Inclusion of Kinetic Proteomics in Multi-Omics Methods to Analyze Calorie Restriction Effects on Aging

Richard Hajime Carson
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Physical Sciences and Mathematics Commons

BYU ScholarsArchive Citation
Carson, Richard Hajime, "Inclusion of Kinetic Proteomics in Multi-Omics Methods to Analyze Calorie Restriction Effects on Aging" (2019). Theses and Dissertations. 9116.
https://scholarsarchive.byu.edu/etd/9116

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Inclusion of Kinetic Proteomics in Multi-Omics Methods to Analyze
Calorie Restriction Effects on Aging

Richard Hajime Carson

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

John C. Price, Chair
David V. Dearden
Ryan T. Kelly
Stephen R. Piccolo
Barry M. Willardson

Department of Chemistry and Biochemistry
Brigham Young University

Copyright © 2019 Richard Hajime Carson
All Rights Reserved
ABSTRACT

Inclusion of Kinetic Proteomics in Multi-Omics Methods to Analyze Calorie Restriction Effects on Aging

Richard Hajime Carson
Department of Chemistry and Biochemistry, BYU
Doctor of Philosophy

One of the greatest risk factors for disease is advanced age. As the human lifespan has increased, so too have the burdens of caring for an increasingly older population suffering from rising rates of cardiovascular disease, kidney disease, diabetes, and dementia. The need for improving medical technology and developing new therapies for age-related diseases is manifest. Yet our understanding of the processes of aging and how to attenuate the effects of aging remains incomplete. Various studies have established calorie restriction as a robust method for extending lifespan in laboratory organisms; however the mechanism is a topic of much debate. Advancing our understanding of calorie restriction holds promise for illuminating biochemical processes involved in the aging process.

One of the best explanations for the lifespan extension benefits of calorie restriction is that it improves cellular protein homeostasis (proteostasis), but because proteostasis is dynamic, it can be difficult to measure. We developed a novel combined omics methodology integrating kinetic proteomics, and applied it to a mouse model placed on calorie restriction. Our unbiased approach integrating just three measurements (kinetic proteomics, quantitative proteomics, and transcriptomics) enabled us to characterize the synthesis and degradation of thousands of proteins, and determine that calorie restriction largely alters proteostasis by slowing global protein synthesis post-transcriptionally. Validating our omics approach, we were able to replicate many previous results found in the literature, demonstrating the differential regulation of various protein ontologies in response to the nutrient stress of calorie restriction. Moreover, we were able to detect differential degradation of the large and small ribosomal subunits under calorie restriction, and proposed a model in which the rate of protein synthesis could be attenuated by the depletion of the large ribosomal subunit relative to the small subunit.

The flexibility of our dynamic combined omics approach was demonstrated by the expansion of measurements to include nucleic acids and lipids. Flux measurements of DNA, ribosomal RNA, and lipids yielded cellular division rates, ribosome turnover, and lipid metabolism insights, respectively. We also adapted this approach to two-dimensional tissue imaging by DESI-MS in a proof-of-concept study to demonstrate its utility for studying regional differences in metabolism. The future integration of metabolomics and lipidomics into our combined omics approach would be facile, and add unprecedented depth to systems-wide studies involving cellular metabolism. Applied to the regulation of cellular homeostasis in humans, this has the potential to open new avenues for elucidating the etiology of aging, understanding the pathology of age-related diseases, and identifying novel targets for therapeutics.

Keywords: proteostasis, aging, calorie restriction, dietary restriction, advanced glycation end-products (AGEs), combined omics, multi-omics, kinetic proteomics
ACKNOWLEDGEMENTS

First of all, I am grateful to Brigham Young University, and the Department of Chemistry and Biochemistry, for this fantastic opportunity to pursue an exciting new research career. I am especially indebted to my advisor, Dr. Price, for his guidance, support, and teaching during my years of graduate school; I would not have been able to complete this journey without him. Although I am grateful to all of the faculty members of BYU whom I have had the privilege to interact with, I wish to single out for special mention: Dr. Barry Willardson, who was the first to welcome me into his lab, and let me make all kinds of absurd mistakes in order to get my “lab hands” back; and Dr. Jonathan Hill, who taught me the intricacies of RNA sequencing analysis and started me in the R programming language.

I wish to also acknowledge some of my fellow students in the program, who served as great role-models and gave me valuable advice, training, encouragement, and support: Takuma Aoba, Rebecca Plimpton, Andrew Mathis, and Bradley Naylor. In the Price Lab, I was privileged to work with many other wonderful people, who aided me in countless ways in the completion of this project: Monique Spiers, Lavender Lin, and numerous undergraduates who gave of their time in the lab. An additional special thanks goes out to Bradley Naylor, who was my humble partner in crime for much of my time here at BYU; every good aspect of this dissertation in some way bears the mark of his influence on my training.

Of course, no acknowledgement section would be complete without thanking my wife, Robin, and our two children for their constant encouragement, support, and sacrifice during my graduate school years. Finally, I would like to thank Sensei Nicholas Strange, Sensei Moroni Alvarez, and Sensei Hyrum Ornano of Z-Ultimate Self-Defense, who helped me make it across the finish line with my sanity more or less intact.
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................................. ii

ACKNOWLEDGEMENTS ........................................................................................................... iii

TABLE OF CONTENTS ............................................................................................................ iv

LIST OF TABLES ....................................................................................................................... viii

LIST OF FIGURES ..................................................................................................................... ix

INTRODUCTION ......................................................................................................................... 1

Aging ........................................................................................................................................ 1

Proteostasis ............................................................................................................................... 4

Slowing the aging process: Calorie restriction ...................................................................... 12

Measurement of proteostasis to understand the biochemistry of aging ............................... 19

References ............................................................................................................................... 25

Chapter 2 ................................................................................................................................. 40

Title ....................................................................................................................................... 42

Abstract ................................................................................................................................. 42

Introduction .............................................................................................................................. 43

Methods ................................................................................................................................ 47

Animal Handling: .................................................................................................................... 48

Kinetic Proteomic Methods .................................................................................................... 50

Quantitative Proteomic Methods .......................................................................................... 55

Transcriptomics Methods ....................................................................................................... 56
Differences in protein intake under dietary restriction resulted in measurably distinct phenotypes.

Dietary protein levels affect protein homeostasis globally in dietary restriction.

Dietary restriction alters the global proteome post-transcriptionally.

Addition of dietary protein alters cellular division rates in dietary restriction.

Individual functional ontologies may be regulated separately from the global proteome change during dietary restriction.

Ribosomal subunits display differential degradation rates under dietary restriction.

Ribosomal RNA quantities do not correlate with ribosomal proteins.

Discussion.

References.
Spatially distinct changes in isotope ratio are specific to individual lipids and independent of concentration .......................................................................................................................... 168

Turnover calculations highlight differences in metabolism between lipids......................... 175

Discussion................................................................................................................................ 177

Acknowledgements ................................................................................................................. 181

References ................................................................................................................................ 181

Chapter 5.................................................................................................................................... 187

Conclusion.................................................................................................................................. 187

APPENDICES ............................................................................................................................. 191

R Scripts for Data Analysis ..................................................................................................... 191

RNA-Seq data analysis code ................................................................................................. 191

Confidence interval calculator function ............................................................................... 193

P-value and Q-value computation for discerning between selective and global regulation 196

Synthesis and degradation rate constant ontology analyses............................................... 198

Linear Regression Statistics................................................................................................. 208
LIST OF TABLES

Chapter 1:

Table 1. Hallmarks of Aging (Lopez-Otin, et al, Cell, 2011) ......................................................... 3

Table 2. Effects of Calorie Restriction in Various Organisms\textsuperscript{56, 60-71} ................................. 14

Table 3. Impacts of Advanced Glycation End-Products (AGEs)\textsuperscript{86-90} ............................. 18

Chapter 2:

Table 1. Ribosomal subunit protein concentration and turnover rate data. ................................. 81

Chapter 4:

Table 1. Instrumental parameters for DESI image acquisition .......................................................... 156

Table 2. The common name, observed ionic mass, and elemental composition of the 4 major singly charged ions observed in our DESI-MS spectra ......................................................... 164

Appendices:

Table S1. Regression value statistics for Chapter 2 .................................................................... 209

Table S2. Regression value statistics for Chapter 3 .................................................................... 210
LIST OF FIGURES

Chapter 1:

Figure 1: The Proteostasis Network................................................................. 5

Chapter 2:

Figure 1: Proteostasis Model. ................................................................. 44

Figure 2: Experimental workflow............................................................ 47

Figure 3: Mouse phenotype data.............................................................. 65

Figure 4: Protein turnover rate and quantitation comparisons between dietary restriction and
ad-libitum. ........................................................................................................ ...................... 67

Figure 5: Comparisons between dietary restriction and ad-libitum mRNA amounts........ 69

Figure 6: Cellular proliferation rates differed between dietary treatment groups. ........... 71

Figure 7: Experimental versus predicted values for turnover rates, quant, RNA-Seq counts, k-
synth, and k-deg............................................................................................. 73

Figure 8: K-synth and k-deg comparisons of unique proteins from significant ontology groups.
........................................................................................................................................ 77

Figure 9: Ribosomal subunit analysis........................................................ 80

Figure 10: Ribosomal rRNA quantitation by RT-qPCR................................. 83

Figure 11: A model for decreasing global translation in dietary restriction............... 86

Chapter 3:

Figure 1: Mouse phenotype data............................................................... 117
Figure 2: Turnover rate comparisons between dietary cohorts................................. 119

Figure 3: Protein quantitative mass spectrometry comparisons between dietary cohorts. ..... 121

Figure 4: DNA turnover rate measurements................................................................. 123

Figure 5: RNA-Seq gene count comparisons between dietary cohorts. ....................... 125

Figure 6: Experimental versus predicted values for turnover rates, quant, and RNA-Seq counts. ................................................................................................................................. 127

Figure 7: Ontology analysis results............................................................................... 130

Figure 8: Ribosomal subunit data. ................................................................................. 132

Figure 9: Mitochondrial, secreted, and endoplasmic reticulum protein rate and quant data. . 134

Supplemental Data: Matched Protein List Analyses

A. Turnover Rate Data .................................................................................................. 139

B. Quant MS Data ...................................................................................................... 140

C. RNA-Seq Data ....................................................................................................... 141

Chapter 4:

Figure 1: Desorption electrospray ionization (DESI) is an ambient ionization mass

spectrometric imaging technique..................................................................................... 151

Figure 2: The workflow is shown here in chronological order........................................ 153

Figure 3: Urine and blood deuterium enrichments for the mice used in this study were

combined to measure an average body water turnover for these mice......................... 154

Figure 4, A-D: Lipid fragmentation from brain samples compared to standards.............. 160
Figure 4A: Fragmentation of 303.25 m/z peak and arachidonic acid standard. .................. 160
Figure 4B: Fragmentation of 327.3 m/z peak and docosahexaenoic acid standard. .......... 161
Figure 4C: Fragmentation of 834.6 m/z peak and phosphatidylserine standard. .......... 162
Figure 4D: Fragmentation of 885.6 m/z peak and phosphatidylinositol standard. ........... 163

Figure 5: Comparison of experimental and optimized simulations of labeled and unlabeled spectra(A). .................................................................................................................... 166

Figure 6: Stable carbon hydrogen bonds which could incorporate deuterium during metabolic labeling ......................................................................................................................... 168

Figure 7: Concentration/isotope ratio images for phosphatidylserine. ............................ 170

Figure 8, A-C: Concentration as the relative ion intensity (right hand column), and isotope ratio images scaled either within the timepoint (right column) or across all the timepoints (middle column) for arachidonic acid (AA, S4A), docosahexaenoic acid (DHA, S4B), and phosphatidyl inositol (PI, S4C). ......................................................................................... 171

Figure 8A: Concentration/ Isotope maps for AA ............................................................... 172
Figure 8B: Concentration/ Isotope maps for DHA........................................................... 173
Figure 8C: Concentration/ Isotope maps for PI............................................................... 174

Figure 9: Deuterium enrichment of all four lipids included was compared in different areas of the brain. .................................................................................................................... 176

Figure 10: Signal purity tested by spatial distribution of neutromers............................... 180
INTRODUCTION

Aging

At its most basic, aging is the process of becoming older. This simple definition, however, belies the complex processes that occur as an organism grows and changes throughout life. At its most broad, aging can be applied to describe the entire life cycle, from birth to death, encompassing all of the often-profound metamorphoses that occur in between. Therefore, for simplicity, the term “aging” will be used here to describe those changes that occur after an organism reaches maturity (adulthood) and declines towards the end of its life.

For human beings, this involves a complex series of changes in phenotype that are generally associated with poorer health. Although the rate of the process varies by individual, there are many unifying characteristics: weakening of the body due to loss of bone and muscle mass, increased wrinkling as skin loses flexibility and depth, decay of eye sight and hearing, loss of cognitive function, decreased ability to combat infections, and a dramatic increase in the risk of age-related diseases such as cancer, cardiovascular disease, kidney disease, diabetes, and dementia. Due to an overall increase in lifespan over the past 100 years resulting largely from improvements in healthcare and nutrition, the average human being will now live almost twice as long as at the beginning of the 20th century\(^1\). However, the cost is an aging society that is dependent upon increasing amounts of health care. Alzheimer’s disease, a form of dementia that primarily affects the elderly, alone was affecting an estimated 50 million people worldwide in 2018, costing over one trillion U. S. dollars; these numbers are expected to rise to over 150 million patients, costing over two trillion U. S. dollars by the year 2030\(^2\). This has been
described as unsustainable³, and has triggered accelerated research into treatments and cures for these age-related conditions.

Basic research into many of these age-related diseases has revealed much about their mechanisms and underlying causes, although there is still much to learn about how they are tied to aging. Aging itself is not a well-understood process, with competing hypotheses about the root causes and mechanisms of aging. In a landmark review article paying homage to Hanahan and Weinberg’s seminal review “Hallmarks of Cancer”⁴,⁵, Lopez-Otin and colleagues detailed “The Hallmarks of Aging”⁶; these are summarized in Table 1. The authors identify initiating changes at the level of an individual cell (genomic instability, telomere attrition, epigenetic alterations, and loss of proteostasis) as likely to be primary drivers of age-related damage, with the other hallmarks being reactions to and complications of these effects.
Table 1 Hallmarks of Aging (Lopez-Otin, et al, Cell, 2011)

<table>
<thead>
<tr>
<th>Hallmark of Aging</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomic instability</td>
<td>The random mutation of DNA as well as environmental damage eventually affects important genes. Abnormal chromosomal segregation and defects in DNA repair machinery may also cause genomic instability.</td>
</tr>
<tr>
<td>Telomere attrition</td>
<td>Telomeres are known to shorten with every cell division, causing the progressive and accumulating loss of the telomere-protecting ends of chromosomes. Eventually, the cell is no longer able to divide.</td>
</tr>
<tr>
<td>Epigenetic alterations</td>
<td>Post-translational modification of histones, DNA methylation, and chromatin remodeling all constitute epigenetic alterations. Certain patterns have been found to correlate with aging.</td>
</tr>
<tr>
<td>Loss of proteostasis</td>
<td>Proteostasis encompasses the array of mechanisms cells use to stabilize and preserve the healthy function of their proteomes. Chaperones, quality control protein machinery, and degradative mechanisms are all involved in proteostasis. Perturbations in proteostasis lead to cellular dysfunction.</td>
</tr>
<tr>
<td>De-regulated nutrient sensing</td>
<td>Dysfunction in nutrient sensing pathways involving the insulin receptor, insulin-like growth factor-1 (IGF-1) receptor, mTOR, AMPK, and/or Sirtuins may accelerate aging.</td>
</tr>
<tr>
<td>Mitochondrial dysfunction</td>
<td>Electron leakage increases and ATP generation decreases as respiratory chain efficiency diminishes due to increasing age. Increasingly damaged mitochondria turn over more slowly as biogenesis and mitophagy both decrease.</td>
</tr>
<tr>
<td>Cellular senescence</td>
<td>Stable arrest of the cell cycle coupled with a corresponding pro-inflammatory &quot;senescence-associated secretory phenotype&quot; that may contribute to aging. Increased age comes with increasing numbers of senescent cells.</td>
</tr>
<tr>
<td>Stem cell exhaustion</td>
<td>Stem cells decrease in number and undergo fewer divisions with increasing age. This causes a decline in the regenerative potential of tissues.</td>
</tr>
<tr>
<td>Altered intercellular communication</td>
<td>Neuroendocrine signaling becomes deregulated as age increases, causing chronic inflammation and declining immunosurveillance against pathogens and malignancies. Additionally, aging effects spread &quot;contagiously&quot; from one tissue to another.</td>
</tr>
</tbody>
</table>
Proteostasis

As a proteomics research group, we have been primarily interested in the loss of proteostasis as a unifying feature of aging across species. Proteostasis describes the balance of biochemical pathways within the cell controlling protein synthesis, folding, processing, transport, and degradation that maintain the appropriate concentrations of proteins in the correct proportions necessary for the proper functioning of the cell. In short, proteostasis is the process whereby the proteome of the cell, all of the proteins within the cell at any single time, is maintained. As its definition implies, proteostasis is very complex, encompassing many different processes. A review by Jonathan Labbadia and Richard I. Morimoto in 2015 laid out a framework to organize the various cellular processes into a “proteostasis network” (PN) composed of four basic areas: protein synthesis, protein folding, disaggregation, and degradation. This framework is shown in Figure 1 below, along with some of the major cellular proteins and structures involved in the proteostasis network. Here, I adopt this organizational scheme to simplify the discussion of the basics of proteostasis and its regulation.
Figure 1: The Proteostasis Network. The maintenance of a functioning proteome requires thousands of participants within the cell, which are collectively termed the proteostasis network (PN). In the figure above, the major components of the PN are categorized and listed adjacent to the proteostatic processes in which they participate.
The first process involved in proteostasis is the synthesis of proteins, in which messenger RNA (mRNA) is translated by the ribosome (Figure 1). This process necessitates the cooperative action of numerous proteins, which provides opportunities for multiple levels of proteostatic regulation. The first step in protein translation, initiation, begins with the binding of the eIF4F complex to the mRNA in a cap-dependent process that recruits the 43S pre-initiation complex, which includes the 40S ribosomal subunit. After locating the start codon (AUG) of the mRNA transcript, the 60S subunit is recruited, and translation begins.

