May, 2016
University of Nevada, Reno

Sugar-Coating Metabolism: A Comparative Analysis of Global Metabolite Profiles in *C. elegans* O-GlcNAcylation Mutants

A thesis submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Science in Biology and the Honors Program

by

Austin J. Marett

Dr. Patricia Berninsone, Ph.D., Thesis Advisor

May, 2016

UNIVERSITY OF NEVADA RENO

THE HONORS PROGRAM

We recommend that the thesis
prepared under our supervision by

AUSTIN J. MARETT

entitled

Sugar-Coating Metabolism: A Comparative Analysis of Global Metabolite Profiles in *C. elegans* O-GlcNAcylation Mutants

be accepted in partial fulfillment of the
requirements for the degree of

BACHELOR OF SCIENCE, BIOLOGY

_____

Patricia Berninsone, Ph.D., Thesis Advisor

_____

Tamara Valentine, Ph. D., Director, **Honors Program**

1. Abstract

In the field of molecular biology, it is becoming increasingly clear that the use of gene knock-out experiments alone is insufficient to characterize the functions of the majority of genes. Approximately 85-90% of the genes in most organisms are ‘silent’ and cannot be explored using these traditional methods. Silent genes are particularly important in the study of posttranslational modifications of proteins, which occurs completely independently of any transcription, gene splicing, or translation. Protein modification is known to be vital to the regulation and control of enzymes, including those involved in the metabolism of sugars. In this study, I examine a specific form of cyclical protein glycosylation, O-GlcNAcylation, in the model organism *Caenorhabditis elegans* using a metabolomics approach. By comparing global metabolite profiles of wild type N2 *C. elegans* and two knock-out mutant strains *oga-1* and *ogt-1*, which respectively possess deletions for the OGA and OGT enzymes involved in O-GlcNAcylation, I demonstrate that interruption of O-GlcNAcylation produces predictable alterations to core metabolites involved in glucose metabolism and the nutrient-sensing hexosamine biosynthetic pathway (HBP). These results suggest that O-GlcNAcylation is directly dependent on, and impacted by, several other fundamental cellular pathways and metabolites. Once identified, such metabolic systems and their intermediates may serve as therapeutic targets in the treatment of diseases that also perturb the same pathways, such as diabetes, neurodegenerative diseases, and various cancers.

Keywords: Metabolomics; *Caenorhabditis elegans*; Post-translational modifications; Protein glycosylation; O-GlcNAcylation; UDP-GlcNAc; OGA; OGT.

2. Acknowledgements

The completion of this thesis is thanks to the efforts of my advisor Dr. Patricia Berninsone and the contributions of the University of Nevada, Reno, the Biology Department, and the Honors Program. This research was also made possible with funding from the Nevada INBRE program and the Honors Undergraduate Research Award. Finally, I would like to thank my friends, family, and lab coworkers, whose support and guidance have been irreplaceable in my academic career.

3. Table of Contents

List of Figures

5. Introduction

When researchers examine disease models and their progression within an organism, one of their objectives is to identify what changes are occurring at the cellular level. However, determining which pathways are altered within a cell, and to what degree those pathways are altered, is extremely complex and often imprecise using conventional methods, making researching diseases with widespread effects, such as diabetes, obesity, or various cancers, very difficult for molecular biologists. While a cell may respond to a disease or stressor by altering any of its many different biochemical pathways, organelles, or signaling cascades, the intermediate molecules involved in metabolism and catabolism (the consumption and production of a cell's 'building blocks') are nearly always impacted. These highly interactive classes of molecules are collectively termed *metabolites*, and the quantitative and qualitative study of how these metabolites change depending on cellular conditions is known as *metabolomics*.

For decades, research on cures and treatment methods for most diseases has hinged on experimentation using model organisms. For research involving genomics, development, signal transduction, neurobiology, aging, and more recently RNAi, one frequently used model organism is the nematode *Caenorhabditis elegans*, which is desirable for its small size, transparent body, small genome, durability, and ease of care (Antoshechkin and Sternberg, 2007). Originally, one of the primary methods of examining the functions of genes and their biological significance using *C. elegans* was by selectively 'knocking-out', or deleting, a desired gene and then observing the mutant nematode strain for any changes in behavior, development, reproduction, or morphology. This practice of studying a mutant strain's appearance, or 'phenotype', is useful for

identifying very noticeable, broad spectrum changes resulting from the deletion of a gene. However, prior research has also shown that approximately 85-90% of all functioning genes in both *C. elegans* and in mammals produce much more subtle effects that are not visibly apparent or quantifiable (Blaise et al., 2007). With respect to *C. elegans*, which has an estimated 20,000 genes, this research suggests that many thousands of genes are impossible to accurately study using traditional knock-out methods (Antoshechkin and Sternberg, 2007).

In recent years, significant progress has been made in the development of new analytical techniques for identifying and quantifying cellular metabolites, including nuclear magnetic resonance spectroscopy (NMR), gas chromatography (GC), liquid chromatography (LC), and mass spectroscopy (MS) (Kolker et al., 2011). These techniques have varying strengths and weaknesses when analyzing different molecules, and are capable of quantifying most metabolites present in biological samples when used in tandem and cross-referenced with one another. These techniques are also nonspecific; they do not require molecules to be labeled or otherwise identified prior to the experiment. Because of their nonspecific nature, researchers are capable of using NMR, GC, LC, and MS to construct profiles of the metabolites present in biological samples and their concentrations. These profiles can then be used to elucidate the metabolic differences between different tissue samples, such as those from mutant *C. elegans* strains, that were previously impossible to effectively quantify.

One avenue of research that has recently gained popularity in molecular biology is the study of *posttranslational modifications* (PTMs). Once translated, proteins (the workhorses of the cell) adopt their appropriate structure and function after undergoing

multiple modifications and folding events, processes collectively referred to as PTMs (Farley and Link, 2009). While protein sequences are encoded by genes, PTMs are what give rise to the overwhelming amount of protein diversity and are fundamental to proper protein activity and cellular metabolism (Kamath et al., 2011). One form of PTM is the addition and removal of carbohydrates to proteins, otherwise known as *protein glycosylation*. There are several subtypes of protein glycosylation, and one of the most common is known as *O-GlcNAcylation*, which is the rapid addition and removal of *UDP-GlcNAc*, a glucose derivative, to proteins (Ghosh et al., 2014). A large proportion of proteins both in the cytosol and the nucleus of the cell have their activity, stability, and interactions regulated by O-GlcNAcylation, making O-GlcNAcylation one of the most globally influential forms of protein glycosylation. However, knock-out mutants for genes involved in O-GlcNAcylation are lethal in mammals, making mammalian animal models unfit for experimentation and the diverse effects of cyclic O-GlcNAcylation poorly understood as a result (Farley and Link, 2009).

The cyclical process of O-GlcNAcylation operates using the two enzymes *OGA* (O-GlcNAcase, which removes O-GlcNAc from proteins) and *OGT* (O-GlcNAc transferase, which adds O-GlcNAc to proteins). While deletions of either of these enzymes in mammalian models are universally nonviable, mutants for this form of glycosylation do exist in the much simpler nematode model organism *C. elegans*. Two mutant strains, *oga-1* and *ogt-1*, possess deletions for the OGA and OGT enzymes respectively, yet unlike their mammalian counterparts they surprisingly suffer minimal phenotypic consequences (Ghosh et al., 2014). With such dramatic differences in the apparent effects of interfering with O-GlcNAcylation between mammals and *C. elegans*,

the question then becomes how does O-GlcNAcylation impact cellular metabolism and its associated pathways? Because O-GlcNAcylation is directly dependent on carbohydrates and other metabolites, I propose that using a metabolomics approach to compare N2 (wild type) strains of *C. elegans* to *oga-1* and *ogt-1* strains will reveal significant differences in concentrations of several metabolites, including various carbohydrates, lipids, and amino acids. These metabolic changes may then explain how this type of protein glycosylation regulates global cellular metabolism. With a formal understanding of the processes most closely related to O-GlcNAcylation and how O-GlcNAcylation impacts metabolite concentrations, a foundation can be laid for the development of possible treatments and interventions for diseases that produce similar aberrations in cellular metabolism, such as diabetes and cancer.

A significant body of prior research, analytical tools, and published findings was necessary to establish the background for this subject. These resources and prior knowledge will now be reviewed, and their implications addressed with regard to the justification and methodology of this study.

6. Literature Review

6.1 Proteins and PTMs

In the context of cell biology, proteins are a class of macromolecules that are responsible for carrying out essentially all of the active processes and activities within the cell. A large study by Kamath et al. (2011) examined protein diversity in multiple organisms and demonstrated that proteins possess a degree of diversity and complexity far beyond other macromolecules that matches their enormous range of functions. In fact, it was estimated that there are three degrees of magnitude more proteins than there are genes in most organisms, meaning humans (who possess approximately 30,000 genes) utilize *millions* of different specialized proteins. This dramatic increase in diversity from genes to proteins, as described by Kamath et al., arises from two factors: alternative splicing of genes and posttranslational modifications (PTMs). Kamath et al. noted that PTMs are responsible for dictating the localization, folding, and life span of proteins, and that every biological protein undergoes a PTM at some point. Most forms of PTMs are also highly conserved across all animal species, supporting the significance of PTMs in an evolutionary context. Finally, PTMs were also shown to play a crucial role in modifying protein activity during stressful cellular conditions, such as those induced by infection or disease (Kamath et al., 2011).

Due to the ubiquitous impact of PTMs on all cellular proteins across all animals, the effects of PTMs on protein function are directly related to the activity and metabolism of the cell itself. A review by Farley and Link (2009) examined the range of PTMs and their activities within cell biology. Their findings confirmed that PTMs occurred

predominantly after proteins were translated from mRNA, and involved the covalent bonding of new molecules to proteins by enzyme activity. These covalent alterations were found to significantly alter the properties of proteins, both physically and chemically, and were directly responsible for determining the activities of proteins, where they localized within the cell, and how stable they were. Some of these PTMs were permanent and necessary for protein maturation, while others were transient and used to regulate functioning or signaling (such as with O-GlcNAcylation). Farley and Link (2009) did demonstrate, however, that nearly *every* protein underwent some form of PTM, and had catalogued over 400 unique forms of PTM.

The overwhelming diversity and significance of PTMs illustrate their importance not only in understanding protein structure and functioning, but also in understanding the intricacies of cell biology as a whole. Prior research has often depended on profiling proteins with respect to their progenitor genes (Blaise et al, 2007); however, more recent studies have routinely demonstrated that considering genetics alone fails to address the majority of protein diversity and entirely ignores the impact of PTMs (Fuchs et al., 2010; Kamath et al., 2011). Yet prior research on the variety of known PTMs has also revealed several obstacles to understanding these processes. One difficulty is the simple fact that examining PTMs by measuring biological metabolites is greatly confounded by the huge variety of genes, proteins, and PTMs, which far exceed the number of observable metabolites (Raamsdonk et al., 2001). Another difficulty is the use of transfection of animal proteins into bacterial models. One study by Salom et al. attempted to express mammalian proteins in bacteria and found that PTMs are not always recreated perfectly in some models, and that there are often differences in folding or missing normal

modifications. However, their study also showed that insertion of the same protein genes into *C. elegans* produced nearly identical protein presentation and PTM action with minimal difficulty (Salom et al., 2014). These findings support the hypothesis that metabolomics analysis of PTMs in the model organism *C. elegans* is both viable and effective.

