved in these processes may be potential biomarkers of disease. Another mass spectrometry technique to consider in validating biomarker candidates is referred to as "targeted proteomics". In this approach, experiments are created to monitor only a select few proteins of interest with high throughput capabilities that mass spectrometry provides. This strategy has the potential to provide greater sensitivity, and allows for detection of lower abundance protein/peptide candidates (31). In the case of the zebrafish proteome, many proteins were implicated in several pathways and processes that were linked to human diseases such as cancer and diabetes. Likewise, several cytoskeletal proteins were shown to be both regulated and modified by HNE treatment. Therefore, applying a targeted proteomics strategy for these selected proteins might provide insight into whether or not these specific proteins may be candidate biomarkers of cytoskeletal diseases such as muscular dysfunctions. Targeted proteomics strategies could also be implemented in cyanobacterial research to validate target membrane proteins revealed in PG-depleted mutants.

Along with targeted proteomics, we can use automated docking software (i.e. BioMedCAChe(TM)) to determine the structural and functional properties of the HNE-regulated proteins. Furthermore, we can determine whether sites of HNE modification cause conformational changes in structure and/or lead to protein dysfunction. These types of computer-modeling systems may be beneficial for approaches towards drug discovery and delivery.

In essence, the approach of discovery-based proteomics is to identify as many of the proteins in a proteome as possible. The challenge with this is that the proteome is very dynamic and large amounts of data are acquired. Interpreting the data and its biological relevance can be difficult. Therefore, integrating multiple "omics" analyses will allow for an in-depth

understanding of the biochemical behavior in processes discovered during lipid stress. It can be concluded that these strategies are complementary, and provide an improved route to deciphering lipid stress in a systems approach.

REFERENCES

1. Kaneko T, Tabata S. Complete genome structure of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. Plant and Cell Physiology. 1997;38(11):1171-6.

2. Zebrafish genome overview [Internet].: Genome Reference Consortium; 2010 [updated Oct. 5, 2010; cited January 2013]. Available from: http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/zebrafish/.

3. Fiehn O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. Comp Funct Genomics. 2001;2(3):155-68.

4. Singh OV, Nagaraj NS. Transcriptomics, proteomics and interactomics: Unique approaches to track the insights of bioremediation. Briefings in Functional Genomics and Proteomics. 2006;4(4):355-62.

5. Kandpal RP, Saviola B, Felton J. The era of 'omics unlimited. BioTechniques. 2009;46(5 SPEC. ISSUE):351-5.

6. Sakurai I, Hagio M, Gombos Z, Tyystjarvi T, Paakkarinen V, Aro EM, et al. Requirement of phosphatidylglycerol for maintenance of photosynthetic machinery. Plant Physiol. 2003 Nov;133(3):1376-84.

7. Gombos Z, Várkonyi Z, Hagio M, Iwaki M, Kovács L, Masamoto K, et al. Phosphatidylglycerol requirement for the function of electron acceptor plastoquinone QB in the photosystem II reaction center. Biochemistry (N Y). 2002;41(11):3796-802.

8. Wada H, Murata N. The essential role of phosphatidylglycerol in photosynthesis. Photosynth Res. 2007 May;92(2):205-15.

9. Bogos B, Ughy B, Domonkos I, Laczko-Dobos H, Komenda J, Abasova L, et al. Phosphatidylglycerol depletion affects photosystem II activity in *Synechococcus* sp. PCC 7942 cells. Photosynth Res. 2010 Jan;103(1):19-30.

10. Sakurai I, Shen JR, Leng J, Ohashi S, Kobayashi M, Wada H. Lipids in oxygen-evolving photosystem II complexes of cyanobacteria and higher plants. J Biochem. 2006 Aug;140(2):201-9.

11. Vuckovic D, Dagley LF, Purcell AW, Emili A. Membrane proteomics by high performance liquid chromatography-tandem mass spectrometry: Analytical approaches and challenges. Proteomics. 2013;13(3-4):404-23.

12. Murphy RC, Axelsen PH. Mass spectrometric analysis of long-chain lipids. Mass Spectrom Rev. 2011 Jul-Aug;30(4):579-99.