Under conditions of stress, the initiation of protein translation can be modulated by components of the cellular integrated stress response system (ISR). eIF2α kinases phosphorylate eIF2α (part of the 43S pre-initiation complex) in response to a variety of conditions (e.g. Gcn2 activates and phosphorylates eIF2α in response to amino acid deprivation); this inhibits the formation of the 43S pre-initiation ternary complex, and reduces cap-dependent protein translation. Cap-dependent translation is also inhibited through eIF4-BP, which under high nutrient conditions is phosphorylated by mTOR; if the activity of mTOR goes down (e.g. in nutrient stress), non-phosphorylated eIF4-BP binds to eIF4 and inhibits the formation of the 43S pre-initiation complex. Simultaneously, the transcription factor ATF4 becomes up-regulated under stress, as part of a cap-independent protein translation process that increases under conditions of cap-dependent translation repression. This transcription factor then promotes the expression of a number of genes that aid the cell in dealing with stress.

Continuation of protein synthesis past the initiation step depends on transfer RNA (tRNA) and the ribosome. Transfer RNA, the bridge between nucleic acids and proteins, is necessary for carrying the proper amino acids to the ribosome for addition to the growing peptide
chain. tRNA is typically heavily modified with diverse chemical modifications, some of which can be regulatory. For example, under oxidative stress conditions, a 5-methylcytosine modification in the anticodon of a tRNA that codes for leucine results in the preferential translation of mRNAs enriched for UUG\textsuperscript{12}. Other codon-specific modifications have been observed to change the rate of protein translation, illustrating that changing the cytosolic pool of tRNA can alter the kinetics of protein synthesis\textsuperscript{8}. Reduction of the tRNA pool occurs through endolytic cleavage under oxidative stress, which has the overall effect of slowing protein translation\textsuperscript{13}. Protein fidelity depends on the correct joining of amino acids to their corresponding tRNAs, the responsibility of the amino-acyl tRNA synthetases. Stress conditions can increase the misacylation of methionine tRNAs, decreasing translational fidelity\textsuperscript{14}; however, translational fidelity has been seen to inversely correlate with translational speed\textsuperscript{15}, which decreases under stress conditions. tRNA metabolism thus offers many opportunities for proteostasis regulation.

As the ribosome is the only protein complex capable of translation, it forms a nexus for general synthesis control of proteins in the cell. The cell spends an extraordinarily large share of its resources (60\% of total mRNA transcription in a yeast cell\textsuperscript{16}) on the generation of ribosomes: four transcripts of rRNA and \~80 ribosomal proteins per ribosome, with over 200 proteins involved in ribosomal biogenesis. Assembly is tightly controlled and coordinated, within the nucleus as well as in the cytosol, with final quality checking leading to targeting of defective pre-ribosomes for degradation\textsuperscript{17}. Regulation of ribosome biosynthesis is subject to signal transduction pathways that respond to stresses in the cellular environment that necessitate adjustment of the proteome. For example, the TOR pathway responds to nutrient stress by
repressing translation of ribosomal proteins (in mammals) and transcription of rRNA and ribosomal protein genes in yeast\textsuperscript{16}.

Regulation of the mature ribosome can be accomplished in part through post-translational modification (PTM) of the ribosomal proteins within the subunits. Although our understanding of the full range of PTMs on the ribosome remains incomplete, notable examples have been characterized\textsuperscript{18}. Phosphorylation of the S6 ribosomal protein by S6 kinase (S6K), the activity of which is largely regulated by mTOR, demonstrates that a change in nutrient state is reflected in the PTM state of the ribosome, presumably for regulatory purposes (the precise effect of the S6 phosphorylation is not completely understood)\textsuperscript{18}. The phosphorylation and subsequent release of L13 from the ribosome can be induced by the activation of the innate immune response; the increase in L13 concentration in the free ribosomal protein pool serves as a downstream regulatory signal\textsuperscript{19}. The phosphorylation of certain ribosomal proteins has been shown to correlate with Parkinson’s disease\textsuperscript{20}, suggesting that some PTMs of the ribosome may lead to disease states, and reinforcing the potential of ribosomal regulation through PTMs.

Ubiquitylation, the attachment of ubiquitin to lysine residues, is another common PTM found on ribosomal proteins; the level of ubiquitylation of the ribosomal proteins is dynamic, and changes with processes such as the cell cycle and the unfolded protein response (UPR), when specific proteins need to be upregulated while translation is globally slowed\textsuperscript{18}.

The process of protein translation is constantly monitored while in progress, for example by the ribosomal quality control complex (RQC), which reacts to ribosomes that have stalled during translation; their defective polypeptide products are tagged for degradation, and the ribosome is “rescued” through splitting into its subunits\textsuperscript{21}. In other situations when the ribosome is damaged or malfunctioning and needs to be recycled, the cell engages in ribophagy, the
degradation of ribosomes through autophagy (the process of autophagy is discussed in more detail below)\textsuperscript{22}. This results in the degradation of the entire ribosome; however, there is evidence that ribosomal proteins can be replaced piece-meal in the ribosomal subunits, ensuring high ribosome quality without the need to sacrifice the entire complex\textsuperscript{23}. Maintenance of the ribosome in this way has been shown to respond to dietary signaling\textsuperscript{23}, indicating regulation in response to cellular needs in order to maintain proteostasis.

Properly functioning ribosomes are also important for proteostasis because the kinetics of translation affect the proper folding of proteins. The ribosome is known to vary its speed as it moves along messenger RNA, going most slowly at rare codons\textsuperscript{24}. During these slow-downs, it is thought that nascent peptides emerging from the ribosome have more time to engage in proper folding before subsequent domains are translated\textsuperscript{25}; this may be necessary if further sections of the protein sequence promote alternative folding or aggregation\textsuperscript{26,27}. Proper folding of proteins is also promoted by chaperones, which associate with the ribosome and engage in co-translational folding, as well as interact with partially folded peptides elsewhere to promote proper folding post-translationally\textsuperscript{7}. For some proteins, transport of the partially folded state to complexes called chaperonins (HSP60 family) is necessary before the protein can assume its properly folded state and function\textsuperscript{28}.

Chaperones are necessary components for maintaining a properly folded proteome post-translationally. Some proteins function properly only in a narrow range of stability, and are sensitive to changing cellular conditions and stresses\textsuperscript{29}. Other proteins may have inherently disordered regions that are prone to aggregation, or may only be stable when bound to the proper substrate or intracellular surface\textsuperscript{29}. The proper function of such proteins relies on the constant monitoring and maintenance provided by the chaperone/chaperonin network of the cell.
Misfolded proteins within the cytosol are primarily recognized and bound by the small heat shock proteins (sHSP’s)\textsuperscript{30}, which prevent their premature aggregation before other chaperones such as HSP70, HSP90, and DNAJ-HSP40, along with numerous co-chaperones, can attempt to refold the proteins into their functional states\textsuperscript{31}. However, some proteins that become misfolded aggregate, which is generally deleterious to cellular health. Neurodegenerative diseases such as Alzheimer’s disease are associated with the accumulation of toxic insoluble protein aggregates within cells\textsuperscript{7}. Chaperones play a beneficial role here as well, either in the attempted disaggregation of protein aggregates\textsuperscript{32}, or by sequestering such aggregated proteins into more benign structures called aggresomes for later processing\textsuperscript{33}.

Proteins eventually outlive their usefulness, become damaged, aggregate, or undergo recycling for raw materials. The two main processes used by the cell to degrade proteins are the ubiquitin-proteosome system (UPS) and autophagy\textsuperscript{34}. A coordinated system of E1 ubiquitin-activating enzymes, E2 ubiquitin-conjugating enzymes, and E3 ubiquitin ligases are tasked with identifying misfolded proteins, and tagging them with ubiquitin\textsuperscript{35}. These proteins are then shuttled to the proteasome for degradation. The 26S proteasome consists of approximately 31 subunits, which are organized into a 20S barrel-shaped core capped by one or two 19S regulatory units responsible for recognizing ubiquitinated proteins\textsuperscript{36}. These proteins are then processed through the proteasome, emerging as separated amino acids; the ubiquitin tags are also removed for recycling by de-ubiquitylating enzymes (DUBs)\textsuperscript{35}.

For large protein complexes, organelles, and protein aggregates too large for the proteasome, degradation proceeds through autophagy\textsuperscript{34}. There are three forms of autophagy: The first, chaperone-mediated autophagy, requires that a chaperone, HSPA8/HSC70, recognize the degradation target, transfer it to the lysosome, and then mediate the engulfing of the target by
the lysosome through binding of a lysosomal-associated membrane protein (LAMP2A)\textsuperscript{34}. The second, microautophagy, is characterized by direct engulfment of substrates by the lysosome through small invaginations of the lysosomal membrane\textsuperscript{37}. The third kind, and the one usually referred to simply as “autophagy,” is macroautophagy, in which degradation substrates are encapsulated in a double-membrane structure called the autophagosome, which later fuses with the lysosome\textsuperscript{38}. In all cases, the degradation of substrates is accomplished through the action of lysosomal hydrolases\textsuperscript{34}. Autophagy thus serves as a means to not only clear the cytosol of unwanted, damaged, or possibly toxic structures, but also to obtain new molecular building blocks for cellular components\textsuperscript{34}.

Breakdowns in any of the component processes of the proteostatic network can result in disease. Mutations in chaperone-coding genes have been documented as causing a wide range of conditions: Defects in heat-shock proteins can cause cataracts (\(\alpha\)-crystallin\textsuperscript{39}), spastic paraplegia (HSP60\textsuperscript{40}), myopathy (\(\alpha\)-crystallin\textsuperscript{41}), cardiomyopathy (\(\alpha\)-crystallin\textsuperscript{41}), Charcot-Marie-Tooth disease (HSP22\textsuperscript{42}, HSP27\textsuperscript{43}), and may also contribute to Parkinson’s disease (Mortalin\textsuperscript{44}). Malfunctions in other chaperones are associated with spastic ataxia (Sacsin\textsuperscript{41}), cardiovascular disease (HSC70\textsuperscript{45}), motor neuropathy (HSJ1\textsuperscript{46}), and myopathy (DNAJB6\textsuperscript{47}). Further downstream, mutations in protein degradation pathway components are associated with amyotrophic lateral sclerosis or ALS (ubiquilin-2\textsuperscript{48}, p62\textsuperscript{49}), spinocerebellar ataxia/Machado-Joseph disease (ataxin-3\textsuperscript{50}), Parkinson’s disease (Parkin\textsuperscript{51}, RME8\textsuperscript{52}), Paget disease (VCP\textsuperscript{53}, p62\textsuperscript{49}), frontotemporal dementia (VCP\textsuperscript{53}), Nakajo-Nishimura syndrome (PSMB8\textsuperscript{54}), and Angelman syndrome (UBE3A\textsuperscript{55}). This is just a partial listing of known mutation-driven conditions, and does not include diseases resulting from other preconditions or causes. However, they illustrate the wide range and severity of symptoms resulting from the malfunction of
proteostatic machinery, and underscore the importance of proper proteome maintenance to the overall health of the cell and organism. Therefore, it is important to understand how cells control and regulate proteostasis. Manipulation of the proteostasis network control nodes within the cell using chemical interventions may lead to treatments for many age-related diseases that are currently difficult to manage or cure. Improvement of overall proteostasis may also lead to extension of health and lifespan in organisms.28

**Slowing the Aging Process: Calorie Restriction**

Calorie restriction is the most reliable method known to researchers for extending lifespan and improving overall health in laboratory animals. Briefly, calorie restriction is the reduction in overall consumption of calories without causing malnutrition. Although the reduction of food intake has been described to increase lifespan since ancient times, the first scientific study on calorie restriction without malnutrition was performed in rats in the early 1930s by McCay, C. M. et al. Since this initial study, calorie restriction experiments have been repeated in a wide variety of model organisms (Table 2). In rodents, it has been shown that a significant reduction of daily caloric intake (up to 60%) can increase lifespan by as much as 50%. While the specifics of these experiments vary, the common thread is that health and longevity are improved in many different model organisms, suggesting that the mechanisms behind calorie restriction are evolutionarily ancient and highly conserved. Calorie restriction thus represents one of the best phenomena for the study of aging.

Major findings from calorie restriction experiments in various organisms are summarized in Table 2, below. Among the benefits in all of the organisms listed are extended lifespan and/or
improved health. However, the magnitude of these benefits depends on modulating factors such as genotype, dietary components, and the specifics of various different calorie restriction regimens (Table 2 – Modulating Factors). Some genotypes of yeast, fruit fly, and rodent have been shown to react less well to calorie restriction (even negatively) relative to the wild-type (Table 2 – Exceptions), and closely related species to these laboratory models also vary in their responses to calorie restriction. While these variations complicate the study of calorie restriction, there is also opportunity to discover more about the mechanisms of calorie restriction by studying the biochemical differences between varying responses.
Table 2 Effects of Calorie Restriction in Various Organisms

<table>
<thead>
<tr>
<th>Organism</th>
<th>Calorie restriction effects</th>
<th>Modulating factors</th>
<th>Exceptions</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeast (S. cerevisiae)</td>
<td>Lifespan extension, postponement of senescent phenotype</td>
<td>Amino acid concentration in media, genotype</td>
<td>Some genotypes respond negatively to DR (Schleit, J. et al. 2014)</td>
<td>Jiang, J. C. et al 2000</td>
</tr>
<tr>
<td>Nematode (C. elegans)</td>
<td>Lifespan-extension, decreased fecundity, improved stress resistance, increased metabolic rate</td>
<td>Diet quality and composition</td>
<td></td>
<td>Houthooft, K. et al 2002;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Garsin, D. A. et al 2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Grandison, R. C. et al 2009</td>
</tr>
<tr>
<td>Rodent: Mouse (M. musculus), Rat (R. norvegicus)</td>
<td>Lifespan extension, prevention of age-related diseases (neoplasms, diabetes, cardiovascular disease, neurodegenerative disorders), suppression of chronic inflammatory responses, and maintenance of more youthful physiology, decreased fecundity</td>
<td>Age of onset, genotype</td>
<td>DBA2, majority of ILSXISS (RI) strains of mice</td>
<td>Agarwal, B. et al 2011;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Forster, M. et al 2003;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Liao, C. et al, 2010</td>
</tr>
<tr>
<td>Primate: Rhesus monkey (M. mulatta)</td>
<td>Fewer disorders of aging (diabetes, cancer, cardiovascular disease, brain atrophy), improved survival rates, lifespan extension, lower fasting blood glucose levels</td>
<td>Age of onset, gender, body weight and fat mass, food intake, diet quality</td>
<td>Other monkey species respond less well to DR (e.g. spider monkeys)</td>
<td>Mattison, J. et al 2017;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Masoro, E. J. et al 1996;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Colman, R. J. et al 2009</td>
</tr>
<tr>
<td>Human (H. sapiens)</td>
<td>Lifespan-extension data are lacking, although studies in long-lived populations such as Okinawans suggest longevity benefits, small studies show evidence of improvement in insulin sensitivity, reduced cardiovascular disease markers, and lower body temperature</td>
<td>Potentially many, including age of onset, gender, body mass, diet composition, and diet quality, but data are lacking</td>
<td>Unknown</td>
<td>Heilbronn, L. K. et al 2006;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Heilbronn, L. K. et al 2003</td>
</tr>
</tbody>
</table>
However, despite almost 100 years of active research, the exact causes of the calorie restriction health and longevity benefits remain elusive. Moreover, as these are broad and complex phenotype readouts at the organismal level resulting from a broad dietary intervention that is anticipated to have multiple organ- and tissue-dependent effects, the mechanisms behind calorie restriction are likely to involve an interplay of multiple biochemical processes that may be difficult to succinctly describe. There has been progress: some key pathways and proteins have been identified as actors in the phenomenon of calorie restriction.

The TOR (target of rapamycin) pathway has been shown to play a central role in the regulation of longevity in many different model organisms. In response to nutrient signalling (such as dietary protein), cellular energy levels, or stress, activated TOR stimulates protein translation at the initiation and elongation steps; activation of TOR has also been shown to affect protein degradation through inhibition of autophagy. Likewise, the inhibition of TOR can lead to decreased protein translation and increased autophagy in the cell. Rapamycin treatment or knock-downs of components of the TOR pathway have been shown to have lifespan-lengthening effects in a variety of model organisms, including yeast, nematodes, fruit flies, and mice. Decreased TOR signaling downstream of calorie restriction may therefore play an essential role in calorie restriction’s lifespan extension benefits.

The TOR pathway acts upon and is regulated by other proteins and protein complexes that have also been implicated in the calorie restriction phenomenon. The insulin-IGF1 pathway activates a series of kinases, one of which, Akt, activates TOR; other kinases also act on TOR downstream targets such as S6K, which in turn phosphorylates the S6 ribosomal protein. Downstream of the insulin-IGF1 pathway are the FOXO class of transcription factors, which have also been shown to be key in calorie restriction. Low energy levels in the cell, such as
may result from calorie restriction, can activate the AMPK pathway, which inhibits TOR activity\textsuperscript{73}. The MAPK pathway, which responds to cellular stress, can synergize with TOR signaling through activation of Mnk1\textsuperscript{72}. Nutrient-sensing sirtuins have also been implicated as actors in calorie restriction\textsuperscript{75}, and act broadly in many parts of the cell, including the AMPK pathway\textsuperscript{76}.