6.2 Protein Glycosylation and O-GlcNAcylation

While Farley and Link (2009) reported there to be over 400 unique forms of PTM, many of these PTMs can be grouped together by their overall types of protein modification. One of the more common subtypes of PTM is known as protein glycosylation, which involves either permanent or transient addition of carbohydrates to proteins. Figure 1 shows how binding of these ‘sugars’ can happen in numerous locations on a protein depending on the type of glycosylation, with O-linked glycosylation binding serine and threonine residues and N-linked glycosylation binding asparagine residues. Prior research examining this type of PTM has also shown that protein glycosylation is highly complex and often involves multiple steps whereas most other PTMs are much simpler and often require only a single step (Kamath et al., 2011). Ultimately, these modifications have significant effects on the activity, localization, stability, signaling, and interactions of proteins within the cell.

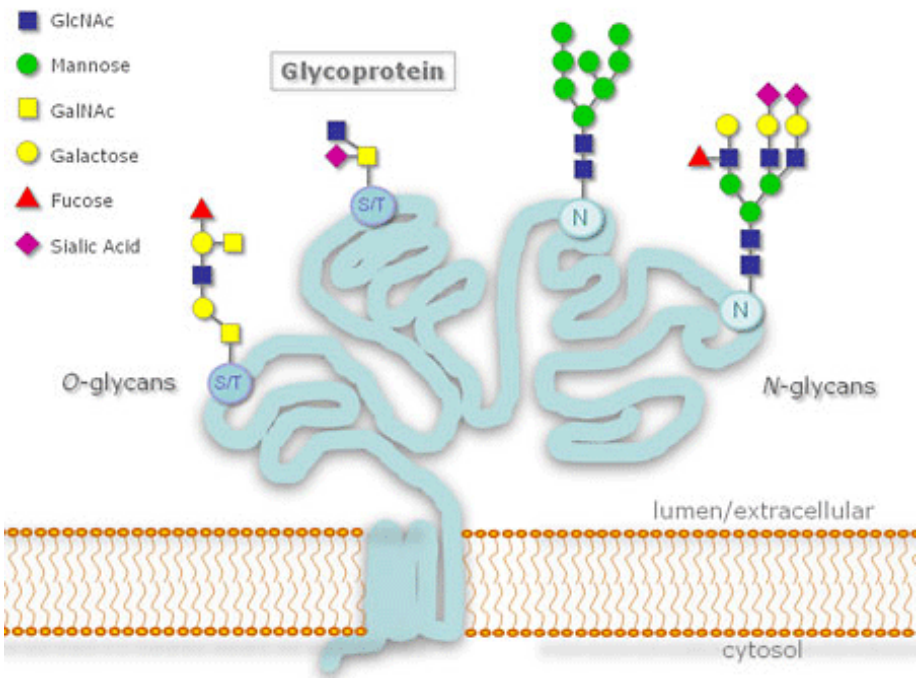

Figure 1: Demonstration of a protein glycosylated with carbohydrates. A single protein may undergo several glycosylation events and have multiple carbohydrates bound to its residues. Glycosylations occurs in multiple steps and is often highly complex (New England Biolabs, 2016).

It is important to understand that the activity of protein glycosylation and other PTMs directly affects, and is altered by, the relative concentrations of the substrate molecules they consume. Because protein glycosylation alters proteins with different modified carbohydrates, it is necessarily dependent upon the levels of those carbohydrates present in the cell, as well as upon any other intermediates and metabolites that utilize or produce those same carbohydrates. This phenomenon was demonstrated in a study by McElwee et al. (2006), who demonstrated that strains of *C. elegans* that were long lived had altered levels of transcription and protein glycosylation, which by

extension altered the concentrations of related metabolites (in their case, the lipids and saccharides involved in the citric acid cycle). This same cascading effect is likely to occur with alterations of any PTMs, and provides important information in metabolomics studies such as this.

Within the category of protein glycosylation, there exist additional subtypes of PTMs. One such subtype, known as O-GlcNAcylation, is relatively common and ubiquitous throughout the cell. As described by Farley and Link (2009), O-GlcNAcylation is a form of protein glycosylation that modifies serine and threonine residues of proteins with O-linked β-N-acetylglucosamine (O-GlcNAc). As diagramed by Figure 2, this molecule is the bound form of UDP-GlcNAc, a nucleotide sugar derived from glucose through the hexosamine biosynthetic pathway (HBP). A unique feature of O-GlcNAcylation is that it is highly dynamic, with O-GlcNAc being rapidly added to proteins by the enzyme OGT and removed by the enzyme OGA. Farley and Link also discussed in their study the similarities between O-GlcNAcylation and phosphorylation, another PMT that is dynamic and even more prevalent in protein activity regulation. They theorized that OGA and OGT may function as redundant protein activation enzymes, or even possibly as antagonists to one another (Farley and Link, 2009). While these hypotheses have yet to be more thoroughly studied, this prior research serves as a useful tool to understanding the significance of O-GlcNAcylation by comparing it to phosphorylation, another common method of signal transduction and regulation used by the cell.

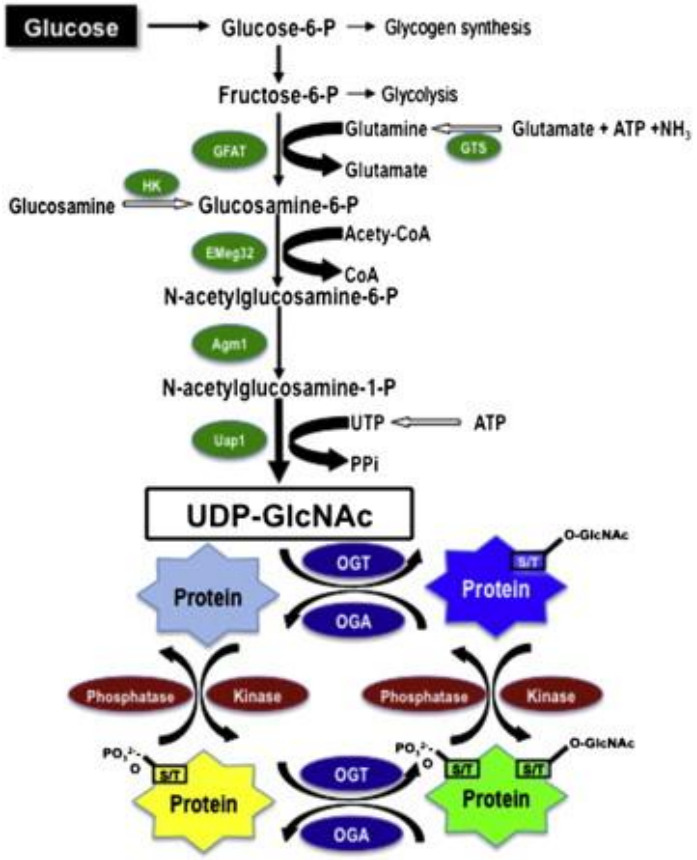

Figure 2: Diagram of the Hexosamine Biosynthetic Pathway (HBP). Several key metabolites serve as intermediates in the HBP, including glucose, Acetyl-CoA, and UTP. The concentrations of these metabolites serve as markers of the HBP's activity. The final product, UDP-GlcNAc, is the primary substrate for O-GlcNAcylation and will be added to proteins by OGT or removed by OGA. Also illustrated is the addition and removal of phosphate through phosphorylation, a highly analogous form of protein regulation (Darley-Usmar et al., 2011).

The scope of O-GlcNAcylation's effects on cellular functioning is so expansive that it is reasonable to see how interrupting its cycle in complex organisms is almost

always nonviable. One study by Ghosh and colleagues demonstrated that O-GlcNAcylation modulates a wealth of cellular processes by modifying more than 4,000 unique proteins (Ghosh et al., 2014). Their study also evaluated the impact of interferences with O-GlcNAcylation on the concentrations of closely related metabolites, such as glucose, L-glutamine, UTP, and acetyl-CoA, all of which are intermediates of the HBP and are required to synthesize UDP-GlcNAc in the cell. Their results had two particularly remarkable findings: first, that dramatically *decreased* levels of UDP-GlcNAc are extremely damaging to mammalian cells, and second, that *increasing* levels of UDP-GlcNAc results in increased levels of enzymes related to HBP functioning. The former result could possibly explain the lethality of knocking-out O-GlcNAcylation, and the latter could demonstrate that UDP-GlcNAc actually alters gene expression by acting like a 'sensor' for cellular nutrition. As Ghosh et al. (2014) explained, this finding is critical as a keystone for the metabolic interactions between different metabolic pathways, and suggests that interrupting O-GlcNAcylation may directly alter the HBP and the 'pool' of available nucleotide sugars in the cell.

The metabolic consequences of interfering with O-GlcNAcylation have also been tightly correlated with glucose metabolism and energy storage. Previous studies by Rahman et al. (2010) and Hanover et al. (2005) both demonstrated that O-GlcNAcylation was a key modulator of glycogen formation and lipid storage. Differences in energy commitment between glycogen and fat storage were correlated to the insulin-like *daf-2* receptor gene, implicating O-GlcNAcylation as a possible contributor to insulin-independent diabetes. This hypothesis was further supported by the findings of Forsythe et al. (2006), which showed that O-GlcNAcylation activity directly altered glycogen

synthase kinase levels and acted as a ‘fine-tuning’ mechanism for insulin-like signaling. This direct correlation between perturbed O-GlcNAcylation and diabetes-like altered metabolism can also be translated to other complex diseases that also demonstrate modified HBP activity and lipid synthesis, such as neurodegenerative disorders and cancers. Further research that explores O-GlcNAcylation and methods of modifying it may therefore provide new techniques for treating these diseases.

6.3 Metabolomics as an Emerging Field of Study

As described before, one the greatest limitations to studying the interactions between metabolic processes and protein modifications has been the limitations of the traditional gene-focused approach to examining knock-out mutants. Because there are far fewer metabolites to measure than there are genes and proteins to experimentally manipulate, the effects of gene deletions or PTM alterations can often be so minute that they appear ‘silent’, or are compensated for by other metabolic pathways (Raamsdonk et al., 2001). Studies show, however, that while these ‘silent’ genes may be impractical to evaluate phenotypically, they often undergo significant changes in the concentrations of their metabolites, which when profiled can provide researchers with a ‘fingerprint’ or ‘signature’ for that gene (Blaise et al., 2007; Fuchs et al., 2010). Another advantage to performing metabolomics studies is that they can be ‘hypothesis-free’; metabolomics approaches can identify and quantify metabolites without researchers knowing ahead of time what they anticipate to find or at what concentrations (Blaise et al., 2007).