13. Subramaniam S, Fahy E, Gupta S, Sud M, Byrnes RW, Cotter D, et al. Bioinformatics and systems biology of the lipidome. Chem Rev. 2011 Oct 12;111(10):6452-90.

14. Rael LT, Thomas GW, Craun ML, Curtis CG, Bar-Or R, Bar-Or D. Lipid peroxidation and the thiobarbituric acid assay: Standardization of the assay when using saturated and unsaturated fatty acids. Journal of Biochemistry and Molecular Biology. 2004;37(6):749-52.

15. Tvrzicka E, Kremmyda LS, Stankova B, Zak A. Fatty acids as biocompounds: Their role in human metabolism, health and disease--a review. part 1: Classification, dietary sources and biological functions. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2011 Jun;155(2):117-30.

16. Pratt DA, Tallman KA, Porter NA. Free radical oxidation of polyunsaturated lipids: New mechanistic insights and the development of peroxyl radical clocks. Acc Chem Res. 2011;44(6):458-67.

17. Butterfield DA, Dalle-Donne I. Redox proteomics. Antioxid Redox Signal. 2012 Dec 1;17(11):1487-9.

18. Poli G, Schaur RJ, Siems WG, Leonarduzzi G. 4-hydroxynonenal: A membrane lipid oxidation product of medicinal interest. Med Res Rev. 2008;28(4):569-631.

19. Takahashi-Umebayashi M, Pineau L, Hannich T, Zumbuehl A, Doval DA, Matile S, et al. Chemical biology approaches to membrane homeostasis and function. Chimia. 2011;65(11):849-52.

20. Uchida K, Stadtman ER. Modification of histidine residues in proteins by reaction with 4-hydroxynonenal. Proc Natl Acad Sci U S A. 1992 May 15;89(10):4544-8.

21. Nadkarni DV, Sayre LM. Structural definition of early lysine and histidine adduction chemistry of 4-hydroxynonenal. Chem Res Toxicol. 1995 Mar;8(2):284-91.

22. Sayre LM, Lin D, Yuan Q, Zhu X, Tang X. Protein adducts generated from products of lipid oxidation: Focus on HNE and one. Drug Metab Rev. 2006;38(4):651-75.

23. Yuan Q, Zhu X, Sayre LM. Chemical nature of stochastic generation of protein-based carbonyls: Metal-catalyzed oxidation versus modification by products of lipid oxidation. Chem Res Toxicol. 2007 Jan;20(1):129-39.

24. Grimsrud PA, Xie H, Griffin TJ, Bernlohr DA. Oxidative stress and covalent modification of protein with bioactive aldehydes. J Biol Chem. 2008 Aug 8;283(32):21837-41.

25. Yan LJ, Sohal RS. Analysis of oxidative modification of proteins. Curr Protoc Cell Biol. 2002 May;Chapter 7:Unit 7.9.

26. Headlam HA, Davies MJ. Markers of protein oxidation: Different oxidants give rise to variable yields of bound and released carbonyl products. Free Radic Biol Med. 2004 May 1;36(9):1175-84.

27. Aldini G, Dalle-Donne I, Colombo R, Maffei Facino R, Milzani A, Carini M. Lipoxidation-derived reactive carbonyl species as potential drug targets in preventing protein carbonylation and related cellular dysfunction. ChemMedChem. 2006 Oct;1(10):1045-58.

28. Dalle-Donne I, Rossi R, Giustarini D, Milzani A, Colombo R. Protein carbonyl groups as biomarkers of oxidative stress. Clin Chim Acta. 2003 Mar;329(1-2):23-38.

29. Guo J, Prokai-Tatrai K, Ngyuen V, Rauniyar N, Ughy B, Prokai L. Protein targets for carbonylation by 4-hydroxy-2-nonenal in rat liver mitochondria. J Proteomics. 2011 Jul 23.

30. Speers AE, Wu CC. Proteomics of integral membrane proteins - theory and application. Chem Rev. 2007;107(8):3687-714.

31. Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, Robinson EW, et al. Mass spectrometry-based proteomics: Existing capabilities and future directions. Chem Soc Rev. 2012;41(10):3912-28.