A common thread among all of these players is their effect on the regulation of proteostasis. The decrease in protein translation due to lower mTOR signaling, from direct sensing as well as the synergy of the insulin-IGF1, MAPK, and AMPK pathways, is expected to result in several benefits to the cell. As protein translation rate slows, translational fidelity increases\textsuperscript{15}, resulting in higher quality proteins. The slower translational speed also means more time for correct co-translational folding as the ribosome moves along the mRNA transcript\textsuperscript{25}, a process also aided by the up-regulation of chaperone and co-chaperone production triggered by stress-induced transcription factors\textsuperscript{77}. The depression of mTOR activity also leads to upregulation of autophagy\textsuperscript{73}, whereby the cell may improve the quality of its existing proteome through the degradation of damaged organelles and protein aggregates. Thus, the improved proteostasis resulting from the effects of these key players may be the direct cause of the health and longevity benefits of calorie restriction\textsuperscript{15,78,79}.

Effective calorie restriction regimens also turn out to be more complex than a simple reduction of calories. Reports in the literature show that the benefits of calorie restriction can be attenuated by shifting the balance of macronutrients in the diet, such as fat, carbohydrates, and protein\textsuperscript{80}. In particular, high protein diets have been shown to be less beneficial under calorie restriction\textsuperscript{80,81}. Dietary protein activates mTOR and other nutrient sensors in the cell, and may thus trigger pathways that conflict with those providing calorie restriction benefits\textsuperscript{80}.
Another well-studied dietary attenuator of calorie restriction benefits are advanced glycation end-products (AGEs), modified proteins formed from the reaction between proteins and sugar\textsuperscript{82}. These are commonly found in processed foods in Western diets, and have been connected to a variety of diseases normally associated with aging (Table 3). One of the most important researchers in this field, Dr. Helen Vlassara, has been researching the effects of AGEs for over 30 years, and has contributed significantly to the understanding of the role of dietary AGEs in diabetes, cardiovascular disease, Alzheimer’s disease, and other age-related diseases. Dr. Vlassara’s group published a study in 2008 showing that AGEs largely negate the normal health and lifespan-extending benefits of calorie restriction\textsuperscript{83}. While the mechanism behind this phenomenon is not completely understood, AGEs are known to interact with receptors of AGEs (RAGE), which have downstream effects of increasing inflammation\textsuperscript{84}. Chronic inflammation is known to cause health problems\textsuperscript{85}, and may be key in the cancellation of calorie restriction benefits. The potential of AGEs to act as a singular dietary signal modifying the proteostasis network in calorie restriction was the driving idea behind our AGE-modified calorie-restriction experiments (Chapter 3): a systems-wide view provided by our combined omics approach held promise as a way to elucidate pathways that lead to CR benefits, as well as reduce them through AGEs.
Table 3. Impacts of Advanced Glycation End-Products (AGEs)\textsuperscript{86-90}

<table>
<thead>
<tr>
<th>Disease</th>
<th>Role of AGEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular disease</td>
<td>AGEs act as a cross-linking agent that causes stiffening of blood vessel walls and entrapment of LDLs. AGEs also promote oxidation of LDLs and conditions leading to atherosclerosis, a disease that can lead to myocardial infarction.</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Diabetic conditions lead to higher-than-normal levels of serum AGEs; these contribute to complications of diabetes, such as nephropathy, retinopathy, periodontitis, and neuropathy.</td>
</tr>
<tr>
<td>Alzheimer's disease</td>
<td>AGEs can be found in the amyloid plaques and neurofibrillary tangles associated with AD; their presence explains the extensive protein cross-linking, induction of oxidative stress, and neuronal cell death observed in the disease.</td>
</tr>
<tr>
<td>Aging</td>
<td>AGEs cause increased oxidative stress and chronic inflammation, which over time may facilitate the emergence of age-related diseases such as diabetes, cardiovascular disease, and Alzheimer's disease.</td>
</tr>
</tbody>
</table>
Calorie restriction, therefore, is not expected to be a simple one-dimensional dietary signal that should yield a single set of predictable biochemical changes. Rather, it suggests that aging can be adjusted through a variety of factors in the diet. Indeed, other studies have shown that aging benefits can be also be gained through intermittent fasting\textsuperscript{91}, high carbohydrate-low protein diets\textsuperscript{80}, and changes in feeding times and frequency\textsuperscript{92}. This vastly expands the usefulness of calorie restriction as a model for studying mechanisms of aging, as each perturbation to the basic regimen should affect a variety of pathways in different ways, all of which will affect aging. Analysis of proteostasis measurements and comparison between these dietary adjustments is expected to be far more efficient in discerning mechanisms involved in aging than studying basic calorie restriction alone.

**Measurement of Proteostasis to Understand the Biochemistry of Aging**

Because of its broad and complex nature, proteostasis is difficult to measure. How does one quantify all of the processes necessary to maintain the entire proteome at any given time? This will necessarily involve omics technologies, as, at present, these are the only methods that attempt to simultaneously measure and characterize entire pools of biological molecules, such as proteins, nucleic acids, and metabolites. At present, the goal of being able to completely describe the intricacies of metabolism in the cell \textit{in vivo or in vitro} remains elusive. But important gains have been made.

The identification and quantification of all of the mRNA being produced in a cell at a given time (transcriptomics) serves as a snapshot of all the genes that are active within a cell\textsuperscript{93}. As such, this can serve as a way of measuring gene expression. Methods for measuring the
transcriptome (the entirety of the mRNA within a cell at any given time) include microarrays\textsuperscript{94}, which test for specific sequences in a targeted manner, and sequencing technologies, which identify and quantify mRNA content in a \textit{de novo} fashion\textsuperscript{93}. Because of the direct relationship between mRNA and protein synthesis, studies in gene expression frequently use transcriptomics as a proxy for protein metabolism within a cell\textsuperscript{95}. However, it has been shown that the levels of mRNA and protein concentration do not always correlate well. For example, Steven Gygi and colleagues in 1999 published a study in which they showed that yeast protein expression levels could not be predicted from their mRNA levels, showing that protein concentrations could vary by up to 20-fold with equivalent mRNA transcription\textsuperscript{96}. This is understandable, since protein metabolism is regulated on a variety of levels, and numerous processes will affect the proteome past the stage of transcription\textsuperscript{96}. Nonetheless, transcription is regulated tightly by the cell through epigenetics\textsuperscript{97} as well as a variety of transcription factors\textsuperscript{98}, which may either act long-term or short-term, and thus serves as an initial point of analysis in the description of proteostasis.

To reliably measure the quantity and types of proteins in a cell, one must use proteomics. This can be done very effectively using mass spectrometry; methods for the identification of proteins from the m/z values in mass spectra continue to be developed, but some, such as the “bottom-up” approach used in our study, are now considered routine\textsuperscript{96}. In a “bottom-up” approach, proteins are first degraded to peptides, which are then sequenced by analysis of mass spectrometry fragmentation data; the amino acid sequences are then mapped via established databases to their original proteins\textsuperscript{99}. Methods also exist to measure the quantity of proteins within mass spectrometry samples, which gives the option for measuring levels of protein expression. Although label-free quantitation\textsuperscript{100} was available, to improve precision and accuracy
in the quantitation methods, we incorporated C\textsuperscript{13}-lysine-labeled liver as an internal reference into our quantitative mass spectrometry samples; the ratio of experimental to heavy-labeled peptides vastly simplified the comparison between different mass spectrometry runs\textsuperscript{101}. These measurements enabled us to accurately measure protein concentration changes between experimental conditions, and formed an integral part of our analysis of proteostasis. Such quantitative measurements are relatively standard in the proteomics field, and form the basis for much of the research. However, one of the major weaknesses behind quantitative mass spectrometry lies in the fact that it yields static results; each measurement is a snapshot of a dynamic process, which limits its utility, and must be considered in any interpretation of the data.

Proteostasis governs a system that is in constant flux, as proteins are constantly cycled through the proteostatic network\textsuperscript{29}. Measurement of protein concentrations alone is therefore insufficient to fully understand the processes in proteostasis; one must also measure the flux, or rate of change in protein concentrations. Indeed, protein concentrations may remain constant even as flux increases or decreases, as long as synthesis and degradation rates match. Kinetic proteomics involves adaptations to standard proteomic experiments that enable the calculation of protein turnover rates, which give a measurement of the flux of proteins through a cell’s metabolic pathways. The underlying principle behind kinetic proteomics is to label the newly synthesized proteins in a way that makes them distinguishable from older proteins, typically by using a heavy isotope, and then tracking this change over time to calculate a rate constant\textsuperscript{102}.

We chose deuterated water as our heavy-isotope source, which had the advantage of incorporating deuterium into not only proteins within the cell, but all newly synthesized biomolecules. In theory, this meant that we could calculate turnover rates for any biochemical species of interest, in addition to proteins, and indeed we adapted our mass spectrometry
protocols to calculate turnover rates of DNA (which serves as a proxy for cellular division rate) and RNA (both total RNA and ribosomal RNA) without having to regenerate biological samples. We also explored the kinetics of lipid metabolism within the brain in a proof-of-concept method development using desorption by electrospray ionization mass spectrometry (DESI-MS). While the results were limited due to technical constraints, we established the feasibility of kinetic lipidomics and explored the possibility of expanding kinetic omic methods to two-dimensional tissue imaging using mass spectrometry, which adds the attractive possibility of exploring regional tissue differences in proteostasis within distinctly structured organs (for more details on the DESI-MS method, see Chapter 4).

Omics methods can generate a wealth of data, and are valuable for answering distinct questions about processes within cells, tissues, and organs. However, when combined in a “multi-omics” manner to look at biological systems as a whole, the number of questions that can be answered about metabolism within the cell expands significantly. For example, the study completed by Rendleman, J. and collaborators, published in 2018, combined transcriptomics, translatomics, proteomics, and RNA-protein interactomics\(^{103}\). In this work, the authors studied the temporal changes that occurred in proteostatic regulation in human cervical cancer cells under conditions of cellular stress caused by protein misfolding and oxidative stress. The integration of their data methods revealed intricacies in the temporal regulation of many genes controlled at multiple levels within the cell, including genes involved in the endoplasmic reticulum unfolded protein response, mitochondrial energy production, and alternative splicing of tRNA synthetases. Their Protein Expression Control Analysis (PECA) tool\(^{104}\) enabled them to convert the data from their multiple omic sets into information on protein expression regulation, allowing them to identify novel regulatory stress event signatures. They identified the regulation
of protein synthesis as a key regulatory strategy, which reinforces our desire to incorporate measurement of turnover into our study of proteostasis.

In our study, multi-omics revealed broad levels of regulation of the proteome in calorie restriction, as well as distinct perturbations in the regulation of specific protein functional groups. We used a simplified model of proteostasis (see Chapter 2, Figure 1) to incorporate our kinetic, quantitative, and transcriptomic data in a kinetic model that enabled us to discern regulatory changes in synthesis from changes in degradation for individual proteins and pathways. Specifically, the concentration of a given protein in the cell can be described classically as a balance between synthesis, degradation, and cellular division rates (see Chapter 2, Equation 1). If the cell is at homeostasis, the change in protein concentration can be assumed to be zero, and the turnover rate of the protein is an average of the synthesis and degradation rates (see Chapter 2, Equation 2). When these equations are combined and solved for the synthesis and degradation constants, it follows that the rate constants for synthesis and degradation can be calculated using three measurements: turnover rate (from kinetic proteomics), protein concentration (from quantitative proteomics), and mRNA concentration (from transcriptomics using RNA-Seq) (see Chapter 2, Equations 3 and 4). While this neglects one of the key aspects of proteostasis, namely protein folding and quality control, it nonetheless provides a very solid foundation for measuring and analyzing basic proteome maintenance in the cell, and thus forms the basis for the present study.

In this dissertation, I will first present my research into the method development for the multi-omics approach to measuring proteostasis, and its application to the calorie restriction phenomenon. Next, I will present a study on attenuation of calorie restriction benefits through the addition of dietary advanced glycation end-products (AGEs), which will utilize parallel
methods and analysis to the basic calorie restriction study. And finally, I will report on my contributions to applying kinetic methods to lipidomics, as well as spatially two-dimensional mass spectrometry (desorption by electrospray ionization mass spectrometry, or DESI-MS), methods which in the future may be incorporated into multi-omic studies to yield further insights into the etiology of aging and how we can attenuate the process through therapeutic interventions.
References


Farrer, M. J., DNAJC13 mutations in Parkinson disease. Human molecular genetics 2014, 23 (7),
1794-1801.

Pestronk, A.; Whyte, M. P.; Kimonis, V. E., Inclusion body myopathy associated with Paget
disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein.

Eguchi, K.; Ida, H.; Yoshiura, K.-I., Proteasome assembly defect due to a proteasome subunit
beta type 8 (PSMB8) mutation causes the autoinflammatory disorder, Nakajo-Nishimura

55. Kishino, T.; Lalande, M.; Wagstaff, J., UBE3A/E6-AP mutations cause Angelman

Academy of Sciences 2011, 1215 (1), 138-143.

57. Schäfer, D., Aging, Longevity, and Diet: Historical Remarks on Calorie Intake


95. Fraser, C. M.; Fleischmann, R. D., Strategies for whole microbial genome sequencing and analysis. *Electrophoresis* 1997, 18 (8), 1207-1216.


CHAPTER 2

“Short-Term Dietary Restriction in Mice Elicits Significant Post-Transcriptional Regulation of Proteostasis”

This chapter represents the bulk of my research work as a graduate student. While my research goal was to investigate the biochemical effects of AGEs (advanced glycation end-products) on calorie restriction (CR), it was necessary to establish the baseline effects of CR on our model organism. It was also useful for MS and RNA-Seq method development, as well as testing multiple data analysis methods for the integration and interpretation of the multiple omic data. In these efforts, I was joined by Brad Naylor, a fellow graduate student in the Price Lab. Brad Naylor was responsible for the mouse work for the mice on the high protein diets, and I was responsible for the mouse work on the low protein diets. We pooled our data for this study and both contributed to study design, experimental work, data analysis, and writing up of results. In terms of specific aspects of the project, Brad performed the quant-MS SILAM experiment and analysis, I did the RNA-Seq experiment and analysis, as well as the RNA motif analysis and RT-qPCR experiments and analysis, and both of us did the kinetic proteomics, each on our respective mouse cohorts. The final write-up of the study, which follows, contains much of Brad’s input, including sections of the Introduction, Methods, and Results, wherein it is a bit difficult to delineate our separate contributions, since we worked closely together on several drafts of the research write-up over the years of the study. However, the final data analyses (especially the Global vs. Selective analysis), the interpretation of the results, the preparation of the final draft and figures, and the development of the differential ribosomal subunit degradation model (see Discussion) were all performed by me, as Brad graduated from the lab prior to the fulfillment of the project. We plan to submit a version of this manuscript for publication very
soon, with a comprehensive list of all those who contributed to this study; while my name will be listed first, Brad Naylor will be either co-first or second author.
Abstract

Advanced age is characterized by decreased proteome quality, leading to increased risk of age-related diseases. Dietary restriction is a proven intervention for increasing lifespan and healthspan in laboratory animals; decades of study suggest that these benefits result from an overall improvement in protein quality and maintenance. However, the mechanisms by which the proteome control machinery is adjusted to do so remain incompletely understood. Using cohorts of mice placed on dietary restriction with varying dietary perturbations, we performed an analysis of kinetic and quantitative proteomic data combined with RNA-sequencing data to investigate nodes of proteome control in vivo. Our data show that a general slowdown of protein turnover results from dietary restriction, and that this is the result of shifts in post-transcriptional regulation. Moreover, this global slowdown is complex and can be modulated by dietary protein intake. Our results support a model in which the large and small subunits of the ribosome are degraded at significantly different rates from each other under dietary restriction, perturbing their normal stoichiometric ratio, and thereby decreasing protein translation in the cell. We speculate that this may be the result of sequestration of the small ribosomal subunit in stress granules formed during dietary restriction. Slower translation rates would naturally lead to improvements in translational fidelity, ribosomal maintenance, and co-translational protein folding, which would result in the improved proteome quality expected under dietary restriction conditions.
**Introduction**

One of the hallmarks of aging is a general loss of proteostasis, which is the balance of biochemical processes within a cell dealing with the synthesis, folding, modification, transport, and degradation of the proteins necessary for life. Imbalances in proteostasis are known causes of disease and aging; with an increasingly older population in much of the world, many age-related diseases, such as cancer and neurodegeneration, have reached epidemic proportions\(^1\). Designing successful interventions to reduce or stop the age related loss of health would benefit from a comprehensive understanding of the regulatory processes responsible for maintaining proteostasis in cells. Healthy cells maintain the quality and quantity of their proteins using many finely tuned controls on epigenomic, transcription, translation, post-translational, and degradation processes. Because cells are dynamic, specialized to their location within the body, and responsive to changes in their environment, proteome maintenance is very complex, with subtle differences in regulation for each protein or protein complex.
Figure 1: Proteostasis Model. At its most basic, proteostasis can be conceptualized as an equilibrium between the processes of protein synthesis and degradation (as well as cell division).
Proteomic dynamics within the cell can be described with a simple generalizable model (Figure 1) in which the quantity of each protein in the cell is maintained in a balance between the processes of synthesis, degradation, and cell division. Understanding the regulation of proteostasis therefore requires large-scale measurements of synthesis and degradation. However, according to the model, we can begin to make deductions about the regulation of synthesis and degradation for thousands of individual cellular proteins by integrating only four separate measurements: protein turnover (kinetic proteomics), protein concentration (quantitative mass spectrometry), mRNA concentration (RNA-Seq), and cellular division rate (see Methods: Multi-Omics Data Analysis). This enables calculations of separate synthesis and degradation rate constants for thousands of individual proteins in vivo and a more detailed study of proteostasis.