Many of these characteristics are significantly advantageous for nonspecific studies intending to screen for entire compositions of tissues or fluids. In their study,

Atherton et al. used this metabolomics structure to analyze *C. elegans* and were able to quantify over 100 metabolites, of which 86 were identified. Their study was conducted to examine homologous genes between *C. elegans* and humans, and demonstrated that existing analytical techniques were sufficient enough to examine a multitude of metabolites in tandem while also being efficient, rapid, and relatively cheap (Atherton et al., 2008). A similar procedure, as shown in Figure 3, was conducted by Blaise et al. (2007), who also constructed metabolite profiles of multiple *C. elegans* strains exposed to varying concentrations of detoxification enzymes and compared them to one another. Their study was able to isolate specific differences in metabolite concentrations and identify which metabolic pathways were most directly impacted by the changing oxidant concentrations (Blaise et al., 2007). These studies exhibit the potential of metabolomics research in associating ‘silent’ genes with their affected metabolic pathways.

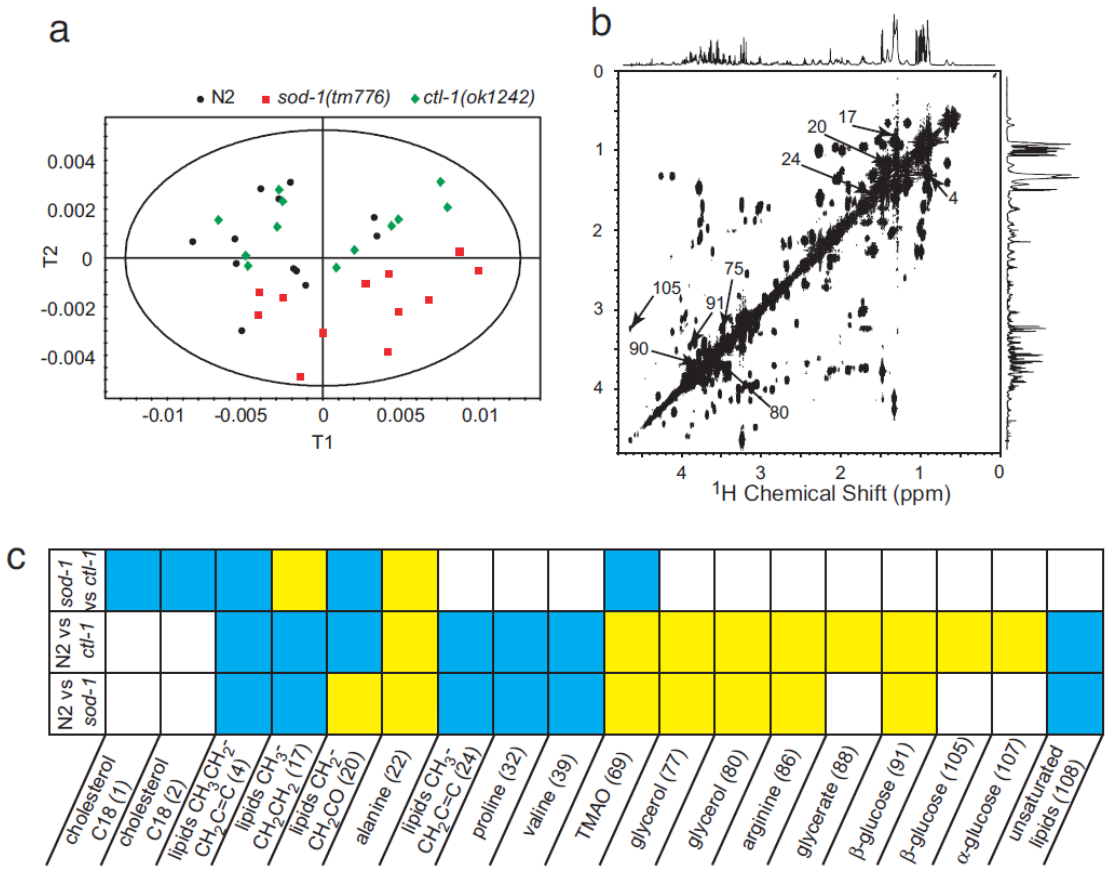

Figure 3: NMR analysis reveals differences in metabolite concentrations. (a) PLS-DA statistical representation of the differences in metabolite concentrations between the three strains. (b) 2D H-H NMR spectra, annotated with identified metabolites. (c) Visual representation of the differences noted between the three strains of *C. elegans* for each quantified metabolite (Blaise et al., 2007).

Metabolomics research is also useful for reflecting the real-time conditions of cells in their environment. While genomics and proteomics studies evaluate the relatively constant sequences and structures of genes and proteins respectively, metabolomics measures metabolites that naturally fluctuate depending on the development, condition, and stress of a cell (Raamsdonk et al., 2001). One of the most confounding yet intriguing

variables in metabolomics is the innate tendency of cells to try and correct for deleted or defective genes, and it was this compensating phenomenon that Morgan et al. (2015) examined in their study. After experimentally inducing defects in the channel proteins of mitochondria in *C. elegans*, they were able to measure changes in metabolite concentrations that reflected the altered cells attempting to shunt metabolites through different pathways to accommodate the loss in mitochondrial efficiency (Morgan et al., 2015). This evidence demonstrates how dynamic the cellular environment is, and also exhibits the strength of metabolomics studies in being able to measure subtle shifts in metabolic processes. These types of compensatory adjustments to cellular metabolism are similar to what would be expected with O-GlcNAcylation interference, and I will be assessing the samples in my metabolomics study for similar shifts in the metabolites most closely related to UDP-GlcNAc and the HBP.

6.4 Methods of Conducting Metabolomics Analysis

As Mahieu stated in a prior study, the goal of metabolomics analysis is, "to assay as many endogenous small molecules in a biological sample as possible" (Mahieu et al., 2014). Multiple different techniques have been employed to do so, each with its own advantages, weaknesses, and specialties. Currently, one of the most predominant tools used for metabolomics analysis is Nuclear Magnetic Resonance Spectroscopy, or NMR, which has been utilized previously in a multitude of metabolomics studies (Atherton et al., 2008; Blaise et al., 2007; Fuchs et al., 2010; Li et al., 2009; Nicholson et al., 1995; Salom et al., 2014).

In principle, NMR spectrometers operate by exposing a sample to a magnetic

field that polarizes the spins of any magnetic nuclei within the sample. The spectrometer then ‘flips’ these spins rapidly, producing a frequency shift, or ‘peak’, that can be measured and reflects the position, bond, and electromagnetic environment of that atom in the molecule (Kolker et al., 2011). As described by Nicholson et al., NMR has several important features that make it a convenient tool for metabolomics: NMR is capable of obtaining signals from *any* molecule that contains magnetic atomic nuclei (eg. ^{1}H, ^{13}C, ^{15}N), NMR spectra provide information on the specific locations of atoms and bonds within a molecule to allow for precise identification, NMR nonspecifically analyzes all of the molecules in a sample at once, NMR analysis does not destroy/consume samples and can be performed multiple times. Many different programs are available for performing NMR to identify specific molecules or characteristics, and modern NMR instrumentation is more precise and less expensive than ever before (Nicholson et al., 1995).

One of the most useful aspects of NMR is the sheer variety of possible analyses that can be performed with it. Basic single-dimensional NMR spectra are rapid, cheap, and provide ample information on the positions of all of a single type of magnetic nuclei present in all of the molecules in a sample. When combined with a second spectrum into a two-dimensional spectrum, these results can be used to identify specific molecules present in solution and even convey rough concentration approximations (Nicholson et al., 1995). Then, programs created to amplify specific signals or examine the coupling between certain atoms, such as 2D COSY, 2D HSQC, or 2D ROSY, can be utilized to provide additional spectra for comparison or reference (Atherton et al., 2008). Though normally used to identify small molecules such as metabolites, these analyses can even be used to identify the structures of much larger macromolecules, such as complex proteins.

A study by Salom et al. (2015) demonstrated this principle by using NMR to determine the crystalline structure of the Bovine rhodopsin protein, a feat that normally requires separate analysis using more refined and precise techniques such as x-ray crystallography.

Though very potent and powerful, NMR also has certain limitations. One of the drawbacks of NMR is that a sample compound must have magnetic nuclei to produce a signal. While sample magnetism is typically not a problem for ^{1}H NMR, spectra requiring less abundant isotopes such as ^{13}C or ^{15}N may require supplementing the model organism with the isotope to produce a useful signal (Salom et al., 2014). Another difficulty is retrieving meaningful data to interpret. Although improved sensitivity and resolution in modern NMR technology allow for more signals to be measured, they also produce more 'background noise' and interference between the signals. Particularly in studies involving tandem analysis of thousands of molecules, this interference can result in spectra filled with signals that are impossible to differentiate or accurately quantify (Nicholson et al., 1995). Though not insurmountable, these factors do require consideration and careful planning when utilizing NMR in metabolomics research.

Another technique that is equally vital to metabolomics research is Mass Spectrometry (MS), which is often combined with other reference studies such as Liquid Chromatography (LC) or Gas Chromatography (GC). Mass Spectrometry is a technique where molecules are fragmented and ionized to determine their mass and composition, and allows for the identification of molecules. Liquid Chromatography and Gas Chromatography use a similar idea of segregation to separate and identify molecules by their size or electrical charge, either in a liquid medium or a heated gaseous medium

(Mahieu et al., 2014). In much the same way as NMR, these techniques can also be modified in many ways to examine specific molecular characteristics, and have been in use for metabolomics research since 2003 (Li et al., 2009). Some of the strengths of MS, GC, and LC is that these techniques are cheap, efficient, and nonspecific, requiring only limited foreknowledge of a researcher's desired molecular targets. On the other hand, some of their drawbacks are that these techniques consume their samples and cannot individually assay all medium types (e.g. a GC analysis can only analyze compounds that can be vaporized without decomposing). These drawbacks, however, can be overcome by combining these analyses together, and prior studies have utilized LC-MS and GC-MS alongside NMR to perform expansive metabolomics research (Atherton et al., 2008; Li et al., 2009; Mahieu et al., 2014).

Although these powerful techniques may be able to provide ample information on the compounds present in a given biological sample, the information is often extremely complex, cumbersome, and difficult to interpret. In order to identify molecules and estimate their concentrations, online databases with reference statistics and measurements for each type of analysis are essential. One such source that is frequently used by researchers is the Systematic Protein Investigative Research Environment (SPIRE), which is an online database that is capable of integrating data from MS studies to identify proteins and other large molecules (Kolker et. al., 2011). Other online sources, such as the Spectral Database for Organic Compounds (SDBS) and the Biological Magnetic Resonance Bank (BMRB), serve similar functions for the aggregation and analysis of NMR spectra (Nicholson et al., 1995). Finally, once metabolites in a metabolomics study have been identified and quantified, studies frequently reference online metabolome

databases to identify which metabolic pathways are predominantly involved with any altered metabolites (Antoshechkin and Sternberg, 2007; McElwee et al., 2006).

6.5 C. elegans as a Model Organism

While metabolomics studies may use these tools and techniques to analyze the metabolites present in nearly any type of biological sample, *C. elegans* is one of the most frequently used model organisms. As described by Antoshechkin and Sternberg, *C. elegans* has several innate characteristics that make it an attractive model for many forms of research. This nematode is biologically very simple, has a rapid life cycle, has a small genome, has a transparent body, occurs naturally in soils throughout the globe, and is extremely durable and easy to maintain (Antoshechkin and Sternberg, 2007). The typical life span of *C. elegans* is approximately three weeks, and *C. elegans* continuously lay eggs for new progeny throughout that time span. However, as shown in Figure 4, their larvae can enter a dauer stage when exposed to harsh conditions that allows them to survive for up to three months. Strains of *C. elegans* can also be frozen in special conditions at -80°C and revitalized several years, even decades, later with minimal damage or sample loss, making them easy to store and excellent models for long-term experimentation (Fuchs et al., 2010). Reproduction in *C. elegans* also involves the majority of adult nematodes being hermaphrodites that self-reproduce, making controlling the inheritance of experimental genotypes extremely easy (Antoshechkin and Sternberg, 2007).

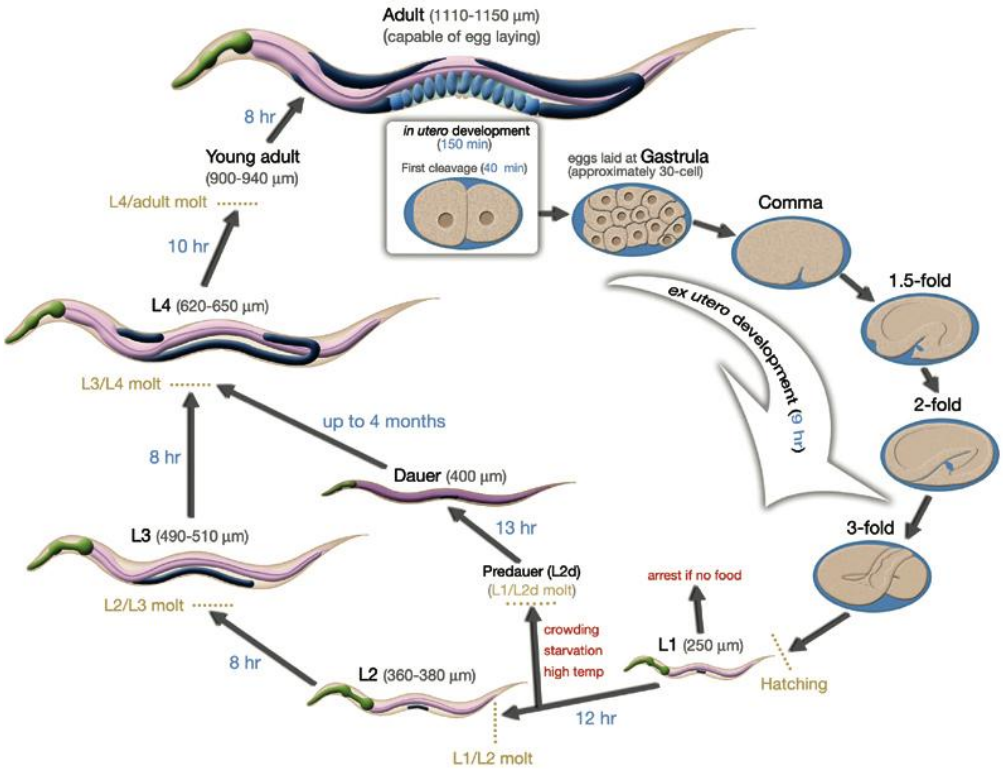

Figure 4: Life cycle of *Caenorhabditis elegans.* Because of their rapid life cycle, *C. elegans* nematodes can reach maturity after hatching within approximately two days in ideal conditions. As larvae, they may also adopt a 'dauer' state to survive poor conditions for several months (Wormatlas, 2006).

Because of its many advantages as a model organism, *C. elegans* has been utilized in a wide array of research from aging and development to signal transduction, immunology, and RNAi (Antoshechkin and Sternberg, 2007). As Antoshechkin and Sternberg describe in their research, the overwhelming number of studies using *C. elegans* has produced a wealth of established knowledge and findings regarding its biology, genomics, and metabolism. For example, *C. elegans* was the first multicellular organism to have its genome completely sequenced. Functional knock-out mutants exist for about one third of the nematode's ~20,000 genes, and online databases have

categorized and mapped over 5,600 different interactions between metabolites and metabolic networks within *C. elegans* (Antoshechkin and Sternberg, 2007). Multiple agencies, such as the *Caenorhabditis* Genetics Center which was established in 1978, operate solely to maintain standardized stocks of *C. elegans* strains from around the world for use in international research. By having ready supplies of both naturally occurring and experimentally mutated strains of *C. elegans,* any number of studies comparing these strains and their metabolic differences can be performed with incredible ease (Fuchs et al., 2010).

One study that demonstrated the usefulness and diversity of *C. elegans* was that conducted by Watson et al. In their study, they conducted a pseudo-reverse metabolomics experiment where they examined how feeding altered bacteria to *C. elegans* and manipulating their metabolites changed their gene expression. Researchers were required to not only monitor the physiology and metabolites of their *C. elegans* samples, but also to observe how their reproduction and fertility changed over generations, how their development and cell activity changed over time, and how their genomes and protein activity shifted depending on the bacteria they were fed (Watson et al., 2014). Many of these tasks could not have been performed in more complex or static model organisms, thus demonstrating how the versatility of *C. elegans* provides a uniquely flexible and adaptable model that is well suited for metabolomics studies.

6.6 Researching the Impact of O-GlcNAcylation on Metabolite Profiles

In regards to the different types of PTMs, no other form of protein glycosylation is as ubiquitous throughout the cell and deeply rooted in metabolic regulation as O-

GlcNAcylation (Farley and Link, 2009). Prior studies show that O-GlcNAcylation, its closely related biochemical pathways (eg. HBP), and its pool of substituent metabolites (eg. glucose, acetyl-CoA, UTP) directly fluctuate under the stressful conditions of several diseases, such as diabetes, obesity, and cancer (Atherton et al., 2008). As mentioned by Ghosh et al. (2014), it is possible to accurately evaluate and quantify the effects of protein glycosylation throughout the cell by using nonspecific analytical techniques such as NMR, LC-MS, and GC-MS to perform comparative metabolomics analysis. In the context of O-GlcNAcylation, the study of O-GlcNAcylation is also enabled by the fact that the model organism *C. elegans* has knock-out mutant strains (*ogt-1* and *oga-1*) that possess deletions for the two main enzymes involved in O-GlcNAcylation (OGT and OGA, respectively). These mutations are almost universally lethal in more complex model organisms, particularly in mammalian models, making *C. elegans* the ideal model for researching disruptions of O-OlcNAcylation (Ghosh et al., 2014).

Previous research by Ghosh et al. has already been performed to examine the consequences of interfering with O-GlcNAcylation in *C. elegans*. Using the same mutants previously described, *ogt-1* and *oga-1*, their findings suggested that O-GlcNAc, the sugar used to modify proteins in O-GlcNAcylation, may in fact act as a biochemical sensor of the cell's nutrition by directly upregulating or downregulating other related pathways, such as the HBP (Ghosh et al., 2014). These findings provide prime suspects for the metabolites and pathways most immediately affected by alterations to O-GlcNAcylation, but do not provide a significant examination of how these changes alter the rest of the cell's metabolism and biochemistry, such as gluconeogenesis or cell respiration. Given the ubiquitous nature of O-GlcNAcylation and the normally nonviable

phenotypes observed when it is disrupted in complex organisms, understanding how these other pathways are also impacted provides invaluable knowledge on the significance and effects of O-GlcNAcylation as a global PTM.

In this study, I will be using NMR spectroscopy to perform metabolomics analysis on three different *C. elegans* strains: N2 (wild type), *oga-1*, and *ogt-1*. I will construct profiles of the global metabolites present in these three strains and compare these profiles to identify any significant differences between the mutant strains and the wild type strain. These profiles and their differences will then be referenced to online databases and metabolomes to identify the biochemical pathways that are most directly impacted by the interruption of O-GlcNAcylation. Finally, these pathways will be examined and possible targets for therapeutic interventions will be identified. These findings may provide a viable platform for future research in *C. elegans* aiming to develop treatments for diseases that cause similar metabolic perturbations, such as diabetes, obesity, and cancer.

7. Methodology

7.1 Nematode Strains

The following strains of *C. elegans* were obtained from the *Caenorhabditis* Genetics Center (CGC) for use in this experiment: wild type N2 (Bristol N2), OGA knockout mutant RB1169 (*oga-1*(ok1207)), and OGT knockout mutant RB653 (*ogt-1*(ok430)). Backup samples of each strain were stored in freezing solution at -80°C as described by previous protocols (Brenner, 1974). Using a sterile technique, populations of *C. elegans* strains were grown on standard agar plates prepared with nematode growth medium (NGM), fed using *Escherichia coli* OP50 lawn cultures, and incubated at 16°C (Brenner, 1974).

Because the obtained strains of *oga-1* and *ogt-1* had not been outcrossed to remove unintentional mutations accumulated from being heavily mutagenized, both strains were backcrossed in-house with wild type N2 using standard backcrossing methods (Ahringer, 2006). The mutant strains were backcrossed with wild type N2 four times, with each backcrossed confirmed using duplex PCR analysis. Primers were constructed to amplify both the wild type OGA and OGT genes as well as their deletion variants in the *oga-1* and *ogt-1* mutants. PCR analysis with these same primers was also used in tandem with NMR analysis to ensure that each sample possessed nematode populations of the desired genotypes with no cross contamination from other strains.

7.2 Nematode Preparation

To perform comparative analysis of the cytosolic metabolites present in different

strains of *C. elegans*, nematode strains were grown en masse and their cellular contents extracted. Worms were grown in large quantities on NGM plates until full plate saturation was achieved. The worms were then rinsed from the plates with M9 sterile salt solution (3.0g/L KH_2PO_4; 6.0g/L Na_2HPO_4; 10g/L NaCl; 2.0mL/L 1M $MgSO_4$; adjusted to pH=6.0), collected into 15mL conical vials, and condensed into pellets via centrifugation at 2000 RPM for four minutes. Supernatant was then drained from each tube down to approximately 2mL without disturbing the worm pellet.

To control for changes in metabolite concentrations resulting from different life stages in *C. elegans*, nematode populations were synchronized to the same developmental stage. To synchronize the worms, 2mL of a bleach solution (1mL 10N NaOH; 4mL 8.25% NaOCl; 15mL M9 solution) was added to each conical vial. Each tube was consistently mixed and monitored until ~80% of the adult worms present were broken and their eggs dispersed, at which time the reaction was quenched by filling the vial to capacity with M9 solution. All vials were then centrifuged at 3000 RPM for five minutes to condense the nematode eggs, and their supernatants were drained down to ~2mL. Tubes were then refilled with M9, and the process of centrifugation and drainage repeated 3-4 times to sufficiently dilute any remaining bleach. Remaining eggs were then suspended in M9 on empty petri dishes and allowed to hatch for 18 hours while shaking at 16°C. Hatched larvae and L1 worms were then transferred to NGM agar plates seeded with OP50 and allowed to grow as normal in 16°C.