To study the perturbations in proteostasis that result from aging, we used a mouse model on dietary restriction (DR) as our experimental group to compare with normal ad libitum (AL) controls. Dietary restriction has been shown repeatedly to be the most robust method for increasing lifespan in laboratory animals; it has also been shown to markedly decrease the incidence of age-related diseases\(^2\). However, despite some recent insights and decades of study, the mechanism behind dietary restriction remains controversial. The source of calories in dietary restriction diets also appears to be important, as demonstrated in multiple studies\(^3,4,5\). For example, an increase in dietary protein has been shown to attenuate the lifespan- and healthspan-expansion benefits of dietary restriction in a dose-dependent manner\(^4\). While this mechanism is incompletely understood, signaling through nutrient sensors such as mTOR appears likely\(^4\). This suggests that dietary restriction benefits, rather than resulting directly from bulk energy reduction, may instead result from the activation of a distinct signaling network, one that can be disrupted with conflicting dietary signals, such as increased dietary protein. Our goal in this
study was to investigate the changes to the proteostasis control network associated with this disruption.

We applied a combination of kinetic proteomics, quantitative proteomics, and transcriptomics to investigate how regulation of many individual proteins (~1000) in the liver is affected by dietary restriction with variable levels of dietary protein. Our data point to a general slowing of protein synthesis under dietary restriction as a result of changes in post-transcriptional regulation. Specific groups of proteins including mitochondrial, endoplasmic reticulum, and secreted proteins escaped this global slowdown of protein synthesis due to selective regulation, likely to gain survival benefits and maintain tissue/organ function. This regulation is complex, affecting both protein synthesis and degradation machinery, and is affected by dietary protein levels. We propose a model in which the large and small subunits of the ribosome experience differential rates of degradation that are exaggerated under dietary restriction and high dietary protein, reducing the stoichiometric ratio of the two ribosomal subunits, and consequently decreasing protein translation. Decreased protein translation simultaneously lowers the synthetic and energy burden on the cell, improves protein quality through multiple mechanisms, and ultimately results in the improved cellular and organismal health observed in dietary restriction.
Methods

Figure 2: Experimental workflow. Mice were fed experimental diets for 10 weeks, then labeled with D₂O and sacrificed at various time-points. Liver was collected and prepared for various analyses: RNA was extracted for RNA-Seq, and protein was collected for Kinetic Proteomics, or mixed with SILAC standard for Quantitative Proteomics.
Animal Handling:

Mouse Housing, Genotype, and Diet

All animal procedure protocols were approved by the Institutional Animal Care and Use Committee (IACUC) of Brigham Young University. Housing for mice was provided in a pathogen-free facility for the duration of the experiment. 10-week old C57/BL6 male mice were purchased from Charles River Laboratory, and fed *ad-libitum* for one week after arrival to acclimate to the facility. Mice were then randomly divided into ad-libitum (AL) and dietary-restricted (DR) dietary cohort groups; the mice were housed individually to ensure equal access to food. Low dietary protein (LP) cohort animals were fed NIH31 chow, and high dietary protein (HP) cohort animals were fed Harlan Teklad8604 chow (LP and HP cohorts were purchased and treated at different times for logistical reasons). AL animals had constant access to food, and DR animals were fed daily a pellet of 3g ± .1g in size (65% of expected AL consumption). Mice were weighed weekly during one of the daily feedings of DR animals to monitor health and weight loss.

Metabolic Labeling

At 10 weeks, (except for the 0-day time points) the mice were given an intraperitoneal bolus injection of sterile D$_2$O saline at 35 µL/g body weight. This injection brought the mice up to 5% Molar Percent Excess (MPE) deuterium as previously described$^6$. Mice were then provided drinking water containing 8% MPE to maintain the 5% MPE in the animals’ body water.
Euthanasia and Sample Collection

Mice were anesthetized with CO₂ and then euthanized by cardiac puncture. Mice were then immediately dissected and all tissues except for blood were flash frozen on solid CO₂ and then stored at -80°C. Blood was stored on ice until it could be centrifuged at 800 x g for 10 minutes at 4°C. The centrifugation separated serum and red blood cells, which were stored in separate containers at -80°C.

The LP-AL and LP-DR cohorts had 9 mice each. The HP-AL cohort had 19 mice and the HP-DR cohort had 20 mice. Due to the difference in cohort sizes, sacrifice times were different. Two LP animals from each diet were sacrificed at each of these time points: 1 day, 3 days, 9 days and 27 days after bolus injection, with one animal from each group sacrificed without receiving a bolus injection or any other D₂O labeling. The HP mice had two animals sacrificed at each of these time points: 9 hours, 1 day, 2 days, 4 days, 8 days, 16 days, and 32 days post injection, with the 4, 8 and 32 day time-points each having a third sacrifice; 2 HP-AL, and 3 HP-DR mice were sacrificed without deuterium injection.

Mitochondrial Respiration

Fresh liver tissue was quickly removed from exsanguinated mice and immediately placed in ice-cold mitochondrial respiration buffer (MiR05: 0.5 mM EGTA, 10 mM KH₂PO₄, 3 mM MgCl₂-6 H₂O, 60 mM K-lactobionate, 20 mM HEPES, 110 mM Sucrose, 1 mg/ml fatty acid free BSA, pH 7.1) and trimmed of connective tissue. Tissue was gently separated and homogenized under a surgical scope (Olympus, ST) to particles of approximately 1 mg. Homogenate was then transferred to a tube with chilled MiR05 and 50 µg/ml saponin and rocked at 4 °C for 30 min, then washed in MiR05 at 4 °C for at least 15 min before use. High-resolution
O₂ consumption was determined at 37 °C using the Oroboros O₂K Oxygraph. Before addition of sample into respiration chambers, a baseline oxygen consumption rate was determined. After addition of sample, the chambers were hyperoxygenated to ~350 nmol/ml., and respiration was determined according to the manufacturer’s protocol. Lastly, residual oxygen consumption was measured by adding antimycin A (2.5 μM) to block complex III action, effectively stopping any electron flow and providing a baseline respiration rate.

**Kinetic Proteomic Methods**

*Measurement of MPE*

Aliquots of serum were distilled in 2.0 mL screw cap tubes overnight in a 90°C sand bath, and the distillates collected. The distillate was diluted 1:300 in ddH₂O, and MPE of deuterium was directly measured against a D₂O standard curve using a cavity ring-down water isotope analyzer (Los Gatos Research [LGR], Los Gatos, CA, USA) according to the published method⁷.

*Trypsin Digestion*

Mouse liver tissue was placed in ammonium bicarbonate (ABC) (25 mM, pH 8.5) along with protease inhibitor cocktail (Sigma) and homogenized using a MP Biomedicals FastPrep-24 bead homogenizer at 6 m/s for 60 seconds. Volumes were calculated to give approximately 10 mg/mL protein. Protein concentration was measured using a bicinchoninic acid (BCA) protein assay (Thermo Fisher). 300-500 μg of protein were placed on 30 kDa centrifugal filters (VWR). 100 μL of a concentrated guanidine solution (6 M guanidine HCl, 100 mM Tris-HCl pH 8.5) was added to each sample and centrifuged at 14,000 x g for 15 minutes. This guanidine wash was repeated, and the flow-through discarded. Disulfide bonds were reduced using a 10 mM
dithiothreitol/6 M guanidine HCl/100 mM Tris-HCl (pH 8.5) solution (100 µL total volume) added directly to the filters, with an incubation of 1 hour at 60°C in a sand bath. After 5 minutes of cooling, cysteine sulfhydryl groups were protected by reaction with iodoacetamide (IAM, 20 mM) for 60 minutes in the dark. Afterwards, the samples were centrifuged at 14,000 x g for 15 minutes, and the flow-through was discarded. The samples were washed twice with ABC (200 µL, centrifuged 15 minutes at 14,000 x g). Pierce MS-Grade trypsin was added (1:50 w:w) in 300 µL ABC to each sample, followed by incubation at 37°C overnight. Resulting peptides were eluted by centrifugation at 14,000 x g for 30 minutes, followed by a wash with 100 µL of ABC and an additional centrifugation for 30 minutes at 14,000 x g. Filters were discarded and the filtrate was dried using a speedvac (Sorval); the dried samples were stored at 4°C until use.

**HPLC Fractionation**

Samples were re-suspended in 10 mM ammonium formate (LC-MS grade pH 9.5) before fractionation using high pH C18 High Performance Liquid Chromatography (HPLC), which is orthogonal to the low pH C18 chromatography used in “LC-MS Data Acquisition,” allowing for improved protein coverage. Fractionation was performed using the 1260 HPLC Infinity (Agilent) and the Gemini 50 x 2.00 mm C18 column with 3µm beads and 110 angstrom pore size. Peptides were eluted using a 10 mM ammonium formate (pH 9.5) H₂O/acetonitrile gradient from 3% B to 60% B over 40 minutes flowing at 1 mL/min. Buffer A was 97% H₂O, 3% acetonitrile, 10 mM ammonium formate pH 9.5. Buffer B was 10% H₂O, 90% acetonitrile, 10 mM ammonium formate pH 9.5. 1 mL fractions were collected and pooled into 8 samples by pooling every 8th fraction. For instance, fractions 2, 10, 18, and 26 were pooled into one sample. Pooled fractions were dried with a speedvac, then resuspended in 200 µL of 80% acetonitrile and decanted into a mass spectrometry vial. Samples were again dried with a
speedvac and suspended in 40 µL of 3% acetonitrile 0.1% formic acid (all LC-MS grade) for LC-MS analysis.

**LC-MS Data Acquisition**

Protein identification and kinetic acquisition were performed on the Agilent 6530 Q-ToF mass spectrometer coupled to capillary and nanoflow Agilent 1260 HPLC using the chipcube nano-spray source\(^6\)\(^\text{9}\). Peptides were eluted from the Agilent C18 Polaris chip at 300 nL/min using an H\(_2\)O-Acetonitrile gradient acidified to pH 4 by use of Pierce LC-MS grade formic acid. Buffer A was 3% acetonitrile, 0.1% formic acid. Buffer B was 97% acetonitrile, 0.1% formic acid. The elution gradient was as follows: 0 minutes, 100% A; 0.1 minutes, 95% A; 27 minutes, 40% A; this was followed by high percentage B column washing and low percentage B equilibration. The Agilent 6530 Q-ToF mass spectrometer was run in 2 Ghz high dynamic range mode. Protein identification runs were performed in MS/MS mode using collision-induced dissociation (CID) with nitrogen gas. MS and MS/MS data were collected at a maximum rate of 4 spectra/second with CID fragmentation on the top 10 most abundant precursors. Dynamic exclusion was set to 0.2 minutes. Kinetic acquisitions were performed in MS only mode and collected at 1 spectra/second. MS only mode increases signal intensity, improves signal-to-noise, and gives more scan points per elution chromatogram, greatly enhancing isotopomer analysis accuracy.

Data were also collected on the Orbitrap Fusion-Lumos mass spectrometer. Samples were resuspended in 0.1% formic acid (Pierce LC-MS grade) in H\(_2\)O (Optima grade Thermo Fisher), and analyzed with a Thermo Lumos Tribrid (Orbitrap). Tryptic peptides were separated using a reverse phase C18 column (Acclaim PepMap™ 100) and a Thermo Easy-Spray source.
Mobile phase for the liquid chromatography was 0.1% formic acid in H₂O (Buffer A) and 0.1% formic acid in 80% acetonitrile (Optima grade Thermo Fisher) with 20% H₂O (Buffer B) on an Easy-nLC 1200 HPLC system. Samples were eluted using a gradient of 5% B to 22% B over 85 minutes, 22% to 32% B over 15 minutes, and a wash of 32% to 95% B over 10 minutes, which was held at 95% B for 10 minutes. Sample loading and equilibration were performed using the HPLC’s built in methods. MS only runs were performed using 2400 V in the ion source, 60000 resolution with a scan range of 375-1700 m/z, 30% RF Lens, quadrupole isolation, 8 *10^5 AGC target, and a maximum injection time of 50 ms. MS/MS scans were performed using the same settings as MS only scans, with 3 seconds allowed per MS/MS after each MS scan, using the following filters: peptide monoisotopic peak determination, intensity threshold of 5*10^3, fragmentation of charge states +2 to +6, dynamic exclusion that excluded a peak after being chosen once within 60 seconds, an error tolerance of 10 ppm high and low, and isotopes excluded. The fragmentation scan used an isolation window of 1.6 m/z, CID fragmentation with an energy of 30%, detection in the linear ion trap in Rapid Scan mode with an AGC target of 1*10^4, a maximum injection time of 35 milliseconds, and used the “Inject Ions for All Available Parallelizable Time” option.

Protein Identification

Peak lists obtained from MS/MS spectra were identified using Mascot version 2.2.04, OMSSA version 2.1.9, X!Tandem version X! Tandem Sledgehammer (2013.09.01.1), MS-GF+ version Beta (v10282), Comet version 2016.01 rev. 2 and MyriMatch version 2.2.140. The search was conducted using SearchGUI version 3.2.710. Protein identification was conducted against a concatenated target/decoy version of the Mus musculus complement of the UniProtKB (created September 2016, 16806 (target) sequences); decoy sequences were created by reversing
the target sequences in SearchGUI. The identification settings were as follows: Trypsin, Specific, with a maximum of 2 missed cleavages 10.0 ppm as MS1 and 0.5 Da as MS2 tolerances; fixed modifications: Carbamidomethylation of C (+57.021464 Da); variable modifications: Oxidation of M (+15.994915 Da), Pyrrolidine from Q (--17.026549 Da), Acetylation of protein N-term (+42.010565 Da), Pyrrolidine from E (--18.010565 Da), Pyrrolidine from carbamidomethylated C (--17.026549 Da); fixed modifications during refinement procedure: Carbamidomethylation of C (+57.021464 Da). Peptides and proteins were inferred from the spectrum identification results using PeptideShaker version 1.15.111. Peptide Spectrum Matches (PSMs), peptides and proteins were validated at a 1.0% False Discovery Rate (FDR) estimated using the decoy hit distribution. Post-translational modification localizations were scored using the D-score and the phosphoRS score with a threshold of 95.0 as implemented in the compomics-utilities package.

**Calculation of Turnover Rates**

Identification files and the MS-only mass spectrometry data were analyzed with the DeuteRater software package\textsuperscript{12}, which provided the protein turnover rates used for later analyses. The resulting kinetic data was filtered to remove data with extreme outliers or other issues: First, rates that were greater than 1 or less than 0.03 were eliminated from further analysis, as these rates represented extrapolations outside of the range of rates that could be calculated confidently with the time points used in the experiment. Since the kinetic proteomics data come from curve fits of relevant measurements, all curves with a Pearson’s R\textsuperscript{2} less than 0.5, or with a covariance (standard deviation/rate value) of greater than 0.2 were also removed from further analysis (the standard deviation was divided by the turnover rate for normalization).
Quantitative Proteomic Methods

Quantitative Mass Spectrometry Sample Preparation

A mouse liver labeled according to a Stable Isotope Labeling by Amino Acids in Mammals (SILAM) protocol (Lys6:13C, Cambridge Isotopes) was purchased and homogenized as described in the “Mass Spectrometry Sample Preparation” section above. For each sample, 50 μg of protein from the SILAM homogenate was combined with 50 μg protein from the sample homogenate (with the exception of the diet A samples; since these underwent HPLC fractionation, 150 ug was used instead of 50 ug). Three biological replicates from each dietary group were prepared. Since isotopic shifts caused by D2O labeling creates difficulties in peptide identification, only 0-day and 1-day time point animals (measured from the date of intraperitoneal injection of D2O) were used for this experiment. The mixtures of SILAM and sample homogenates were otherwise prepared identically to those described within the “Kinetic Proteomic Methods” section.

Mass Spectrometry Analysis

Samples were re-suspended to 1 μg/μL concentration in 0.1% formic acid. Samples were analyzed on a Thermo Lumos Tribrid instrument paired with a Thermo Easy-nLC 1200 HPLC. Solvent A was 0.1% formic acid, 100% water; Solvent B was 0.1% formic acid, 80% acetonitrile, 20% water (all solvents were LC-MS grade). Gradient was as follows: 0-45 minutes was 5-45% Buffer B, 45-50 minutes 45-100% Buffer B, 50-60 minutes 100% Buffer B. The MS scans were collected at 60,000 Resolution, with 20 MSMS scans between MS scans. MSMS fragmentation was done with a 1.6 m/z isolation window, and CID fragmentation with 28%
fragmentation energy. Fragment ions were measured in the Linear Ion Trap with the “Normal” resolution setting.

Data Analysis

SILAM data was analyzed using the MaxQuant software package\textsuperscript{13} to quantify peptides and proteins. All data was analyzed at the same time, with each dietary group analyzed as a single experimental group. Constant modifications were carbamidomethylation for cysteine; variable modifications allowed were oxidation of methionine and acetylation of the N-terminus. The identification database was the same as that used for the kinetic analysis. Lys6 was added as a heavy label; all other settings left at default. Unlabeled divided by labeled signal intensities were determined and used to determine amounts of peptide in the sample relative to the SILAM standard. The relative peptide abundances were then compared between samples.

Transcriptomics Methods

RNA-Seq

Following the manufacturer’s protocol, we used the Direct-zol RNA MiniPrep Plus kit from Zymo Research to extract RNA from the liver tissue of all animals in this study. Briefly, 10-20 mg of frozen liver tissue from each mouse was processed using a Trizol extraction, and the resulting RNA frozen at -80°C. Since we wished to identify changes in the transcriptome that resulted only from dietary differences, RNA from three different mice in the same dietary cohort sacrificed at different time points were pooled to form a single sample for RNA-Seq, thus masking any differences resulting from temporal variations. Total RNA was extracted from three pools per dietary cohort (4\textsuperscript{g} of total RNA in 50 \textsuperscript{L} of RNase-free water). Each pool was submitted as a single sample to the BYU Sequencing Center. The BYU Sequencing Center
constructed cDNA libraries after a poly-A pull-down to enrich for mRNA, and then sequenced the libraries using an Illumina HiSeq 2500 sequencing platform. RNA-Seq gene count data were analyzed in R\textsuperscript{14} using the DESeq\textsuperscript{2} package (see Appendices) in Bioconductor 3.3 from Bioconductor.org\textsuperscript{15}, downloaded on October 20, 2016.

\textit{RNA Motif Analysis}

RNA motif analysis was performed using the HOMER motif analysis tool\textsuperscript{16} to detect any translational preference due to RNA sequence motifs in the mRNA. Output from HOMER was processed using in-house R scripts\textsuperscript{14}.

\textit{DNA Turnover Rate Analysis}

Since DNA turnover rate serves as a useful proxy for cellular division rate, we measured the DNA turnover rates for all dietary cohorts by GC-MS using a previously published method\textsuperscript{6}. Briefly, DNA was extracted from 30-90 mg of frozen liver samples using the DNeasy kit (Qiagen, Valencia, CA), and hydrolyzed to deoxyribonucleosides overnight at 37°C using S1 nuclease and potato acid phosphatase. The hydrolysates were reacted with pentafluorobenzyl hydroxylamine and acetic acid, and then acetylated with acetic anhydride and 1-methylimidazole. Dichloromethane extracts were dried, resuspended in ethyl acetate, and analyzed via GC/MS on an Agilent DB-17 column with negative chemical ionization, using helium as a carrier gas and CH\textsubscript{4} as the reagent gas. The fractional molar isotope abundances at m/z 435 (M0 mass isotopomer) and 436 (M1) of the pentafluorobenzyl triacetyl derivative of purine dR were quantified using ChemStation software. Excess fractional enrichment of D\textsubscript{2}O was calculated relative to an unlabeled pentafluorobenzyl triacetyl purine derivative, and fraction new DNA was determined for each sample by comparing excess fractional enrichment to the
calculated theoretical maximum. GraphPad Prism was used to perform least-squares fits to first-order rate curves on plots of fraction new DNA versus time point. Standard deviation and confidence intervals were calculated based on results from two biological replicates (each measured three times) per time point. DNA turnover rates were analyzed and displayed using R scripts\textsuperscript{14}.

**Multi-omics Data Analysis**

*Calculation of $K_{\text{synth}}$ and $K_{\text{deg}}$ Values*

The model shown in Figure 1 suggests that the change in protein concentrations in cells can be simplified as a balance between synthesis rate (which increases protein concentration), and degradation and dilution (cell division) rates (which decrease protein concentration). Equation 1 formalizes this relationship:

**Equation 1:**

$$\frac{d[\text{Protein}]}{dt} = k_{\text{synthesis}}[\text{mRNA}] - k_{\text{degradation}}[\text{Protein}] - k_{\text{division}}$$

Although protein concentration changes can be measured, separate measurements of *in vivo* protein synthesis and protein degradation rates are not currently possible. By measuring individual protein turnover rates in the same tissues used for the change in concentration, it becomes possible to determine individual synthesis and degradation rates for each of the observed proteins. Canonically, turnover can be described as the average of synthesis and degradation rates (Equation 2, below).

**Equation 2:**
If the cells in the biological sample are terminally differentiated, as they are in most tissues of adult mammals, then the division rate can be assumed to be very close to zero, and dropped from Equation 1. At homeostasis, the change in protein concentration over time should be minimal, thus enabling us to set d[Protein]/dt equal to zero in Equation 1. Algebraic rearrangement then enables us to solve for the k-synth and k-deg values for every protein as a combination of turnover rates, and either protein concentration or mRNA concentration (which can be provided using RNA-Seq experiments). These equations are shown below (Equations 3 and 4), and were used in calculations of synthesis and degradation rate constants for this study.

**Equation 3:**

\[
\frac{k_{\text{synthesis}}[mRNA]}{2} = \frac{\text{Protein Turnover} \text{(day}^{-1})}{[\text{Protein}]} = \frac{k_{\text{degradation}}[\text{Protein}]}{2}
\]

**Equation 4:**

Global Versus Selective Regulation Determination

Linear regression analyses were performed on turnover rate (Figure 4), protein concentration (Figure 4), and RNA concentration (Figure 5). The global regression constants were used to calculate expected values for each protein observed in the low protein dietary-restricted (LP-DR) and high protein dietary-restricted (HP-DR) cohorts relative to the “control” diets: low protein ad-libitum (LP-AL) for LP-DR, and high protein ad-libitum (HP-AL) for HP-
DR. This was done to identify protein specific regulation versus global regulation of the entire proteome during low protein and high protein dietary restriction. A paired t-test between predicted and experimental values was then analyzed using R scripts\textsuperscript{14} (see Appendices) for each individual protein in LP-DR vs. LP-AL and HP-DR vs. HP-AL for turnover rate, quantitative MS, and RNA-Seq data sets, generating a list of p-values for each analysis. To reduce false positives, a q-value test was run in R using the qvalue package\textsuperscript{17} (see Appendices) for each analysis, with an FDR level set to 0.01. Any proteins designated as not significantly different were excluded from downstream analyses. A threshold of +/- 20\% change was used to further screen proteins that were significantly different from their predicted values; these were divided into up- and down-regulated categories, respectively. Volcano plots generated using R (Version 1.1.463)\textsuperscript{14} were used to display the results of these tests (Figure 7).

**Gene Ontology Analysis**

Ontology groupings were determined using DAVID\textsuperscript{18,19}. Proteins/gene lists were obtained from the Global Versus Selective Regulation Determination analysis above. Ensembl IDs from RNA-Seq data were translated to UniProt accession numbers prior to ontology analysis. Each category was input into DAVID\textsuperscript{18,19} against a background comprised of all protein ID’s in that particular analysis (i. e. combined from up-regulated, down-regulated, and not significantly different groups). Ontology groups with Benjamini-Hochberg coefficients less than or equal to 0.05 were retained for analysis; all others were disregarded. Box plots generated using R (Version 1.1.463)\textsuperscript{14} were used to compare significant ontology groups (see Figure 8).

**Ribosomal Subunit Analysis**
Turnover rate and quantitative MS data for all of the ribosomal proteins was extracted from the larger data sets, and separated into two categories based on whether the ribosomal protein was part of the large subunit (60S) of the ribosome or the small subunit (40S) of the ribosome. A Mann-Whitney non-parametric rank order test\textsuperscript{20} was performed using Microsoft Excel (version 14.7.7) on turnover and quantitative MS data to detect whether the large subunit was turning over significantly faster or if its concentration was significantly lower than the small subunit. One-tailed p-values of 0.05 or smaller indicated significance.

Ribosomal Subunit Quantitative PCR (RT-qPCR)

cDNA samples were prepared according to the manufacturer’s protocol using the iScript cDNA Synthesis Kit from Bio-Rad; sample RNA came from three biological replicates from each dietary cohort (the RNA samples were randomly selected from previously prepared RNA used for RNA-Seq; for preparation method, see RNA-Seq method section). The qPCR samples were prepared in a 96-well qPCR plate using SYBR Green Master Mix from Bio-Rad. Primers were ordered from Eurofins for the following targets: 28S rRNA (Forward primer: 5’-GCCTTGGATTGGTCCACCCAC-3’; Reverse primer: 5’-AGGGATAACTGCTTGTCG-3’), 18S rRNA (Forward primer: 5’-CTTAGAGGGACAAGTGGCG-3’; Reverse primer: 5’-ACGCTGAGCCAGTCAGTGA-3’), Vault RNA (Forward primer: 5’-GCTGAGCGG TartTGCACA-3’; Reverse primer: 5’-GTCTCAACAAACACTCATG-3’), and TATA (Forward primer: 5’-ACAGCTTTCCACCTTATGCT-3’; Reverse primer: 5’-GATTGCTGACTGAGGCCTGC-3’). The qPCR reaction was run on a QuantStudio5 qPCR machine (Applied Biosystems, ThermoScientific) using the standard settings for SYBR green chemistry (QuantStudio Design&Analysis software, version 1.4.3); each biological replicate was
run in duplicate. Data were analyzed using qbase+ software, version 3.2 (Biogazelle, Zwijnaarde, Belgium – www.qbaseplus.com).
Results

Differences in protein intake under dietary restriction resulted in measurably distinct phenotypes.

Mouse weights were monitored throughout the dietary interventions in order to monitor animal health and gather data on overall phenotypic effects of the various diets. Shown in Figure 3 below, there was no statistical difference between the two ad-libitum diets, suggesting that the level of protein in the diets did not affect weight gain in at-will feeding. The dietary-restricted cohort on the low-protein diet (LP-DR) initially showed a delay of 3 to 4 weeks before gaining weight, but afterwards seemed to gain weight at the same rate as the ad-libitum (LP-AL and HP-AL) cohorts. This suggests that the dietary restriction regimen, after an initial adjustment period in which the metabolism of the organism shifts to accommodate lower nutrition, does not by itself seriously impede normal growth in the mouse. In contrast, mice on HP-DR diets suffered significant weight loss when they were subjected to dietary restriction. Over the course of two weeks, the mice lost on average 20% of their body mass, and while they eventually began to gain weight at a rate similar to the LP-DR mice, they never fully recovered all of their initial weight loss. This suggests a substantial difference in the metabolism of LP-DR versus HP-DR animals in this experiment, and how they react to dietary restriction. Such weight loss has been observed in previous studies, and provides verification that our dietary restriction interventions were successful.

The mitochondrial respiration data that we collected immediately after sacrifice also showed substantial differences between the low and high protein diets. Pair-wise t-tests were conducted between the ad-libitum cohorts, the dietary-restricted cohorts, and between the
dietary-restricted cohorts and their ad-libitum counterparts in low and high protein diets. While the differences between LP-DR and LP-AL cohorts for D (ADP) and S (Succinate) testing were only statistically suggestive (p < 0.10), the HP-DR was significantly lower than both HP-AL and the LP-DR in all categories of the mitochondrial measurement, suggesting that mitochondria use less oxygen in high protein dietary restriction. This does not translate to a decrease in efficiency, however; the ratios of D/GM and S/GM for the high protein diets are comparable, if not higher than the low protein diets (Figure 3). No statistically significant differences in efficiency were detected between the diets, which suggests that higher protein levels do not impair mitochondrial respiration; that the mitochondrial respiration in the high protein dietary restriction cohort is lower in absolute terms may be due to modified expression of oxidative phosphorylation pathway proteins (see “Individual functional ontologies may be regulated separately from the global proteome change during dietary restriction” section below).
Figure 3: Mouse phenotype data. Mouse weight and mitochondrial respiration data are shown for the four dietary cohorts. The upper left panel shows the weekly mouse weight data over the 10 weeks of dietary treatment. Mouse weights were normalized to starting weights before averaging. The ad-libitum cohorts, LP-AL and HP-AL, were significantly different from the dietary restriction cohorts, LP-DR and HP-DR, which in turn were significantly different from each other. The bottom two panels are bar graphs displaying the mitochondrial respiration data of the dietary cohorts. “GM” is the mitochondria at rest with oxygen, “D” is the addition of ADP and oxygen to show the function of the Electron Transport Chain (ETC) excluding complex 2, and “S” shows complex 2 added to the rest of the ETC. The lower right panel displays the relative efficiency of the ADP effect (D/GM) and the Succinate effect compared to mitochondrial leak (GM). All error bars represent 95% confidence intervals. Horizontal bars and asterisks indicate significant differences.


Dietary protein levels affect protein homeostasis globally in dietary restriction.

Using our kinetic proteomics method, we measured the protein turnover rates for over 1000 mouse liver proteins for all of the dietary intervention cohorts. We found that on average, the turnover rates of the proteins slowed significantly in both LP-DR and HP-DR (Figure 4, below, top panels), with HP-DR decreasing global protein turnover rates less than LP-DR. This agrees well with previous studies in which dietary and calorie restriction have been shown to decrease the overall rates of protein synthesis and degradation\textsuperscript{22}. Generally, protein quantity decreased for the \textasciitilde 1000 proteins measured in the LP-DR group. HP-DR showed almost no change in cellular protein quantity (Figure 4, below, bottom panels). According to our model (Figure 1), the combination of turnover rates and protein quantification enables us to describe these global differences as the result of changes in either protein synthesis or degradation. In this case, our data supports a fundamental difference in the global protein metabolism between the dietary protein conditions: With low protein, protein synthesis appears to be slowed globally, while high protein shows both decreased synthesis and degradation.
Figure 4: Protein turnover rate and quantitation comparisons between dietary restriction and ad-libitum. In the scatterplots above, each point represents the intersection of the turnover rates (upper two panels) or the quantitation (lower two panels) between dietary restriction (y-axis) and ad-libitum (x-axis) cohorts for a single protein. For each plot, the dotted line represents the identity line (y = x) with a slope of one, the dashed line indicates the linear regression fit line (equation and R-squared values displayed; all slope p-values << 0.00001 – see Table S1 in Appendices), and the error bars indicate 95% confidence intervals. Note that if the dashes and dots lie along the same line, this indicates no average difference between dietary restriction and ad-libitum cohort measurements.
Dietary restriction alters the global proteome post-transcriptionally.

We used RNA-Seq to measure the change in mRNA concentration between our dietary cohorts. The results show almost no change between ad-libitum and dietary restriction interventions, regardless of whether the mice were on low protein or high protein diets (Figure 5). Given that the turnover rates and concentrations of the proteins changed so markedly under dietary restriction, this lack of global perturbation in the transcriptome is remarkable. We concluded from this that the regulatory changes induced by dietary restriction occurred post-transcriptionally.

One possibility for the post-transcriptional regulation is that RNA sequence motifs, such as micro-RNA motifs and transcriptional activating factors that are known to regulate mRNA longevity and translation\(^23\), may be influencing downstream protein translation rates and levels. An analysis of the RNA-Seq data showed no correlation between known motifs and the levels of protein expression downstream in our study (data not shown). This indicates that the post-transcriptional regulation under dietary restriction is unlikely to result from RNA sequence-based regulation. Our global analysis of turnover, quantitation, and RNA-Seq data points toward regulatory changes happening at protein translation, post-translational processing, or degradation steps.
Figure 5: Comparisons between dietary restriction and ad-libitum mRNA amounts. In the scatterplots above, each point represents the intersection of the normalized log of the mRNA counts measured by RNA-Seq between dietary restriction (y-axis) and ad-libitum (x-axis) cohorts for a single gene. For each plot, the dotted line represents the identity line (y = x) with a slope of one, the dashed line indicates the linear regression fit line (equation and R-squared values displayed; confidence interval statistics listed in Table S1 in Appendices), and the error bars indicate 95% confidence intervals. Note that if the dashes and dots lie along the same line, this indicates no average difference between dietary restriction and ad-libitum cohort measurements.
Addition of dietary protein alters cellular division rates in dietary restriction.

Since protein concentration can be decreased through cellular division as well as degradation and other processes, we used a GC-MS method to measure DNA turnover rates, which is an accepted metric of cell division\textsuperscript{24}. DNA normally replicates only once per cellular division; thus, the turnover rate of DNA should be identical to the cellular division rate. Our data showed that the cellular division rates were largely identical between cohorts, with the exception of the high protein dietary restriction (HP-DR); the cellular division rates in this dietary cohort were significantly decreased compared to low protein dietary restriction (LP-DR) and high protein ad-libitum (HP-AL) (Figure 6). While at first glance this may seem to explain the difference in average protein quantitation concentration (Figure 4, bottom panels) between LP-DR and HP-DR, an examination of the scale of the turnover rates shows that the DNA turnover rate is one to two orders of magnitude smaller than the protein turnover rates, and thus cannot account for the difference alone. It seems more likely to us that the high protein levels interfere with normal dietary restriction metabolic reprogramming, decreasing cellular health, which would cause delays in cell proliferation.
Figure 6: Cellular proliferation rates differed between dietary treatment groups. Average DNA turnover rates are shown as fraction/day for DNA in LP-AL, LP-DR, HP-AL, and HP-DR cohorts. Error bars indicate 95% confidence intervals. Asterisks and horizontal brackets indicate significant differences between high protein dietary restriction (HP-DR) and high protein ad-libitum (HP-AL), as well as low protein dietary restriction (LP-DR).
Individual functional ontologies may be regulated separately from the global proteome change during dietary restriction.

The turnover and quantitation data indicated that the entire proteome was undergoing global shifts in metabolism due to dietary restriction and protein intake levels. This was not unexpected, and indeed was logical considering the broad phenotypic changes observed with dietary restriction; however, the mechanisms causing this global shift in regulation are likely to be masked by the overarching changes in the proteome. Nevertheless, we sought to investigate those protein functional groups and pathways that may be involved in causing this global shift. We reasoned that the regulation of these proteins may be perturbed differently from the overall global regulation; therefore, proteins that were only shifted globally should be very similar in relative turnover and quantitation changes, while the selectively regulated proteins should be significantly different. To separate these regulation levels, we first calculated turnover, quantitation, and gene count (RNA-Seq) values for the dietary restriction cohorts based on the ad-libitum controls and the regression line values in Figure 4 and Figure 5. These predicted values were then compared to the values measured experimentally, and the differences tested for statistical significance, as detailed in the Methods. We conservatively allowed for a +/- 20% noise level, and selected proteins that were significantly outside of this range for further analysis (the volcano plots in Figure 7 illustrate this selection process).
Figure 7: Experimental versus predicted values for turnover rates, quant, RNA-Seq counts, k-synth, and k-deg. Q-values for the comparisons between experimental and predicted values for the various measurements are
plotted on the y-axis, while the x-axis shows fractional change between the experimental and predicted values. The blue lines indicate a Q-value of 0.05 on the phred scale, and the red lines indicate the cutoff of a +/- 20% difference. Numbers of significant proteins are indicated; these numbers may or may not correlate with the displayed points on the graph, as some of those displayed points were eliminated by the q-value test, but still included in the graphic. Outliers past the ranges of the x-axis and y-axis are included in the numbers, but were not displayed in order to keep the plot scales uniform. In each graph, the predicted values were calculated based on linear regression fits between dietary-restricted and ad-libitum cohort values. The top row shows experimental versus predicted rate data for low protein (LP) dietary restriction and high protein (HP) dietary restriction cohorts. Subsequent rows display the same analysis applied to quantitative MS data, RNA-Seq data, k-synth values, and k-deg values.
Among the differentially expressed proteins, three broad functional categories of proteins emerged: mitochondrial, endoplasmic reticulum, and secreted. The k-synth and k-deg values for these proteins are plotted in Figure 8, below. Key differences were revealed between the high and low protein cohorts: Endoplasmic reticulum proteins were not significantly shifted in LP-DR, but for HP-DR, their synthesis rate constants were increased and degradation rates decreased relative to the global regulation. This may reflect the need for more protein processing capacity under high protein conditions, and would agree with our previous hypothesis that mTOR signaling is elevating translation in response to the high protein diet. Synthesis rate constants for secreted proteins in both LP-DR and HP-DR were elevated, while degradation rate constants remained mostly unchanged relative to global regulation. We hypothesize that regardless of the changes in the cellular metabolism, the liver cells need to maintain their secretory functions\(^{25}\) to maintain overall tissue and organismal homeostasis.