Synchronized nematode populations were monitored over the following 48-72 hours. Once the majority of the individuals had reached the L4 stage and before any adults began to lay eggs, the nematodes were rinsed from their plates with M9 solution

and again concentrated into 15mL conical tubes via centrifugation at 2000 RPM for four minutes. Each vial contained a population of approximately 2000 L4 individuals. The same method of draining supernatant, suspension in M9, and repeat centrifugation was applied to each tube 3-4 times to remove remaining bacteria and other contaminants. The final nematode pellets were then flash frozen in liquid N_2 and filled to 4mL with 80:20 Methanol:M9 solution to extract hydrophilic cellular metabolites. Samples were allowed to melt and mix completely with the methanol solution. To expose the nematode cellular contents, samples were then sonicated until 80% of nematode tissues were disrupted (Duty Cycle of 60; Output Control of 6). Remaining unbroken cellular tissues were then condensed via centrifugation at 3000 RPM for 10 minutes. Supernatant solutions containing metabolite extracts were removed, transferred to 1.5mL Eppendorf tubes, and dehydrated overnight in a vacuum concentrator. Dehydrated samples were retrieved, sealed, and stored at -20°C.

To perform NMR analysis, samples were rehydrated in 650μL NMR buffer solution (6.0mL D_2O and 0.24g trimethylsilyl propionic acid (TMSP) internal standard per 9 aliquots of buffer). Samples were then centrifuged again at 3000 RPM for 10 minutes to remove remaining particulate matter, and 600μL of each sample was transferred to an individual 5mm 600MHz Wilmad NMR tube.

7.3 Pilot ^{1}H NMR Spectroscopy

To verify the nematode population per sample required to obtain significant and consistent results, ^{1}H NMR spectroscopy was performed on samples prepared with 1,000, 2,000, or 4,000 N2 wild-type worms. Each population category was run in triplicate with

three replicate samples prepared for each. All samples were prepared and analyzed in tandem to control for any differences resulting from age related oxidization or degradation of the samples.

^{1}H NMR spectroscopy was carried out using a Varian MR-400 NMR spectrometer operating at a frequency of 400 MHz for the ^{1}H frequency, provided by the UNR Chemistry Department's Shared Instrument Laboratory (SIL). Triplicate samples of each population of worms were analyzed using a one dimensional NOESY pulse sequence to suppress occluding signals from residual H_2O (Atherton et al., 2008). Signal spectra were generated for each sample from 16 transients with 16K total data points each across a spectral width of 14ppm. All spectra were collected, standardized to the internal TMSP standard, and compared using Mnova NMR Suite (Mestrelab Research, S. L., 2016). Spectra were also adjusted within the Mnova NMR Suite using manual baseline correction and Fourier transform functions to more accurately represent the intensities and chemical shifts of measured signal peaks.

7.4 High Resolution Comparative NMR Spectroscopy

The same extraction procedure outlined previously in this methodology was used to prepare samples of the N2, *oga-1*, and *ogt-1* strains for NMR analysis. For each strain, three completely independent samples were prepared in isolated conditions to prevent cross-contamination. Each sample was comprised of approximately 2000 L4 nematodes of the appropriate strain. Following sonification and prior to dehydration using a vacuum concentrator, aliquots were taken from all samples and stained with Coomassie Blue dye. Per standard Bradford Assay protocols, a spectrophotometer was then used to generate a

standard curve of the concentrations of total proteins present in each sample, and the samples were adjusted accordingly by volume to control for differences in relative protein amounts (Bradford, 1976).

^{1}H NMR spectroscopy was then performed using a Varian V-500 NMR spectrometer operating at a frequency of 500 MHz for the ^{1}H frequency, also provided by the UNR SIL. Samples were analyzed using standard one dimensional ^{1}H NMR parameters (Atherton et al., 2008). Spectra for these samples were constructed using 128 transients with 16k data points each across a spectral width of 12ppm. Using the same Mnova NMR Suite software, all spectra were normalized according to the internal TMSP standard and adjusted using manual baseline correction and Fourier transform functions. Post-collection signal suppression was also used to isolate and remove lingering H_2O signal peaks. Finally, predominant and recurring landmark peaks were identified using peak-picking software, and the areas beneath their curves were measured with internal integration software. These areas were used as relative measures of the concentrations of the corresponding metabolites for their signal peaks.

7.5 Data Analysis

Chemical shifts and integrated areas of landmark signal peaks were collected and recorded from each spectra of the strain triplicates. The integrated areas for each signal peak were then averaged between the three spectra for each strain. To analyze the relative changes in intensity for each signal between the mutant strains, areas for signal peaks in the *oga-1* and *ogt-1* spectra were also normalized to the N2 signal peak intensities. These normalized peaks were then compared both to the wild-type N2 and to each other to

determine which signals showed significant alterations.

To relate the observed changes in signal peaks to possible metabolic differences between the strains, signal peaks were matched with likely correlated metabolite candidates. This was accomplished by comparing obtained spectra and signal peaks to known spectra from online databases, including: Madison Metabolomics Consortium Database, Spectral Database for Organic Compounds SDBS, and NMRShiftDB (Cui et al., 2008; SDSB Compounds and Spectral Search, 2016; NMRShiftDB, 2016). Landmark signals were approximated to their most likely representative class of molecule or, when possible, constituent metabolite. Finally, these metabolites were compared and their changes in intensity between *C. elegans* O-GlcNAcylation strains examined to hypothesize on possible metabolic perturbations.

8. Results

8.1 Demonstrating the Reliability of NMR

Spectra from ^{1}H NMR spectroscopy performed on 1,000, 2,000, and 4,000 N2 worm populations in triplicate were obtained, averaged, and compared (Figure 5). Analysis of the signal peaks present in these spectra focused on the relative positions and intensities of the peaks within the chemical shift range of {0, 4.5}, between the signals of the internal standard at δ=0.00 and the H_2O peak at δ=4.650. When compared to spectra constructed using samples of 2,000 nematodes (Figure 5B), spectra constructed using samples of 1,000 nematodes showed significantly reduced peak intensity and decreased definition of signals against background noise (Figure 5A). Contrasting spectra of 2,000 nematode samples and spectra of 4,000 nematode samples, however, demonstrated no significant differences. Thus there was no empirical support for increasing the required nematode samples sizes beyond 2,000 L4 individuals for effective NMR spectra results.

These spectra were also significant for their consistency between samples. First, the chemical shifts and presentations of signal peaks (whether they appeared as singlets, doublets, triplets, etc.) remained virtually unchanged between spectra. This consistency was true even amongst the notably diminished signals of the spectra produced using samples of 1,000 nematodes. Second, the peak intensities of the major signals remained relatively constant when comparing the spectra of the 2,000 and 4,000 N2 nematode samples. The consistency of these results allowed for the establishment of 2,000 L4 nematodes as an appropriate sample size for further NMR analysis.

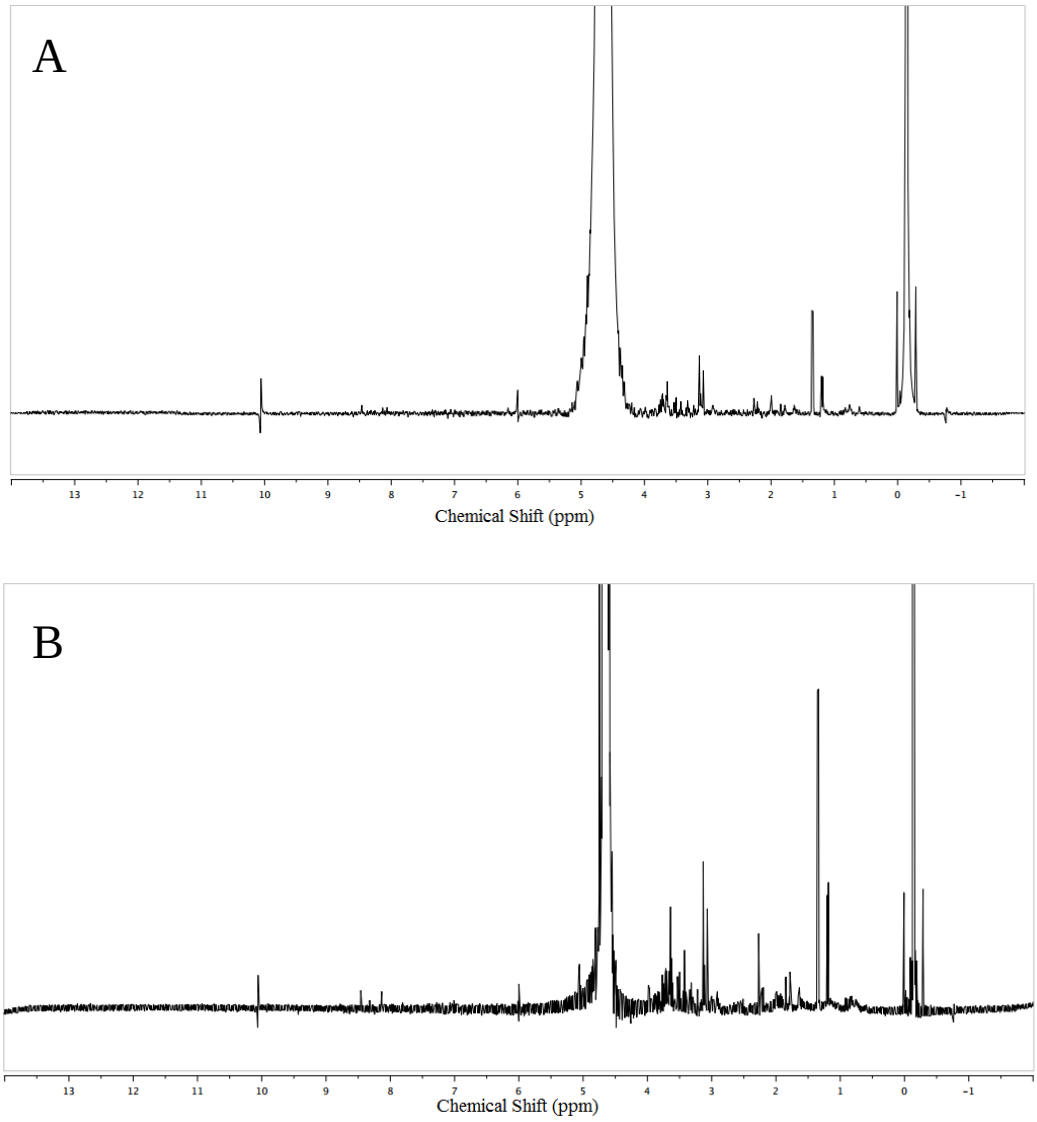

Figure 5: Sample size dependency of signal intensity in ^{1}H NMR. Proton NMR spectra were generated and compared between samples of 1,000 (A) and 2,000 (B) N2 nematodes. Examination of the consistency of signal peaks, signal definition, and peak intensity were used to determine result accuracy. Compared to samples of 1,000 worms, the 2,000 worm samples provided significantly greater resolution and peak intensity. Samples using 4,000 worms (not shown) were unchanged compared to 2,000 worm samples.