Mitochondrial proteins displayed an increase in synthesis rate constants in LP-DR, while their degradation rate constants remained unchanged relative to global regulation. This indicated that the mitochondrial proteins were under selective synthesis control, and that maintaining higher levels relative to other cellular proteins in dietary restriction confers benefits for cellular health, presumably by keeping energy flux at healthy levels. In contrast, HP-DR caused a decrease in both synthesis and degradation rate constants of mitochondrial proteins. Decreased degradation of mitochondrial proteins through mitophagy, the primary degradation pathway for mitochondria\(^{26}\), would be predicted to manifest as decreased turnover and increased concentration, and we see this in our data. We reason that the corresponding reduction in synthesis may thus be necessary in order to maintain homeostasis. Such a slowdown in the metabolism of mitochondrial proteins is interesting, given the attention that oxidative damage
caused by ROS (reactive oxygen species) created in the mitochondria has received in previous studies of dietary restriction. ROS are thought to accelerate aging through the induction of oxidative damage; one dietary restriction hypothesis is that the health and longevity benefits result from the reduction of this oxidative stress. If mitochondria lag in turning over, it is likely that their proteins will become more damaged and less functional over time, leading to deficiencies in mitochondrial function, which may produce more ROS and subsequent damage to the cell. This may provide an explanation for the attenuation of dietary restriction benefits by high dietary protein intake seen in previous studies.
Figure 8: K-synth and k-deg comparisons of unique proteins from significant ontology groups. Combined dot and box plots are shown of the k-synth and k-deg constants for the three significant ontology groups determined through the online ontology analysis tool DAVID18,19: endoplasmic reticulum, mitochondria, and secreted proteins were all significantly shifted from the global proteome regulation of dietary restriction. High protein (HP) is indicated in dark grey, while low protein (LP) is indicated in light grey. Each dot represents the k-synth or k-deg value for a unique protein. The interquartile range is indicated by the reduced width towards the median; non-overlap of interquartile ranges is indicative of a significant difference. Here, the emphasis is on direction of change from expected values. The dotted red line indicates the predicted (global) values for high protein, and the dashed blue line indicates the predicted (global) level for low protein dietary cohorts. Only unique protein id’s matched between dietary cohorts in each category were kept for this analysis; protein id’s that were shared between ontology categories were eliminated from these plots.
Ribosomal subunits display differential degradation rates under dietary restriction.

Adjustments in energy flux through manipulation of mitochondrial metabolism may provide part of the answer for dietary restriction benefits, but this does not speak directly to the cause of the globally decreased protein quantity and turnover observed in our data. To this end, we began to examine other possible nodes of regulation; the ribosome immediately attracted our attention in its role as a central control locus for protein synthesis. In previous work, we reported that certain ribosomal proteins had faster turnover rates than the ribosome as a whole, and hypothesized that some proteins were being replaced piecemeal as part of a ribosomal maintenance mechanism. The mice in that study were fed high-protein ad-libitum or high-protein dietary restriction diets. Since a low-protein dietary restriction diet is reportedly healthier than a high protein diet, and increased ribosomal maintenance should contribute to improved proteome quality, we reasoned that we should see magnified differences in the turnover rates and quantities of these “fast turnover” ribosomal proteins compared to the high protein diet.

An examination of the ribosomal proteins with the fastest turnover rates showed little overlap between low and high protein dietary cohorts. Moreover, the ribosomal protein turnover rates appeared to be changing less in LP-DR compared to HP-DR. While this did not directly refute the partial replacement hypothesis we previously proposed, the data appeared to imply that something else was occurring. A cursory examination of the ribosomal proteins turning over fastest in LP-DR and HP-DR, while not overlapping, seemed to show a bias towards large subunit ribosomal proteins. We applied a Mann-Whitney rank-order test (see Methods) to test whether the large (60S) and small (40S) ribosomal subunits were turning over at different rates. The results (displayed in Figure 9) confirmed that the large subunit proteins were turning over significantly faster than the small subunit proteins in both LP-DR and HP-DR, as well as HP-AL.
diets. When we applied the same statistical test to the quantitative MS data, the large subunit proteins were seen to be lower in concentration than the small subunit proteins in LP-AL as well as LP-DR, but in high protein diets, only HP-AL had a significant reduction in large subunit proteins compared to small subunit proteins. This implies that under low protein dietary conditions, it is normal for non-stoichiometric ratios of ribosomal subunits to be maintained in the cell, but under dietary restriction this ratio is skewed more heavily in favor of the small subunit as the large subunit undergoes increased degradation. Under high protein conditions, this dietary restriction reprogramming appears to be mitigated by a competing signal that brings the degradation of the large and small subunits more in line with each other.
Figure 9: Ribosomal subunit analysis. Turnover rates (panel A) and quantitation (panel B) for the ribosomal proteins observed in each of the dietary cohorts (data listed in Table 1 below). For each of the dietary cohorts (indicated on the x-axis), the ribosomal proteins were separated by whether they were part of the large subunit (60S (L)) or the small subunit (40S (S)) of the ribosome, and then plotted separately: large subunit protein data is shown in light grey, and small subunit protein data is shown in dark grey. The interquartile range is indicated by the reduced width towards the median, and significant differences (determined by one-tailed t-test) are indicated by asterisks and horizontal lines.
Table 1: Ribosomal subunit protein concentration and turnover rate data. Data for the ribosomal subunit analysis in Figure 9 is shown below. Quant MS (ratio relative to internal standard) and turnover rate (fraction per day) data are listed for 73 ribosomal proteins. All data except missing values were plotted ("NA" indicates missing data). "+/- 95% C.I." indicates the computed 95% confidence interval value for the associated value.
Ribosomal RNA quantities do not correlate with ribosomal proteins

The differences observed in ribosomal protein metabolism could be due to changes in degradation of the ribosome as a unit or the individual proteins. Therefore, in order to test whether the large and small ribosomal subunits were being differentially degraded in response to dietary restriction, we performed RT-qPCR on two ribosomal RNA’s (rRNA’s): the 16S rRNA (small subunit), and the 28S rRNA (large subunit). The results are shown in Figure 10 below. Based on the ribosomal protein quantitation data (Figure 9), we expected that the high protein cohorts should show comparable values for the large and small subunit rRNA’s (although the differences were statistically significant, the large range of the data suggests that perhaps they are not very different), and the qPCR bears this out. However, we were very surprised to see that the low protein cohorts showed the opposite trend to the protein quantitation: the large subunit rRNA was in greater quantity than the small subunit rRNA, ad libitum relative to dietary restriction. This result directly contradicts the hypothesis that differential degradation of ribosomal subunits may be responsible for slowing translation in response to dietary restriction. However, this data is only preliminary, and the RT-qPCR experiment needs to be repeated for verification. Possible implications of these RT-qPCR data are presented in the Discussion.
Figure 10: Ribosomal rRNA quantitation by RT-qPCR. The 16S and 28S rRNA molecules in each dietary cohort were quantified and compared by RT-qPCR. The left panel shows the normalized expression levels in arbitrary units. The right panel shows the rRNA ratio of dietary restriction (DR) to ad libitum (AL) in both high protein (HP) and low protein (LP) diets: 16S rRNA DR/AL and 28S rRNA DR/AL. Error bars in both panels indicate 95% confidence intervals. In the ratio comparisons, only the low protein 28S rRNA values were significantly different between AL and DR.
Discussion

Dietary restriction has long been known to extend lifespan and healthspan for a variety of model organisms. The mechanism behind dietary restriction remains debated, but one observation that has consistently emerged from dietary restriction studies is the slowdown of protein synthesis compared to an *ad-libitum* diet\(^{22a,29,30}\). Logically, this provides the cell with several benefits: Decreased protein synthesis lowers the burden for raw materials and energy in the cell, an advantage when adjusting to the lower nutrient conditions of DR. Multiple studies have reported that lower translation rates enable higher quality proteins to be made through improved translational fidelity\(^{31}\), increased co-translational folding time, improved chaperone to nascent protein ratios, and lowered burdens on the protein synthesis quality-control machinery such as the ribosomal quality-control complex\(^{32}\). We have previously proposed that ribosomes undergo maintenance to swap out damaged or malfunctioning proteins\(^{28}\); decreasing the protein synthesis burden would enable more ribosomes to be available for maintenance, increasing the quality of the available ribosome pool, which would presumably increase the quality of proteins produced during translation. As an additional consequence, the burden on degradation machinery would also decrease, as overall protein quality improved, enabling selective targeting of dysfunctional proteins for degradation that would otherwise form aggregates. The formation of protein aggregates within cells is a characteristic of many age-related diseases\(^{33}\). Thus, decreasing the protein synthesis burden on the cell is an attractive candidate for a global mechanism behind the health benefits of dietary restriction.

There are several possible mechanisms whereby global protein synthesis could be slowed. Previous studies have determined that dietary restriction triggers nutrient sensors in the cell that alter protein synthesis\(^{34}\). Gcn2 is a nutritional sensor that activates in response to an
increase in uncharged tRNAs in the cytosol, a change that occurs during nutritional stress\textsuperscript{34}. In response, Gcn2 rapidly inhibits the formation of the ternary translation initiation complex\textsuperscript{34}. mTOR is another well-characterized nutrient sensor that increases protein translation in response to amino acid signaling; under dietary restriction, mTOR is less active, and translation is slowed\textsuperscript{35}. Other pathways, such as the insulin/IGF-1 and MAPK pathways, also have direct effects on translational levels in response to nutritional signals\textsuperscript{36}.

Dietary restriction can also be viewed as the application of a cellular stress, to which the cell may respond by formation of stress granules\textsuperscript{37}. Stress granules are non-membrane bound aggregations of mRNA and associated proteins that represent incomplete initiation complexes\textsuperscript{38}. Though the literature on stress granule formation in dietary restriction is sparse, there is evidence of stress granule formation under similar conditions of more specific nutrient depletion\textsuperscript{39,40}. This formation of stress granules may represent an additional mechanism for altering the rate of protein synthesis in cells under dietary restriction. Interestingly, a recent proteomics study of stress granules reported that while small ribosomal subunit proteins can be found within stress granules, large ribosomal subunit proteins are conspicuously absent\textsuperscript{41}. 
Figure 11: A model for decreasing global translation in dietary restriction. Dietary restriction (DR) increases the formation of stress granules, which segregate mRNA associated with translation initiation factors and small ribosomal subunits (40S) away from the general cytosol, inhibiting protein translation and degradation of the small ribosomal subunit proteins.
Given these findings, plus the data in the current study, we propose here a model in which the formation of stress granules during dietary restriction represents an additional mechanism for slowing protein translation through the sequestration of mRNA and associated proteins, including the small (40S) ribosomal subunit (Figure 10). In this model, the stress granules inhibit translation directly by preventing the association of the large ribosomal subunit with the translation initiation complex. Protein synthesis is also indirectly hindered since the sequestration of the small (40S) ribosomal subunit prevents its degradation through cytosolic ribophagy, causing the ratio of 60S to 40S ribosomal subunits to fall below a more efficient near-stoichiometric match.

Our data contain several lines of support for such a phenomenon. The kinetic proteomics (Figure 4), combined with the quantitative proteomics results (Figure 4), indicated that dietary restriction was affecting primarily protein synthesis; however, the lack of significant changes in our transcriptomic data (Figure 5) pointed to a post-transcriptional regulatory mechanism. Mitochondria, endoplasmic reticulum, and secreted protein groups escaped this global synthetic slowdown (Figure 8); there is some literature support that sequestration into stress granules can be sequence-specific\textsuperscript{42,43}, although this is still debated.

The strongest support for this model comes from the turnover and quantitative MS data for the ribosomal proteins (Figure 9). When taken together, the increase in turnover rate and drop in quantity for the large (60S) ribosomal subunit relative to the small (40S) ribosomal subunit points to preferential degradation of the 60S subunit versus the 40S subunit. Although we believe that the simplest method is through sequestration of the 40S ribosomal subunit into stress granules where it is protected from cytosolic degradation relative to the 60S subunit, one
can envision a number of alternative ways the cell could increase the degradation of the large ribosomal subunit relative to the small subunit. For example, ribophagy could be targeted towards the large subunit through post-translational modification of the large subunit ribosomal proteins, or the large subunit as a whole. There could also be a more passive bias in ribophagy (possibly based on steric hindrance) towards degradation of ribosomal subunits that are unassociated with mRNA: given that the small ribosomal subunits are a vital component of the ternary complex that initiates translation, and the large subunit is the last component to be added before translation begins, it is logical that the small subunit would spend more time associated with mRNA. This provides one possible explanation for the difference in degradation rates that we see in the high protein *ad libitum* (HP-AL) cohort: increased mTOR signaling in response to higher dietary protein would lead to increased translation, thus increasing the number of small subunits associated with mRNA and decreasing the pool of free small ribosomal subunits.

The increased level of dietary protein itself may cause translational stress to the cell, as mTOR increases the level of protein translation past a healthy level under dietary restriction. mTOR is a major node of protein synthesis regulation that responds to changes in amino acid availability, especially leucine\(^4\); increased amino acid concentrations in the cell have been shown to upregulate mTOR signaling, stimulating the translation of proteins\(^3\). We hypothesize that the additional dietary protein in dietary restriction results in activation of mTOR, increasing protein synthesis despite the dietary restriction conditions, resulting in competing nutritional signals that result in the protein concentration within the cell remaining more or less constant (Figure 4). Evidence of increased mTOR activity can also be seen in the selective decrease in degradation of mitochondria under high protein dietary restriction (Figure 8); mTOR is also known to be a control node for the regulation of cellular autophagy\(^2\). The increased protein
translation burden caused by increased mTOR activity can logically be expected to cause a strain on cellular resources under lowered nutrient conditions, and may lead to decreased cellular fitness and attenuation of dietary restriction health benefits in the long term, as seen in other studies⁴. This extra level of stress would thus be expected to increase the number of stress granules, magnifying the sequestration of 40S ribosomal subunits.

To test this hypothesis, we performed an RT-qPCR experiment to measure the expression levels of 16S and 28S rRNA. These are constitutive components of the small and large ribosomal subunits, respectively, and the quantities of each should correlate tightly with the amounts of the small and large subunits themselves. If the sequestration model is correct, we expect that the concentration of 28S rRNA relative to 16S rRNA will decrease under dietary restriction, since it is only the small ribosomal subunit that is sequestered in stress granules. Our qPCR results showed that while the high protein dietary cohorts had rRNA levels that approximately correlated with the ribosomal protein concentrations (i.e. there was little difference between AL and DR diets), the low protein cohorts instead displayed increased 28S rRNA relative to 16S rRNA during dietary restriction. This is a preliminary experiment, and needs to be repeated for verification; for example, simple sample confusion may have resulted in erroneous data. Other possibilities include biases in the reverse transcription of 28S rRNA versus 16S rRNA (ribosomal RNA is known to be heavily modified⁴⁵ and this may be problematic), or primer bias in the qPCR could have resulted in skewed amplification of the rRNA. Temporal biases in ribosomal levels in the cell due to diurnal cycling⁴⁶ could also be a factor. However, if the data are true, they would either directly disprove the proposed model, or indicate other possible scenarios, such as a disconnect between ribosomal protein and ribosomal RNA metabolisms in response to dietary signaling. Ribosomal biogenesis is tightly controlled.
and coordinated in the cell\textsuperscript{47}, so this possibility would necessitate a rigorous set of research experiments to produce supporting evidence.

Given the large energy and material investment the cell makes in creating the ribosomal subunits, it seems counter-intuitive that a reduction in the 60S:40S subunit ratio by increased degradation of the large subunit proteins would lead to a healthier, more efficient proteostatic equilibrium. However, this may enable the overall quality of the large ribosomal subunit pool to increase, which may increase the proofreading accuracy of the ribosome. Future experiments specifically targeting the ribosomal proteins to assess post-translational modifications and damage due to oxidation would be useful. Other future experiments determining individual rRNA turnover rates would be used to support the 40S and 60S subunit protein turnover rates; the ribosomal proteins cover a spectrum of turnover rates, and the subunit turnover rates as a whole are inferred. Calculation of mRNA turnover rates would shed light on whether mRNA is being sequestered and presumably protected from degradation under dietary restriction (although reported exchanges of mRNA in stress granules with p bodies and the cytosol\textsuperscript{48} may complicate the picture). (We did have data on the turnover rates for total RNA, which were suggestive of higher turnover for all species of RNA under dietary restriction, but without a reliable way to distinguish the turnover rates of the various species of RNA from each other, more definitive conclusions were not possible.) More comprehensive proteomics studies of stress granules including kinetic measurements would be illuminating, as well as a measurement of the number and size of stress granules in dietary restriction through staining and microscopy.