8.2 Determining Metabolite Identity and Relative Abundance

Spectra from higher resolution ^{1}H NMR spectroscopy of triplicate samples of N2, *oga-1*, and *ogt-1* nematode samples were obtained and analyzed using MNova NMR Suite software. After comparing these spectra, 16 key signal peaks, as well as the internal standard and H_2O peaks, were picked, identified by their chemical shifts (δ), and had the areas of their peaks integrated (Figure 6). The areas of these peaks were averaged among spectra for each strain, and then compared to the average areas of the same peaks in the other strains. In both the N2 and *oga-1* strains, notable variances in the metabolite concentrations were common. However, patterns in signal peak intensity between the strains did emerge with nine out of sixteen *oga-1* peaks having less intensity than corresponding N2 peaks and thirteen out of sixteen *ogt-1* peaks having greater intensity than corresponding N2 peaks.

These key signal peaks were also identified by their likely corresponding metabolites using NMR spectra databases. Identifications were as follows: 1. Cholesterol (δ=0.620); 2. Palmitate (δ=0.835); 3. Acetate (δ=1.215); 4. Lactate (δ=1.360); 5. Citrulline (δ=1.595); 6. Unidentified Metabolite (UIM; δ=1.799); 7. Glutamate (δ=2.020); 8. Butyrate (δ=2.280); 9. Citrate (δ=2.550); 10. Asparagine (δ=2.915); 11. Creatine (δ=3.085); 12. β-Glucose (3.235); 13. α-Glucose (3.439); 14. Choline (3.656); 15. Phenylalanine (δ=7.305); 16. Deoxyadenosine (δ=8.336); 17. TMSP internal standard (δ=0.0); and 18. H_2O (δ=4.650). For all of these signal peaks, the correlated metabolites were the most well supported of several candidates, but ^{1}H NMR spectral data alone were insufficient to exclude the possibility of a Type I error. Specific isoforms of each metabolite were unable to be determined due to this lack of experimental power, and in

the case of signal ‘5’ no probable metabolite candidate was able to be identified.

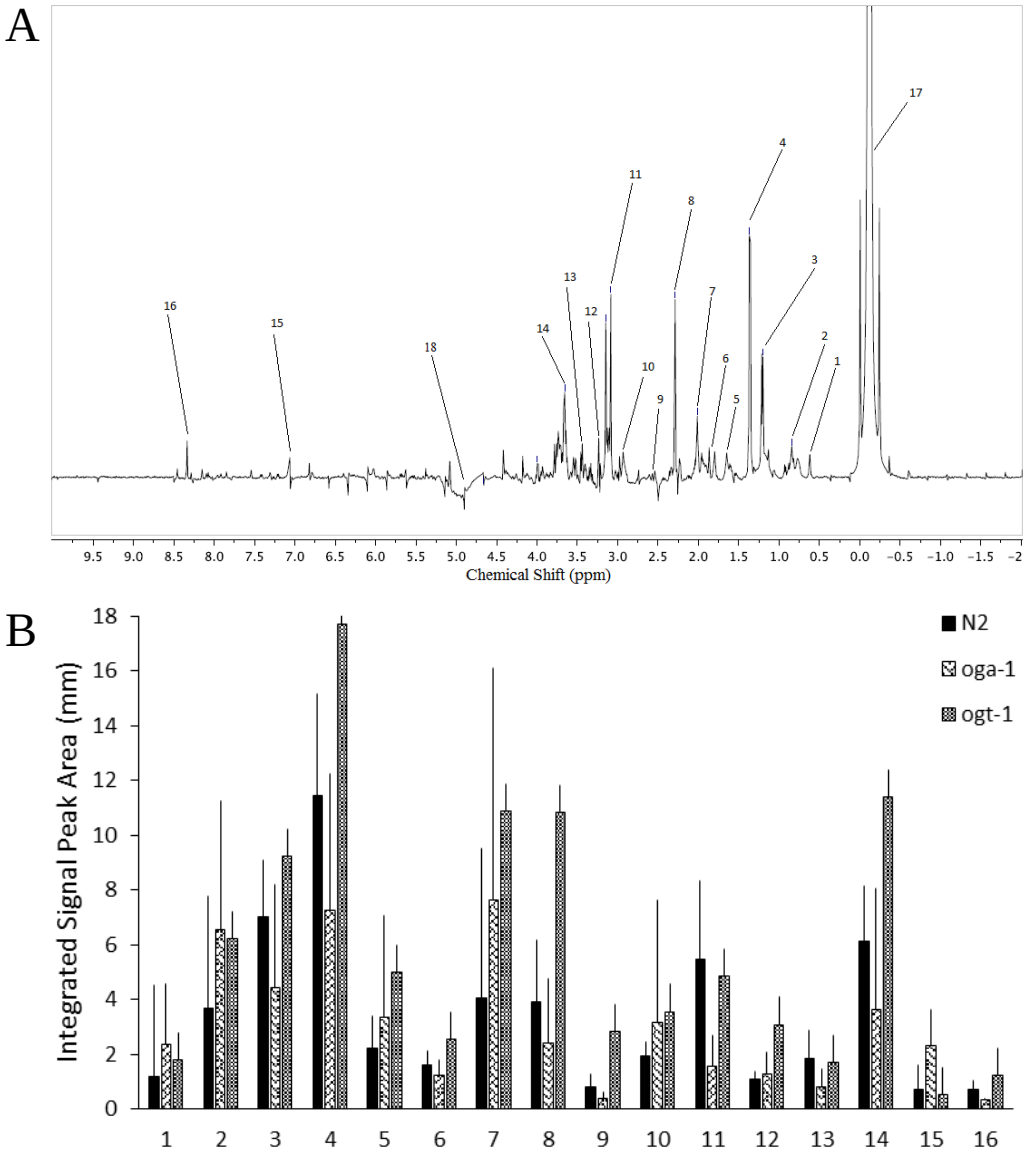

Figure 6: Identification and quantification of key signal peaks. Landmark signal peaks were labeled in the spectra for each strain’s triplicate samples (A). Labeled peaks were then integrated and their areas averaged for comparison between strains (B). Overall peak intensity remained largely unchanged for some signals (eg. 1, 2, and 3), while others showed significant differences (eg. 7, 8, and 14). Compared to the control N2 strain, *oga-1* metabolite signals were most often reduced while *ogt-1* signals were more often increased.

8.3 Quantifying Metabolite Differences between Mutant Strains

To better quantify the changes in metabolite concentrations in the *oga-1* and *ogt-1* mutants, the difference in metabolite abundance between the wild type N2 and both individual strains for each metabolite was found and normalized by the corresponding abundance of the N2 control (Figure 7). These normalized abundances better represented the magnitudes of the changes in metabolite abundance in the mutant strains.

When normalized by the 'expected' N2 metabolite abundances, the *oga-1* strain typically demonstrated repressed metabolite abundances. Metabolite reductions were greatest (over 50% lower than control levels) in citrate, creatine, α-glucose, and deoxyadenosine, with notable reductions (over 25% lower than control levels) also occurring in acetate, lactate, butyrate, and choline. Comparatively, increases in metabolite levels were far less common, occurring in seven out of sixteen metabolites. Slight increases (between 50-100%) were noted in palmitate and citrulline, a moderate increase (over 100%) was noted in cholesterol, and a significant increase (over 200%) was noted in phenylalanine.

In contrast, *ogt-1* samples demonstrated a pattern of general metabolite increase when normalized by N2 metabolite levels. All but three metabolites showed some increase in abundance over N2 levels, with slight increases in cholesterol, palmitate, lactate, asparagine, choline, and deoxyadenosine, moderate increases in citrulline, glutamate, butyrate, and β-glucose, and a significant increase in citrate. Of the three metabolites that were reduced in *ogt-1* samples (creatine, α-glucose, and phenylalanine), only phenylalanine showed a truly noticeable change in abundance with just over a 25% reduction.

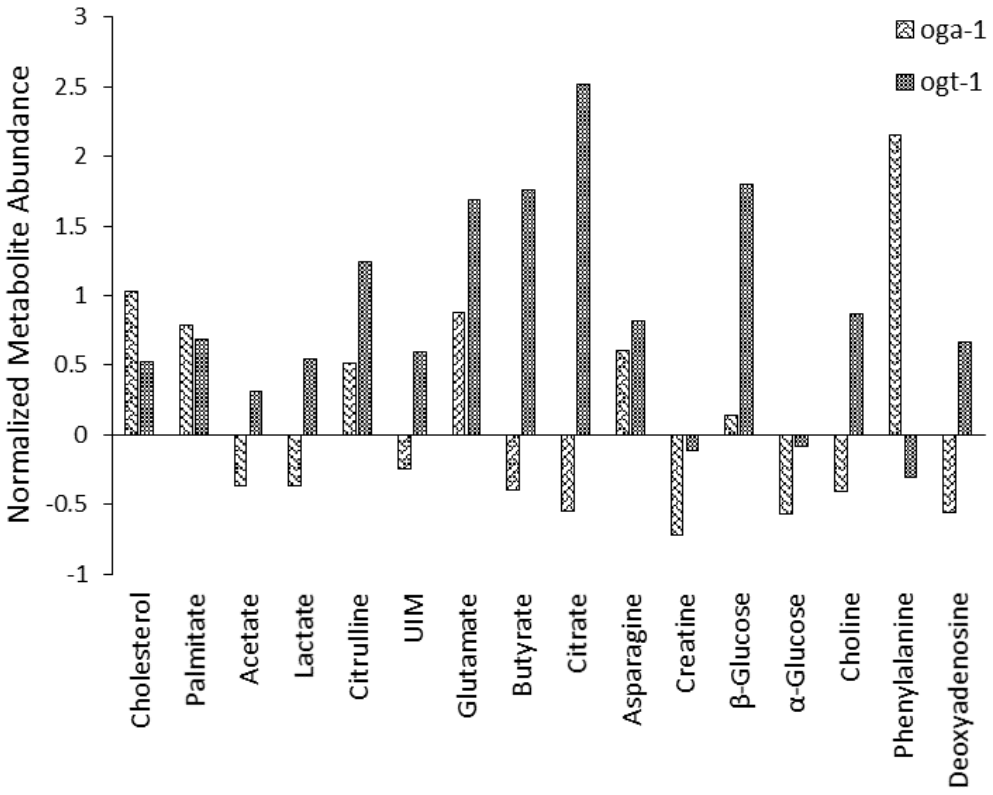

Figure 7: Metabolite level differences in *C. elegans* O-GlcNAcylation mutants. For each metabolite signal peak, the difference between the mutant strains and wild type N2 was found, and the result normalized by the intensity of the metabolite in N2. When examining the *oga-1* strain, one metabolite was increased one-fold (cholesterol) and another increased two-fold (phenylalanine), while four metabolites were reduced by more than 50% (citrate, creatine, α-glucose, and deoxyadenosine). However, the *ogt-1* strain showed no similar reductions in metabolites while instead having four metabolites increased one-fold (citrulline, glutamate, butyrate, and β-glucose), and one metabolite increased over two-fold (citrate).

It is also important to compare the changes in metabolite concentration between both mutant strains. For half of the listed metabolites, the measured *oga-1* and the *ogt-1* metabolite abundances were elevated or reduced in an agonistic fashion. However in the

other half of the metabolites, the two strains demonstrated shifts in metabolite abundances that directly opposed one another. It is worth noting that some metabolite changes, such as the decreases of *ogt-1* abundances of creatine and α-glucose, were small enough to be considered insignificant. In these cases, the metabolites are considered to have remained neutral or unaltered, as opposed to being reduced or increased in the mutant strains.