Taken together, these future directions would either confirm or disprove the slowdown of protein synthesis through sequestration of ribosomal subunits into stress granules. Interestingly, the model itself presents a “which came first: the chicken or the egg?” kind of question, one
which is not answerable at present: Does dietary restriction trigger stress granule formation to alter protein expression in order to confer survival benefits, with the ribosomal subunit turnover differential as a side effect? Or is the goal to reduce long-term translation rates by altering the 60S:40S ratio through degradation, with sequestration the mechanism for achieving this goal? Although it may seem rather academic, as researchers pursue dietary restriction mimetics for increasing human longevity and health, this may become important in deciding which cellular proteins to target for intervention.

In summary, we present a methodology that needs only four measurements to monitor proteostasis regulation at the systems scale. We applied this approach to short-term dietary restriction in mice under low and high dietary protein conditions. The data from our experiments allowed us to recapitulate many of the phenomena that have previously been observed in dietary restriction, albeit in combination here rather than in isolation. Moreover, we propose that under dietary restriction, the cell decreases global protein synthesis through the sequestration of the small ribosomal subunits into stress granules. This global slowdown in translational burden lowers energy and raw material demands on the cell, and improves protein folding and quality control\textsuperscript{49}. Ultimately, these improvements in proteome metabolism and maintenance lead to improvements in cellular health and longevity observed under dietary restriction.
References


CHAPTER 3

“Exploratory Investigation of AGE-Mediated Attenuation of Calorie Restriction Benefits in Mice Using a Combined Omics Method”

This chapter represents the focus of my original research aim, which was to investigate the effect of AGEs (advanced glycation end-products) on calorie restriction. Previous studies had shown that these protein modifications attenuated and even canceled the life-extending benefits of calorie restriction; however, the mechanism was unknown. This presented a novel opportunity for application of our kinetic proteomic method, along with other omics methodology, to produce a comprehensive mapping of the AGE effect on cellular metabolism. Unfortunately, the data were mostly negative: we did not see a profound change in calorie restriction caused by dietary AGEs. Because of logistical reasons, this study could not be easily repeated and refined, so the analysis presented in this chapter represents more of a work in progress, wherein the data is best viewed as preliminary for a more rigorous and ambitious future experiment. The complete data analysis was useful and necessary, however, to confirm the lack of a detectable profound AGE effect under these experimental conditions, hence its inclusion in this dissertation. Combined with the results presented in Chapter 2, this study provides the basis for experimental modification, as well as baseline data, for continuing the investigation into the negative consequences of AGE consumption during calorie restriction.

Unlike Chapter 2 and Chapter 4, in which I benefitted greatly from close collaboration with coauthors, the experiments and data analysis were primarily executed by myself. By itself, the results of this study do not merit publication as a scientific paper. It is my hope that this work will be continued by a future student who will be able to use the data to gain fundamental
insight into how to mitigate the potential damage that AGE-heavy Western diets are causing to an aging population.
Title

“Exploratory Investigation of AGE-Mediated Attenuation of Calorie Restriction Benefits in Mice Using a Combined Omics Method”

Abstract

Calorie restriction (CR) is a time-tested method for improving the health and lifespan of laboratory animals. Studies show, however, that the addition of protein modifications termed “advanced glycation end-products” (AGEs) to the diet of calorie-restricted animals can lead to the attenuation of these health benefits and a shortening of lifespan. As these compounds are widely prevalent in modern Western diets, this has significant implications for human health. We hypothesized that these AGEs constituted an abnormal dietary signal that disrupted protein homeostasis and decreased the quality of the cellular proteome, and used a multi-omic approach comprised of kinetic proteomics, quantitative proteomics, and transcriptomics to investigate the liver proteomes of C57/BL6 male mice on calorie restriction with additional AGEs. Our results showed that the measurable proteomes of mice on a calorie-restricted diet supplemented with a small amount of AGE are not significantly different from isocaloric control mice with either a small amount of additional dietary protein, or no additional protein at all. This suggests either that the effects of AGEs in the diet were too subtle to detect, and were overwhelmed by a dominant CR phenotype, or the experimental doses were too small to cause a significant difference between the dietary cohorts in the liver tissue. Additional experimentation will be necessary to distinguish between these possibilities.
**Introduction**

Human lifespans have increased markedly in the past century, accompanied by dramatic increases in age-related diseases such as cancer\(^1\). Cardiovascular disease, kidney disease, type II diabetes, and neurodegenerative diseases such as Alzheimer’s are increasingly common\(^2\), and place a heavy burden upon aging populations; developing more effective treatments and therapies for these diseases is therefore imperative.

Calorie restriction has attracted a great deal of attention for its ability to improve both the health and lifespan of model organisms. A calorie-restricted diet involves a significant reduction in total calories (up to 60%) without malnutrition, and although not currently considered practical for widespread human application, modifications such as short-term calorie restriction and intermittent fasting hold promise\(^3\). From a research standpoint, calorie restriction provides a model for healthy aging that can be studied not only to learn more about the etiology of aging, but also to discover targets for pharmacological interventions that would not only forestall deleterious aging processes, but prevent age-related diseases by treating the root cause: age-related physiological decline.

The importance of diet in the aging process can also be seen in the discovery that dietary AGEs (advanced glycation end-products) can attenuate and even cancel the benefits of a calorie-restricted diet in laboratory animals\(^4\). Advanced glycation end-products (AGEs) result from the irreversible non-enzymatic reaction of reductive sugars with peptides, proteins, and lipids, and can be formed both *in vivo* and *ex vivo*\(^5\). These AGEs can be found in high amounts in the typical Western diet\(^6\), which is abundant in sugars, fats, and animal protein. Perhaps not coincidentally, first world countries are currently suffering from epidemic levels of insulin
resistance, diabetes, and obesity\textsuperscript{7}, as well as a rise in other traditionally age-related diseases among younger people\textsuperscript{8}. Could AGEs be accelerating deleterious aging processes, thereby increasing these disease rates?

The ingestion of AGEs has been shown to increase systemic oxidative stress and inflammation levels, which have been tied to increased susceptibility to insulin resistance, diabetes, cardiovascular disease, and kidney disease\textsuperscript{4}. Inflammation also plays a large role in cancer: chronic inflammation can drive tumorigenesis, suppress anti-tumor immune responses, stimulate angiogenesis, and advance metastasis\textsuperscript{9}. Although the mechanism for AGE-mediated cancellation of calorie restriction benefits is incompletely understood, it is likely to be related to increased inflammation, a hallmark of aging\textsuperscript{10}.

The purpose of the current study was to explore these effects on the proteome of calorie-restricted mice. Our hypothesis is that calorie restriction functions by improving protein homeostasis (proteostasis), with ripple effects extending to tissue, organ, and overall health. AGEs, by extension, are predicted to disrupt the beneficial reprogramming of proteome regulation, mitigating the effects of calorie restriction. The specifics of this disruption, and the magnitude of the changes in the proteome, were the focus of this study.
Methods

Animal Handling:

*Mouse housing, genotype, and diet*

All animal procedure protocols were approved by the Institutional Animal Care and Use Committee (IACUC) of Brigham Young University. Housing for mice was provided in a pathogen-free facility for the duration of the experiment. 10-week old C57/BL6 male mice were purchased from Charles River Laboratory, and fed ad-libitum for one week after arrival to acclimate to the facility. Mice were then randomly divided into dietary cohort groups of 9 mice each: ad-libitum (AL), calorie-restricted (CR), calorie-restricted with additional protein (CR PROT), and calorie-restricted with additional AGE protein (CR AGE); the mice were housed individually to ensure equal access to food. All animals were fed NIH31 chow, with the CR animals receiving a variant fortified with additional nutrients. AL animals had constant access to food, and CR animals were fed daily a pellet of 3g ± .1g in size (65% of expected AL consumption). Additional protein and AGEs were added to the E and E2 diets by pipetting a small amount (4.2 mg in PBS buffer) of either BSA (bovine serum albumin) or AGE-modified BSA directly onto the pellet rations and letting them dry before feeding (the additional calories were judged to be minimal, and the diets judged to be isocaloric to the normal CR diet). Mice were weighed weekly during one of the daily feedings of CR animals to monitor health and weight loss.

*Movement Data Collection*

The activity of the mice was measured using an in-house constructed motion sensor apparatus. This apparatus consisted of four low-power infrared light emitters and four infrared
sensors connected to a computer equipped with the LabVIEW program\textsuperscript{11}; with the emitters lined up on the outside of a mouse cage on one side, and the sensors positioned opposite of the emitters on the outside of the other side of the mouse cage, interruptions in the infrared beams could be detected whenever the mouse inside the cage crossed the beams. The number of beam crosses was recorded over a twenty-four period for each mouse measured in order to minimize variations due to nocturnal/diurnal activity level differences. The total number of beam crosses was divided by the total time of the measurement in order to eliminate differences between measurements due to minor time differences, as well as to provide an average movement rate. Each mouse was measured at least twice, and any irregularities, such as apparatus or computer malfunction, that occurred during a measurement, no matter how minor, eliminated that measurement from the downstream analysis.

\textit{Metabolic Labeling}

At 10 weeks, (except for the 0-day time points, one per dietary cohort) the mice were given an intraperitoneal bolus injection of sterile D\textsubscript{2}O saline at 35 µL/g body weight. This injection brought the mice up to 5\% Molar Percent Excess (MPE) deuterium as previously described\textsuperscript{12}. Mice were then provided drinking water containing 8\% MPE to maintain the 5\% MPE in the animals’ body water.

\textit{Euthanasia and Sample Collection}

Two animals from each diet were sacrificed at 1 day, 3 days, 9 days and 27 days after bolus injection, with one animal from each group sacrificed without receiving a bolus injection or any other D\textsubscript{2}O labeling. Mice were anesthetized with CO\textsubscript{2} and then euthanized by cardiac puncture. Mice were then immediately dissected and all tissues except for blood were flash
frozen on solid CO₂ and then stored at -80°C. Blood was stored on ice until it could be centrifuged at 800 x g for 10 minutes at 4°C. The centrifugation separated serum and red blood cells, which were stored in separate containers at -80°C.

**Mitochondrial Respiration**

Fresh liver tissue was quickly removed from exsanguinated mice and immediately placed in ice-cold mitochondrial respiration buffer (MiR05: 0.5 mM EGTA, 10 mM KH₂PO₄, 3 mM MgCl₂-6 H₂O, 60 mM K-lactobionate, 20 mM HEPES, 110 mM Sucrose, 1 mg/ml fatty acid free BSA, pH 7.1) and trimmed of connective tissue. Tissue was gently separated and homogenized under a surgical scope (Olympus, ST) to particles of approximately 1 mg. Homogenate was then transferred to a tube with chilled MiR05 and 50 μg/ml saponin and rocked at 4 °C for 30 min, then washed in MiR05 at 4 °C for at least 15 min before use. High-resolution O₂ consumption was determined at 37 °C using the Oroboros O₂K Oxygraph. Before addition of sample into respiration chambers, a baseline oxygen consumption rate was determined. After addition of sample, the chambers were hyperoxygenated to ~350 nmol/ml, and respiration was determined according to the manufacturer’s protocol. Lastly, residual oxygen consumption was measured by adding antimycin A (2.5 μM) to block complex III action, effectively stopping any electron flow and providing a baseline respiration rate.

**Kinetic Proteomic Methods**

**Measurement of MPE**

Aliquots of serum were distilled in 2.0 mL screw cap tubes overnight in a 90°C sand bath, and the distillates collected. The distillate was diluted 1:300 in ddH₂O, and MPE of deuterium was directly measured against a D₂O standard curve using a cavity ring-down water
isotope analyzer (Los Gatos Research [LGR], Los Gatos, CA, USA) according to the published method.13.

**Mass Spectrometry Sample Preparation**

50-100 mg of mouse liver tissue was placed in 1.0 mL ammonium bicarbonate (ABC) (25 mM, pH 8.5) along with protease inhibitor cocktail (Sigma, 1% v/v final concentration) and homogenized using a MP Biomedicals FastPrep-24 bead homogenizer (2 x 3mm steel beads, 6 m/s for 60 seconds). Protein concentration of the homogenate was measured using a bicinchoninic acid (BCA) protein assay (Thermo Fisher). For each sample, 50-100 µg of protein was placed on a 30 kDa centrifugal filters (VWR), and 100 µL of a concentrated guanidine solution (6 M guanidine HCl, 100 mM Tris-HCl pH 8.5) added to the sample. The filter was then centrifuged at 14,000 x g for 15 minutes. This guanidine wash was repeated, and the flow-through discarded. Disulfide bonds were reduced using a 10 mM dithiothreitol/6 M guanidine HCl/100 mM Tris-HCl (pH 8.5) solution (100 µL total volume) added directly to the filters, with an incubation of 1 hour at 60°C in a sand bath. Cysteine sulfhydryl groups were protected by reaction with iodoacetamide (IAM, 20 mM) for 60 minutes in the dark. Afterwards, the samples were centrifuged at 14,000 x g for 15 minutes, and the flow-through discarded. The samples were then washed twice with ABC (200 µL, centrifuged 15 minutes at 14,000 x g), and the flow-through discarded. Pierce MS-Grade trypsin was added (1:50 w:w) in 300 µL ABC to each sample, followed by incubation at 37°C overnight. Resulting peptides were eluted by centrifugation at 14,000 x g for 30 minutes, followed by a wash with 100 µL of ABC and an additional centrifugation for 30 minutes at 14,000 x g. Filters were discarded and the filtrate was dried using a speedvac (Sorval); the dried samples were stored at 4°C until use. With the exception of diet A samples, which were HPLC fractionated (see Methods in Chapter 2), these
samples were then resuspended in Orbitrap A buffer (0.1% formic acid (Pierce LC-MS grade) in H2O (Optima grade Thermo Fisher)) to 1 µg/µL concentration in mass spectrometry vials for measurement.

**LC-MS Data Acquisition**

MS data were collected on a Thermo Lumos Tribrid (Orbitrap) mass spectrometer. Tryptic peptides were separated using a reverse phase C18 column (Acclaim PepMap™ 100) and a Thermo Easy-Spray source. Mobile phase for the liquid chromatography was 0.1% formic acid in H2O (Buffer A) and 0.1% formic acid in 80% acetonitrile (Optima grade Thermo Fisher) with 20% H2O (Buffer B) on an Easy-nLC 1200 HPLC system. Samples were eluted using a gradient of 5% B to 22% B over 85 minutes, 22% to 32% B over 15 minutes, and a wash of 32% to 95% B over 10 minutes, which was held at 95% B for 10 minutes. Sample loading and equilibration were performed using the HPLC’s built in methods. MS only runs were performed using 2400 V in the ion source, 60000 resolution with a scan range of 375-1700 m/z, 30% RF Lens, quadrupole isolation, 8 *10^5 AGC target, and a maximum injection time of 50 ms. MS/MS scans were performed using the same settings as MS only scans, with 3 seconds allowed per MS/MS after each MS scan, using the following filters: peptide monoisotopic peak determination, intensity threshold of 5*10^3, fragmentation of charge states +2 to +6, dynamic exclusion that excluded a peak after being chosen once within 60 seconds, an error tolerance of 10 ppm high and low, and isotopes excluded. The fragmentation scan used an isolation window of 1.6 m/z, CID fragmentation with an energy of 30%, detection in the linear ion trap in Rapid Scan mode with an AGC target of 1*10^4, a maximum injection time of 35 milliseconds, and used the “Inject Ions for All Available Parallelizable Time” option.
Protein and Peptide identification

PEAKS Studio software (version 8.5) was used for de novo sequencing database searching to identify proteins in the raw MS data. Peptides were identified from MS/MS spectra by searching against the Swiss-Prot mouse database (downloaded January 2018) with a reverse sequence decoy database concatenated. Variables for the search were as follows: enzyme was set as trypsin with one missed cleavage site. Carbamidomethylation of cysteine was set as a fixed modification, while N-terminal acetylation and methionine oxidation were set as variable modifications. A false positive rate of 0.01 was used to filter peptide and protein identification results. Minimum length of peptide was set to 7 amino acids, and at least 2 peptides were required for protein identification. A mass error of 20 ppm was set for the precursor mass, and the mass error was set as 0.3 Da for MSMS.

Calculation of Turnover Rates

Identification files from PEAKS and MS-only mass spectrometry data were analyzed with the DeuteRater software package, which provided the protein turnover rates used for later analyses. The resulting kinetic data was filtered to remove data with extreme outliers or other issues: First, rates that were greater than 1 or less than 0.03 were eliminated from further analysis, as these rates represented extrapolations outside of the range of rates that could be calculated confidently with the time points used in the experiment. Since the kinetic proteomics data come from curve fits of relevant measurements, all curves with a Pearson’s R² less than 0.5, or with a covariance (standard deviation/rate value) of greater than 0.2 were also removed from further analysis (the standard deviation was divided by the turnover rate for normalization).
Quantitative Proteomic Methods

Quantitative Mass Spectrometry Sample Preparation

A mouse liver labeled according to a Stable Isotope Labeling by Amino Acids in Mammals (SILAM) protocol (Lys6\(^{13}\)C, Cambridge Isotopes) was purchased and homogenized as described in the “Mass Spectrometry Sample Preparation” section above. For each sample, 50 µg of protein from the SILAM homogenate was combined with 50 µg protein from the sample homogenate (with the exception of the diet A samples; since these underwent HPLC fractionation, 150 µg was used instead of 50 µg). Three biological replicates from each dietary group were prepared. Since isotopic shifts caused by D\(_2\)O labeling creates difficulties in peptide identification, only 0-day and 1-day time point animals (measured from the date of intraperitoneal injection of D\(_2\)O) were used for this experiment. The mixtures of SILAM and sample homogenates were otherwise prepared identically to those described within the “Kinetic Proteomic Methods” section.