9. Discussion

9.1 O-GlcNAcylation Modulates CAC Activity

Several of the metabolites most noticeably altered in the O-GlcNAcylation mutant strains are directly involved or closely related to the Citric Acid Cycle (CAC). This cycle, which is the principle process for generating chemical energy from fats, proteins, and sugars, is central to aerobic metabolism and hinges on several enzymes and intermediary metabolites (Wagner and Fell, 2001). One measured metabolite, citrate, is a critical component of the CAC, while three other metabolites, lactate, glutamate, and asparagine, are nearby derivatives of the CAC (Figure 8). This cycle is most often referenced in regards to the metabolism of carbohydrates via the consumption of their derivatives pyruvate and acetyl-CoA. However, acetyl-CoA may also be derived from several other precursor metabolites via protein and fatty acid metabolism. The CAC also has several intermediate metabolites that may be shunted toward the catabolism of fatty acids, amino acids, and nucleotides. These factors make the CAC integral to cellular metabolism.

As the first product of acetyl-CoA metabolism, citrate acts as an excellent benchmark for CAC activity. In the *ogt-1* strain, levels of this metabolite were significantly increased with respect to the wild type N2, while the *oga-1* strain showed a significant decrease in this metabolite. These findings suggest that CAC activity is inhibited by O-GlcNAcylation, and that the cycle is more active in the absence of any O-GlcNAcylation. Another metabolite lactate is typically produced in high activity conditions when CAC activity is unable to compensate for energy demands (Foll et al, 1999). The *ogt-1* mutants demonstrated an overall increase in levels of lactate, while *oga-*

1 showed an overall decreased. These changes suggest that the enzymatic production of lactate is also inhibited by O-GlcNAcylation activity. It is also possible that an upstream enzyme responsible for the conversion of glucose to pyruvate is inhibited by O-GlcNAcylation. O-GlcNAcylation directed inhibition would also explain the matching rise in CAC activity, which is also dependent on the availability of pyruvate.

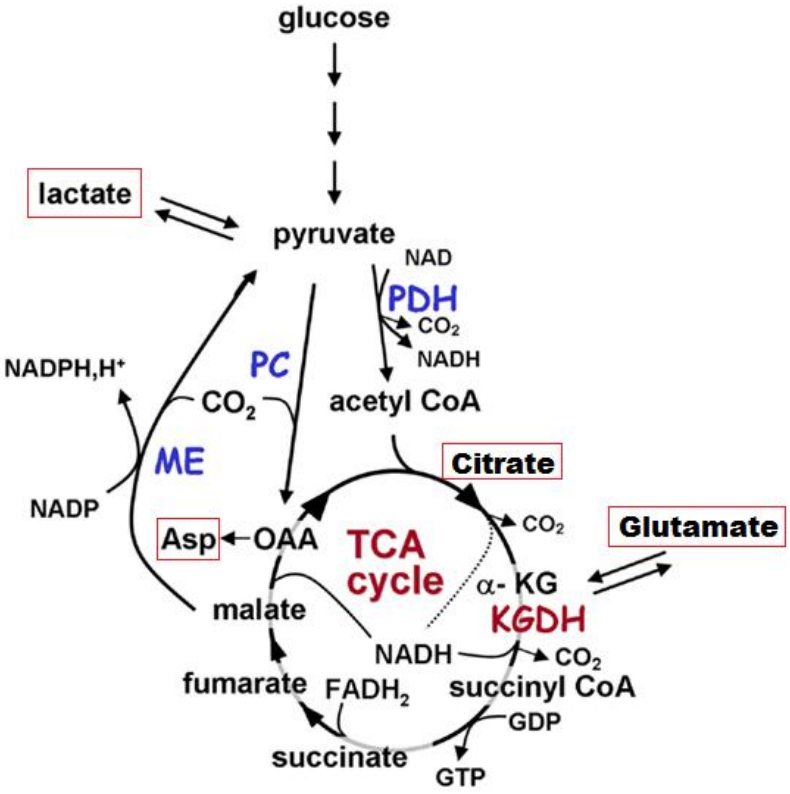

Figure 8: Measured metabolites involved in the Citric Acid Cycle. The Citric Acid Cycle consumes, transforms, and produces several different metabolites critical to multiple metabolic pathways. Citrate is directly involved in this cyclical process, while glutamate, asparagine, and lactate interact with the CAC and other metabolic pathways. Changes in the abundances of these metabolites reflect alterations in the activity of the CAC and other branching metabolic pathways (Quek et al., 2009).

Measured levels of asparagine and glutamate also help reveal the activity of the CAC. Besides participating in energy retrieval through CAC activity, both of these metabolites are also amino acids incorporated in the structures of proteins. In the *ogt-1* Upregulation of asparagine and glutamate supports the earlier findings that O-GlcNAcylation inhibits CAC activity, and in its absence these metabolites are produced in greater quantities. However, both glutamate and asparagine were slightly elevated in the *oga-1* strain as well. These findings appear to contradict the explanation of the CAC being inhibited when more heavily O-GlcNAcylated. Therefore, matching changes in the relative levels of asparagine and glutamate may instead be due to regulation of other CAC independent enzymes involved in amino acid synthesis. Further measurement of additional metabolites involved in the specific metabolism of glutamate and asparagine could help distinguish which of the two alternative hypotheses is more accurate.

The relationship between energy production via the CAC and activity of O-GlcNAcylation was demonstrated through the levels of creatine present in the mutant strains. Creatine, which serves as an alternative source of chemical energy during phases of extreme acute cellular activity, also acts as a benchmark of an organism's resting 'reserve' energy (Wu & Morris, 1998). Levels of measured creatine were largely unchanged in the *ogt-1* strain, but were dramatically decreased in the *oga-1* strain. A lower abundance of creatine in *oga-1* suggests that they have little 'reserve' energy, a finding that coincides well with the previously mentioned results demonstrating that CAC was upregulated by O-GlcNAcylation. Overall, these findings point strongly toward an intimate between O-GlcNAcylation and regulation of CAC-associated metabolism.

9.2 Disruption of Nutrient Reserves

Due to the reliance of the production of UDP-GlcNAc on carbohydrate precursors, it was hypothesized that alterations to O-GlcNAcylation would also impact sugar abundances. When I examined the measured levels of α-glucose and β-glucose (which serve as nonspecific measures of several carbohydrate isoforms), the *ogt-1* strain showed an overall increase in sugar levels while the *oga-1* strain shows an overall decrease. This suggests that the *oga-1* strain is metabolizing available sugars at a greater rate, which would be expected given the lack of a functional OGA enzyme to remove bound UDP-GlcNAc from proteins. The opposite trend is seen in the *ogt-1* strain, again affirming relationship between O-GlcNAcylation activity and the amount of available carbohydrates.

Alterations in nutrient abundances were also noted in phenylalanine, citrulline, and deoxyadenosine. Levels of phenylalanine and citrulline, amino acids involved in separate metabolic pathways, were both elevated in the *oga-1* strain. In the *ogt-1* strain, citrulline was also elevated while phenylalanine was reduced. While nested within different metabolic pathways, these results implicate amino acid biosynthesis as also being impacted by O-GlcNAcylation even when the constituent amino acids are not directly involved in the CAC. Nucleotides were also shown to be impacted by O-GlcNAcylation, such as the measured metabolite deoxyadenosine. These nucleotides are dependent on products of the previously described pentose phosphate pathway, which is in turn impacted by other nutrients including simple sugars. Deoxyadenosine was significantly decreased in the *oga-1* strain, and slightly increased in the *ogt-1* strain, which correlates well with the changes in the HBP intermediary acetyl-CoA. This pattern

of inhibition by O-GlcNAcylation also mirrors that of the abundances of measured sugars, further supporting the impact of O-GlcNAcylation on the levels of available carbohydrates and other cellular nutrients.

9.3 Altered Regulation of Fatty Acid Metabolism

Prior to being transformed into citrate for fueling the CAC, acetyl-CoA may also shunt to other pathways. Another metabolite it may form is acetate, a molecule that is used as a fundamental building block for the synthesis and extension of fatty acids (Foll et al., 1999). In the mutant strains, acetate abundance was largely unchanged in *ogt-1* and decreased in *oga-1*. The previously discussed inhibition of CAC activity would normally result in an excess of available acetyl-CoA, so it is unlikely that a lack of a precursor metabolite is to blame for the decreased acetate abundance. Instead, it is more likely that the *oga-1* strain is experiencing an increase in fatty acid synthesis and consuming more acetate. Increased fatty acid synthesis is supported by the observed levels of butyrate and choline, both of which are small metabolites also involved in supporting fatty acid synthesis (Zeisel & da Costa, 2009). The *ogt-1* strain demonstrated significant increases in the abundances of both of these metabolites, while the *oga-1* strain showed clear decreases. These patterns correspond well with the changes in acetate levels, further supporting the hypothesis that protein O-GlcNAcylation upregulates fatty acid synthesis.

The metabolism of fatty acids also involves the production of more complex intermediate metabolites. One such metabolite is cholesterol, which is produced from acetate through the mevalonate pathway and is ultimately used to construct a variety of hormones, sterols, and plasma membrane structures (Payne & Hales, 2004). Another

metabolite indicative of fatty acid synthesis is palmitate, a ubiquitous fatty acid found in nearly all multicellular organisms (Bar-Yoseph et al., 2013). Both of these metabolites are derived from byproducts of the CAC, such as acetate, and both were elevated in the mutant strains. The degree of elevation in each was more significant in the *oga-1* strain, however, suggesting again that protein O-GlcNAcylation either enhances fatty acid metabolism or inhibits fatty acid catabolism.

10. Conclusions

In the scope of protein PTMs, O-GlcNAcylation is unique in the breadth of its effects and the diversity of proteins it targets for modification. Combined with the facts that O-GlcNAcylation can activate or inhibit its targets depending on the protein, and that it is simultaneously cooperating and competing with other PTMs such as phosphorylation, O-GlcNAcylation is far from a simple or easily understood activity. This is particularly true when considering the role of O-GlcNAcylation on moderating metabolism, of which the effects are numerous and frequently interrelated.

The results of this research reveal much about the multifaceted role of protein O-GlcNAcylation. Previous research hypothesized that O-GlcNAcylation directly modulated glucose metabolism and hexosamine biosynthetic pathway activity. Findings from this study support a similar conclusion with O-GlcNAcylation activity lowering levels of both carbohydrates and nucleotides. The clear inhibitory effect of O-GlcNAcylation on the CAC and creatine production also demonstrates the importance of O-GlcNAcylation on cell respiration. Energy metabolism regulation was also evidenced in the observed changes to fatty acid synthesis, both in the upregulation of their synthesis and the downregulation of their degradation.

Together, these findings describe a complex network of O-GlcNAcylation modifications that moderate carbohydrate and nutrient metabolism, cell respiration activity, and lipid biosynthesis. Such implications correspond well with the results of prior studies on O-GlcNAcylation, particularly the hypotheses that O-GlcNAcylation is intimately involved in insulin-like signaling and energy storage determination proposed by Forsythe et al. (2006). While the HBP, gluconeogenesis, CAC activity, and fatty acid

synthesis have all been individually associated with O-GlcNAcylation, the results of this study have now demonstrated their simultaneous changes within the *oga-1* and *ogt-1* O-GlcNAcylation mutants. With the effects on the metabolic pathways and their interactions with one another established, identification of common signaling cascades or enzymes involved in these pathways, such as the previously described *daf-2* insulin-like receptor, will allow for further manipulation of O-GlcNAcylation's complex effects.

To build further on this research, additional studies are required to more accurately identify an even broader range of metabolites with greater precision. Using techniques such as LC-MS, GC-MS, and ^{13}C NMR could provide these measurements, and performing the analysis using greater numbers of samples per strain would also grant greater experimental power to determine which changes are most significant. Although limitations in time and resources prevented these techniques from being incorporated into this current study, the results demonstrated here provide a more precise focus for their future implementation. New extraction procedures may also be designed to specifically isolate metabolites from the discussed pathways, which have now been shown to be significantly altered by O-GlcNAcylation.

While much research in disease treatment attempts to discover a singular or straightforward 'cure', such a solution may not always be feasible. In many diseases like diabetes, obesity, neurodegenerative disorders, and especially cancer, the number of factors that contribute to the onset and propagation of the disease create a complex scenario that cannot be traced to a single treatable alteration. However, exploring how metabolites themselves fluctuate under stress provides a basis for developing new treatments to counteract the adverse effects of these diseases. New techniques may be

developed to modify the metabolism of any of the metabolites discussed in this study, or any metabolites addressed in future research, to effectively treat these complex diseases. Indeed, just as metabolic abnormalities produce complex problems, so to may metabolomics research provide equally elegant solutions.

11. Bibliography

Ahringer, J. (2006) Reverse genetics. *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.47.1, http://www.wormbook.org.

Antoshechkin, I., & Sternberg, P. W. (2007). The versatile worm: genetic and genomic resources for Caenorhabditis elegans research. *Nature Reviews Genetics, 8*(7), 518-532. doi:10.1038/nrg2105

Atherton, H. J., Jones, O. A. H., Malik, S., Miska, E. A., and Griffin, J. L. (2008). A comparative metabolomic study of NHR-49 in *Caenorhabditis elegans* and PPAR-a in the mouse *FEBS Letters, 582*(12), 1661-1666. doi:10.1016/j.febslet.2008.04.020

Bar-Yoseph, F., Lifshitz, Y., and Cohen, T. (2013). Review of sn-2 palmitate oil implications for infant health. *Prostaglandins Leukotrienes and Essential Fatty Acids, 89*(4), 139-143. doi: 10.1016/j.plefa.2013.03.002

Blaise, B. J., Giacomotto, J., Elena, B., Dumas, M.-E., Toulhoat, P., Segalat, L., and Emsley, L. (2007). Metabotyping of *Caenorhabditis elegans* reveals latent phenotypes. *PNAS, 104*(50), 19808-19812. doi:10.1073/pnas.0707393104

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry, 72*(1), 248-254.

Brenner, S. (1974). Genetics of Caenorhabditis-elegans. *Genetics, 77*(1), 71-94.

Cui, Q., Lewis, I. A., Hegeman, A. D., Anderson, M. E., Li, J., Schulte, C. F., Westler, W. M., Eghbalnia, H. R., Sussman, M. R., and Markley, J. L. (2008). Metabolite identification via the Madison Metabolomics Consortium Database. *Nature Biotechnology, 26*(1), 162-164.

Darley-Usmar, V. M., Ball, L. E., and Chatham, J. C. (2011). Protein O-linked β-N-acetylglucosamine: A novel effector of cardiomyocyte metabolism and function. *Journal of Molecular and Cellular Cardiology, 52*, 538-549. doi: 10.1016/j.yumcc.2011.08.009

Farley, A. R., & Link, A. J. (2009). Identification and Quantification of Protein Posttranslational Modifications. *Guide to Protein Purification, Second Edition, 463*, 725-763. doi:10.1016/s0076-6879(09)63040-8

Foll, R. L., Pleyers, A., Lewandovski, G. J., Wermter, C., Hegemann, V., and Paul, R. J. (1999). Anaerobiosis in the nematode *Caenorhabditis elegans*. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology, 124*(3), 269-280. doi:10.1016/s0305-0491(99)00130-3

Forsythe, M. E., Love, D. C., Lazarus, B. D., Kim, E. J., Prinz, W. A., Ashwell, G., Krause, M. W., and Hanover, J. A. (2006). *Caenorhabditis elegans* ortholog of a diabetes susceptibility locus: *oga-1* (*O*-GlcNAcase) knockout impacts *O*-GlcNAc

cycling, metabolism, and dauer. *PNAS, 103*(32), 11952-11957. doi: 10.1073/pnas.0601931103

Fuchs, S., Bundy, J. G., Davies, S. K., Viney, J. M., Swire, J. S., and Leroi, A. M. (2010). A metabolic signature of long life in *Caenorhabditis elegans BMC Biology, 8*(14), 1-12. doi:10.1186/1741-7007-8-14

Ghosh, S. K., Bond, M. R., Love, D. C., Ashwell, G. G., Krause, M. W., and Hanover, J. A. (2014). Disruption of O-GlcNAc Cycling in *C. elegans* Perturbs Nucleotide Sugar Pools and Complex Glycans. *Frontiers in Endocrinology, 5*(197), 1-7. doi:10.3389/fendo.2014.00197

Hanover, J. A., Forsythe, M. E., Hennessey, P. T., Brodigan, T. M., Love, D. C., Ashwell, G., and Krause, M. (2005). A *Caenorhabditis elegans* model of insulin resistance: Altered macronutrient storage and dauer formation in an IGT-1 knockout. *PNAS, 102*(32), 11266-11271. doi: 10.1073/pnas.0408771102

Kamath, K. S., Vasavada, M. S., and Srivastava, S. (2011). Proteomic databases and tools to decipher post-translational modifications. *Journal of Proteomics, 75*(1), 127-144. doi:10.1016/j.jprot.2011.09.014

Kolker, E., Higdon, r., Welch, D., Bauman, A., Stewart, E., Haynes, W., Broomall, W., and Kolker, N. (2011). SPIRE: Systematic protein investigative research environment. *Journal of Proteomics, 75*(1), 122-126. doi:10.1016/j.jprot.2011.05.009

Li, J., Cai, T., Wu, P., Cui, Z., Chen, X., Hou, J., Xie, Z., Xue, P., Shi, L., Liu, P., Yates III, J. R., and Yang, F. (2009). Proteomic analysis of mitochondria from *Caenorhabditis elegans*. *Proteomics, 9*(19), 4539-4553. doi:10.1002/pmic.200900101

Mahieu, N. G., Huang, X. J., Chen, Y. J., and Patti, G. J. (2014). Credentialing Features: A Platform to Benchmark and Optimize Untargeted Metabolomic Methods. *Analytical Chemistry, 86*(19), 9583-9589. doi:10.1021/ac503092d

McElwee, J. J., Schuster, E., Blanc, E., Thornton, J., and Gems, D. (2006). Diapause-associated metabolic traits reiterated in long-lived*daf-2* mutants in the nematode *Caenorhabditis elegans*. *Mechanisms of Ageing and Development, 127*(5), 458-472. doi:10.1016/j.mad.2006.01.006

"Mnova NMR." *Mestrelab Research.* Mestrelab Research, S. L., 2016. Web. 5 Apr. 2016. http://mestrelab.com/software/mnova/nmr/

Morgan, P. G., Higdon, R., Kolker, N., Bauman, A. T., Ilkayeva, O., Newgard, C. B., Kolker, E., Steele, L. M., and Sedensky, M. M. (2015). Comparison of proteomic and metabolomic profiles of mutants of the mitochondrial respiratory chain in *Caenorhabditis elegans*. *Mitochondrion, 20*, 95-102. doi:10.1016/j.mito.2014.12.004

Nicholson, J. K., Foxall, P. J. D., Spraul, M., Farrant, R. D., and Lindon, J. C. (1995). 750-MHz H-1 and H-1-C-13 NMR-Spectroscopy of Human Blood-Plasma. *Analytical Chemistry, 67*(5), 793-811. doi:10.1021/ac00101a004

New England Biolabs. (2016). Glycobiology. Retrieved from https://www.neb.com/applications/glycobiology

NMRShiftDB. (2016, January 1). Retrieved from http://nmrshiftdb.nmr.uni-koeln.de/portal;jsessionid=83E5450F43239EA34EE3CC9810B19D44

Payne, A. H. & Hales, D. B. (2004). Overview of steroidogenic enzymes in the pathway from cholesterol to active steroid hormones. *Endocrine Reviews, 25*(6), 947-970. doi: 10.1210/er.2003-0030

Quek, L. E., Wittmann, C., Nielsen, L. K., and Kromer, J. (2009). OpenFLUX: efficient modeling software for ^{13}C-based metabolic flux analysis. *Microbial Cell Factories, 8*(25), 1-15. doi: 10.1186/1475-2859-8-25

Raamsdonk, L. M., Teusink, B., Zhang, N., Walsh, M. C., Berden, J. A., Brindle, K. M., Broadhurst, D., Hayes, A., Kell, D. B., Oliver, S. G., Rowland, J. J., Westerhoff, H. V., and Van Dam, K. (2001). A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology, 19*(1), 45-50. doi:10.1038/83496

Rahman, M. M., Stuchlick, O., El-Karim, E. G., Stuart, R., Kipreos, E. T., and Wells, L. (2010). Intracellular protein glycosylation modulates insulin mediated lifespan in C. elegans. *Aging, 2*(10), 678-690.

Salom, D., Cao, P., Yuan, Y., Miyagi, M., Feng, Z., and Palczewski, K. (2014). Isotopic labeling of mammalian G protein-coupled receptors heterologously expressed in *Caenorhabditis elegans Analytical Biochemistry, 472*, 30-36. doi:10.1016/j.ab.2014.11.008

SDBS Compounds and Spectral Search. (2016, March 18). Retrieved from http://sdbs.db.aist.go.jp/sdbs/cgi-bin/direct_frame_top.cgi

Wagner, A. & Fell, D. A. (2001). The small world inside large metabolic networks. *Proceedings of the Royal Society B-Biological Sciences, 268*(1478), 1803-1810. doi:10.1098/rspb.2001.1711

Watson, E., MacNeil, L. T., Ritter, A. D., Yilmaz, L. S., Rosebrock, A. P., Caudy, A. A., and Walhout, A. J. M. (2014). Interspecies Systems Biology Uncovers Metabolites Affecting C. elegans Gene Expression and Life History Traits *Cell, 156*, 759-770. doi:10.1016/j.cell.2014.01.047

Wormatlas. (2006). Introduction to *C. elegans* Anatomy. Retrieved from http://www.wormatlas.org/ver1/handbook/anatomyintro/anatomyintro.htm

Wu, G. Y. & Morris, S. M. (1998). Arginine metabolism: nitric oxide and beyond. *Biochemical Journal, 336*(1), 1-17.

Zeisel, S. H. & da Costa, K. A. (2009). Choline: an essential nutrient for public health. *Nutrition Reviews, 67*(11), 615-623. doi: 10.1111/j.1753-4887.2009.00246.x