Mass Spectrometry Analysis

Samples were re-suspended to 1 µg/µL concentration in 0.1% formic acid. Samples were analyzed on a Thermo Lumos Tribrid instrument paired with a Thermo Easy-nLC 1200 HPLC. Solvent A was 0.1% formic acid, 100% water; Solvent B was 0.1% formic acid, 80% acetonitrile, 20% water (all solvents were LC-MS grade). Gradient was as follows: 0-45 minutes was 5-45% Buffer B, 45-50 minutes 45-100% Buffer B, 50-60 minutes 100% Buffer B. The MS scans were collected at 60,000 Resolution, with 20 MSMS scans between MS scans. MSMS fragmentation was done with a 1.6 m/z isolation window, and CID fragmentation with 28%
fragmentation energy. Fragment ions were measured in the Linear Ion Trap with the “Normal” resolution setting.

Data Analysis

SILAM Data was analyzed using the MaxQuant software package\textsuperscript{16} to quantify peptides and proteins. All data was analyzed at the same time, with each dietary group analyzed as a single experimental group. Constant modifications were carbamidomethylation for cysteine; variable modifications allowed were oxidation of methionine and acetylation of the N-terminus. The identification database was the same as that used for the kinetic analysis. Lys6 was added as a heavy label; all other settings left at default. Unlabeled divided by labeled signal intensities were determined and used to determine amounts of peptide in the sample relative to the SILAM standard. The relative peptide abundances were then compared between samples.

DNA Turnover Rate Analysis

Since DNA turnover rate serves as a useful proxy for cellular division rate, we measured the DNA turnover rates for all dietary cohorts by GC-MS using a previously published method\textsuperscript{12}. Briefly, DNA was extracted from 30-90 mg of frozen liver samples using the DNeasy kit (Qiagen, Valencia, CA), and hydrolyzed to deoxyribonucleosides overnight at 37°C using S1 nuclease and potato acid phosphatase. The hydrolysates were reacted with pentafluorobenzyl hydroxylamine and acetic acid, and then acetylated with acetic anhydride and 1-methylimidazole. Dichloromethane extracts were dried, resuspended in ethyl acetate, and analyzed via GC/MS on an Agilent DB-17 column with negative chemical ionization, using helium as a carrier gas and CH\textsubscript{4} as the reagent gas. The fractional molar isotope abundances at m/z 435 (M0 mass isotopomer) and 436 (M1) of the pentafluorobenzyl triacetyl derivative of
purine dR were quantified using ChemStation software. Excess fractional enrichment of D$_2$O was calculated relative to an unlabeled pentafluorobenzyl triacetyl purine derivative, and fraction new DNA was determined for each sample by comparing excess fractional enrichment to the calculated theoretical maximum. GraphPad Prism was used to perform least-squares fits to first-order rate curves on plots of fraction new DNA versus time point. Standard deviation and confidence intervals were calculated based on results from two biological replicates (each measured three times) per time point. DNA turnover rates were analyzed and displayed using R scripts$^{17}$.

**RNA-Seq**

Following the manufacturer’s protocol, we used the Direct-zol RNA MiniPrep Plus kit from Zymo Research to extract RNA from the liver tissue of all animals in this study. Briefly, 10-20 mg of frozen liver tissue from each mouse was processed using a Trizol extraction, and the resulting RNA frozen at -80°C. Since we wished to identify changes in the transcriptome that resulted only from dietary differences, RNA from three different mice in the same dietary cohort sacrificed at different time points were pooled to form a single sample for RNA-Seq, thus masking any differences resulting from temporal variations. There were three pools per dietary cohort, containing RNA from three mice each (4ug of total RNA in 50 μL of RNase-free water). Each pool was submitted as a single sample to the BYU Sequencing Center. The BYU Sequencing Center constructed cDNA libraries after a poly-A pull-down to enrich for mRNA, and then sequenced the libraries using an Illumina HiSeq 2500 sequencing platform. RNA-Seq gene count data were analyzed in R$^{17}$ using the DESeq2 package$^{18}$ (see Appendices) in Bioconductor 3.3 from Bioconductor.org$^{18}$, downloaded on October 20, 2016.
Multi-omics Data Analysis

Global Versus Selective Regulation Determination

Linear regression analyses were performed on turnover rate (Figure 2), quantitative MS (Figure 3), and RNA-Seq data (Figure 5). The global regression constants were used to calculate expected values for each protein observed in the calorie-restricted (CR), calorie-restricted plus AGEs (CR AGE), and calorie-restricted plus protein (CR PROT) cohorts relative to the “control” diets: ad-libitum (AL) for CR, CR for CR PROT and CR AGE, and CR PROT for CR AGE. This was done to identify protein specific regulation versus global regulation of the entire proteome during calorie restriction, protein, and AGE dietary signaling. A paired t-test between predicted and experimental values was then run using R scripting17 (see Appendices) for each individual protein in CR vs AL, CR PROT vs CR, CR AGE vs CR, and CR AGE vs CR PROT, for turnover rate, quantitative MS, and RNA-Seq data sets, generating a list of p-values for each analysis. To reduce false positives, a q-value test was run in R17 using the qvalue package19 (see Appendices) for each analysis, with an FDR level set to 0.01. Any proteins designated as not significantly different were excluded from downstream analyses. A threshold of +/- 20% change was used to further screen proteins that were significantly different from their predicted values; these were divided into up- and down-regulated categories, respectively. Volcano plots generated using R (Version 1.1.463)17 were used to display the results of these tests (see Figure 6).

Gene Ontology Analysis

Ontology groupings were determined using DAVID20,21. Proteins/gene lists were obtained from the Global Versus Selective Regulation Determination analysis above. Ensembl
IDs from RNA-Seq data were translated to UniProt accession numbers prior to ontology analysis. Each category was input into DAVID\textsuperscript{20,21} against a background comprised of all protein ID’s in that particular analysis (i.e. combined from up-regulated, down-regulated, and not significantly different groups). Ontology groups with Benjamini-Hochberg coefficients less than or equal to 0.05 were retained for analysis; all others were disregarded.

\textit{Ribosomal Subunit Analysis}

Turnover rate and quantitative MS data for all of the ribosomal proteins was extracted from the larger data sets, and separated into two categories based on whether the ribosomal protein was part of the large subunit (60S) of the ribosome or the small subunit (40S) of the ribosome. A Mann-Whitney non-parametric rank order test\textsuperscript{22} was performed using Microsoft Excel (version 14.7.7) on turnover and quantitative MS data to detect whether the large subunit was turning over significantly faster or if its concentration was significantly lower than the small subunit. One-tailed p-values of 0.05 or smaller indicated significance.
Results

Calorie restriction resulted in a measurably distinct phenotype compared to ad-libitum controls.

Weight and movement for the mice were measured during the dietary treatments, and mitochondrial respiration was measured immediately after sacrifice. Together, these measurements provided some insight into the overall health of the mice and the outward effects of the dietary treatments. Calorie restriction has been reported many times before to reduce weight gain relative to ad libitum diets\textsuperscript{23}, and we saw that effect in our study as well. In Figure 1, it can be seen that all of the mice gained weight over the 10-week course of the dietary interventions; however, all of the calorie restriction cohorts gained weight more slowly than the ad-libitum mice (AL). After 2 weeks, mice on diets CR, CR PROT, and CR AGE had all become statistically significantly lighter than their AL counterparts; there were no statistically significant differences in mouse weight between diets CR, CR PROT, and CR AGE throughout the course of the dietary interventions. Since the dosage of additional protein and AGEs was kept very low (0.14% w/w, see Methods) in order for the calorie restriction diets to be iso-caloric, this was not particularly surprising.

Likewise, there was not much difference in the activity levels of the mice, measured over twenty-four hour periods using a custom beam-crossing motion sensor assay. In the upper right panel of Figure 1, only the AL and CR PROT cohorts were statistically different from each other, suggesting that the protein addition to CR may have had a stimulating effect on mouse activity; however, we have other data (not shown) from mice that were under dietary restriction rather than calorie restriction that is very similar to the CR PROT activity level. Given that dietary restriction and calorie restriction are often used interchangeably in the literature\textsuperscript{24}, this suggests
that, rather than being a real effect, stochastic effects in our measurements may be exaggerating measured differences at the reported n values. Further verification is required before we can confidently report that protein-augmented calorie restricted causes an increase in activity levels of mice.

The mitochondrial respiration data (Figure 1, bottom two panels) showed the most variation between the dietary cohorts. Pair-wise two-tailed t-testing showed that calorie restriction was significantly different from CR plus protein (CR PROT) under the GM testing conditions (mitochondrial leak) significantly different from all of the other diets under the D testing conditions (ADP), and nearly statistically significant (p = 0.059) from ad libitum (AL) under the S testing conditions (succinate). No other significant differences were found, suggesting that the addition of protein or AGEs to calorie restriction affects mitochondrial respiration, returning oxygen utilization by mitochondria to near ad-libitum levels (?).

Efficiency levels (Figure 1, bottom right) relative to GM suggest that CR decreases O2 flux relative to ad-libitum, evidencing decreasing efficiency, and that the addition of protein or AGEs to the CR diet mitigates this effect in the D (ADP) testing, but not in the S (succinate) testing, indicating that the protein/AGE effect may be complex. However, the lack of significance between the efficiency measurements makes this result suggestive only. More in-depth data on the mitochondria at the protein level was required to confirm this organelle as a regulatory focus node for calorie restriction.
Figure 1: Mouse phenotype data. Mouse weight, movement, and mitochondrial respiration data are shown for the four dietary cohorts. The upper left panel shows the weekly mouse weight data over the 10 weeks of dietary treatment. Mouse weights were normalized to starting weights before averaging. The ad libitum cohort (AL) is significantly different from the other three cohorts after week 2, but the three CR dietary cohorts were never statistically different from each other over the course of the 10 weeks. The upper right panel is a dot plot of the activity level of the mice; the only significant difference is between AL and CR PROT. The bottom two panels are bar graphs displaying the mitochondrial respiration data of the dietary cohorts. “GM” is the mitochondria at rest with oxygen, “D” is the addition of ADP and oxygen to show the function of the Electron Transport Chain (ETC) excluding complex 2, and “S” shows complex 2 added to the rest of the ETC. The lower right panel displays the relative efficiency of the ADP effect (D/GM) and the Succinate effect compared to mitochondrial leak (GM). All error bars represent 95% confidence intervals.
Calorie restriction results in decreased turnover rates. Protein- and AGE-supplemented calorie restriction diets also have decreased protein turnover rates.

Our examination of the proteomes of the various cohorts began with a global look at turnover rates. As shown in Figure 2, below, each of the three calorie restriction diets showed an overall decrease in protein turnover throughout the proteome relative to ad-libitum diet (AL). This is consistent with the results from the previous study using dietary restriction instead of calorie restriction, and also correlates with other studies, showing the relative interchangeability between dietary and calorie restriction. Also consistent with the previous dietary restriction analysis is the fact that the addition of protein (AGE-modified (CR AGE) as well as non-modified (CR PROT)) perturbed the turnover rates less than unmodified calorie restriction (CR). As with the previous dietary restriction study, this indicates a global slowdown of protein metabolism in the cell.

Interestingly, CR PROT and CR AGE proved very similar to each other as well as to CR, even though they are quite different when measured relative to AL. This is probably attributable to the fact that proteins were matched only between the two diets being compared rather than between all diets. Although this decreased the reliability between dietary comparisons, this was more acceptable than matching the proteins across all of the diets, as this resulted in an unacceptably small list (see Supplemental Data). Duplicating the analyses with the matched proteins list resulted in very similar regression values.
Figure 2: Turnover rate comparisons between dietary cohorts. In the top three panels, each of the calorie restriction cohorts (y-axis) is compared with ad-libitum diet (x-axis) for proteins matched between the two diets in each graph. In the bottom three panels, the three calorie restriction cohorts are compared to each other: CR PROT vs. CR, CR AGE vs. CR, and CR AGE vs. CR PROT. Red dashed lines indicate linear regression fits (all slope p-values << 0.00001 – see Table S2 in Appendices for regression line statistics), while dotted blue lines indicate identity lines with a slope of one. Error bars indicate 95% confidence intervals.
Calorie restriction may result in lower overall concentration of cellular proteins. Protein and AGE supplementation did not significantly alter the CR protein concentration.

A global comparison of the protein concentrations across the proteome of the various cohorts suffered from both very noisy data and small degrees of overlap between data sets, even though a SILAC-based method was chosen to mitigate both of these effects (a label-free quantitative MS data set had even poorer results (data not shown)). As shown in Figure 3 below, although the three calorie restriction diets (CR, CR AGE, CR PROT) show consistently lower overall protein quantity compared to ad-libitum diet (AL), poor correlation values between the data sets makes this an unreliable measurement. This may be due to the fact that the AL sample was fractionated by HPLC before measurement on the LUMOS Orbitrap, while the other three diets were not. While this increased the number of protein identifications from the AL samples significantly, such a difference in sample preparation may have introduced variables that made the measurements more unreliable and noisy.

However, since the calorie restriction diet samples were prepared and analyzed identically, the quantitative mass spectrometry measurement comparisons in the bottom row of Figure 3 can be considered reliable. Despite the inability to use the AL data, it is still apparent that there is little to no difference between the three calorie restriction diets. This suggests that adding small amounts of protein or AGEs to calorie restriction has almost no effect on the overall quantity of proteins in the proteome.
Figure 3: Protein quantitative mass spectrometry comparisons between dietary cohorts. In the top three panels, each of the calorie restriction cohorts (y-axis) is compared with ad-libitum diet (x-axis) for proteins matched between the two diets in each graph. In the bottom three panels, the three calorie restriction cohorts are compared to each other: CR PROT vs. CR, CR AGE vs. CR, and CR AGE vs. CR PROT. Red dashed lines indicate linear regression fits (all slope p-values < 0.05 – see Table S2 in Appendices for regression line statistics), while dotted blue lines indicate identity lines with a slope of one. Error bars indicate 95% confidence intervals.
Differences in turnover rate and quantity of proteins are not due to changes in cellular division.

One possible cause for changes in protein turnover and quantity in a cell is dilution through cellular division. To account for this, we measured the turnover rate of DNA within our samples. Since DNA should only replicate once per cell division, the turnover rate serves as a measurable proxy for the cellular division rate. One major advantage of our D$_2$O isotopic labeling method is that DNA becomes metabolically labeled similarly to proteins, and the turnover rate can thus also be measured using mass spectrometry (described in detail in Methods). The DNA turnover rates for ad-libitum (AL), calorie restriction (CR), calorie restriction plus AGEs (CR AGE), and calorie restriction plus protein (CR PROT) are shown in Figure 4 below. Although it is tempting to speculate that calorie restriction decreases cellular division rates, and this is attenuated slightly by the addition of protein or AGEs, none of the DNA turnover rates is significantly different from any other. Moreover, all of the DNA turnover rates are much slower than the protein turnover rates measured in our study: the slowest protein rates are an order of magnitude larger. We thus conclude that cellular division is not a major factor in perturbations to protein turnover rate and quantity due to our dietary interventions.
Figure 4: DNA turnover rate measurements. Average turnover rates are shown as fraction/day for DNA in diets AL, CR, CR AGE, and CR PROT. Error bars indicate standard errors.
Perturbations to protein regulation under calorie restriction are post-transcriptional.

In order to gain insight into gene expression regulation under calorie restriction conditions, we conducted next-gen RNA sequencing (RNA-Seq) analyses on each of the dietary cohorts. Relative expression of matched genes between dietary cohorts was then examined for any significant differences. All of the dietary cohorts were extremely similar to each other, indicating that the effects of calorie restriction, with or without additives, were due to changes in post-transcriptional regulation. This correlates very well with the results of the dietary restriction study in the previous chapter, and provides further support to a post-transcriptional hypothesis. Unfortunately, while the data are useful for measuring relative gene expression levels within and between data sets, normalization procedures in the data analysis procedure preclude making any global RNA quantitative comparison; left unanswered is the question of whether the overall level of mRNA in a cell goes up or down under calorie restriction conditions. It is possible that mRNA in the cell could be affected globally, altering downstream global protein metabolism. Future quantitative PCR (qPCR) or RNA quantitative mass spectrometry measurements are necessary to address this possibility.
Figure 5: RNA-Seq gene count comparisons between dietary cohorts. In the top three panels, each of the calorie restriction cohorts (y-axis) is compared with ad-libitum diet (x-axis) for genes matched between the two diets in each graph. In the bottom three panels, the three calorie restriction cohorts are compared to each other: CR PROT vs. CR, CR AGE vs. CR, and CR AGE vs. CR PROT. Red dashed lines indicate linear regression fits (see Table S2 in Appendices for regression line statistics), while dotted blue lines indicate identity lines with a slope of one. Error bars indicate 95% confidence intervals.
In calorie restriction, mitochondrial and ribosomal protein regulation is significantly different from global regulation of translation.

Thus far, the evidence from the proteomic and transcriptomic measurements indicated that the entire proteome was experiencing a global shift in post-transcriptional regulation. While this was not unexpected, and is indeed significant for the consideration of calorie restriction mechanisms and effects, we asked whether any protein groups or organelles were being perturbed differently from the global regulation, and what these could tell us about how calorie restriction works in cells. To this end, we used the previous linear regressions to generate predicted values for protein turnover rates, quant, and RNA-Seq counts: the control value for each protein was input into the regression line formula as the x-value, and the output y-value was computed as the predicted value for the other diet. These predicted values were then compared to the actual values measured experimentally, and the difference between the predicted and experimental values calculated and normalized, as detailed in the Methods. A paired student’s t-test was used to generate p-values, and a follow-up q-value test was used to minimize false positives due to multiple comparisons. We then selected those significant results that fell 20% or more outside of the predicted range for further analysis. The volcano plots in Figure 6 below summarize this process graphically.
Figure 6: Experimental versus predicted values for turnover rates, quant, and RNA-Seq counts. Q-values for the comparisons between experimental and predicted values for the various measurements are plotted on the y-axis, while the x-axis shows fractional change between the experimental and predicted values. The horizontal dashed blue line indicates a Q-value of 0.05 on the phred scale, and the red vertical lines indicate the cutoff of a +/- 20% difference. Numbers of significant proteins are indicated; these numbers may or may not correlate with the displayed points on the graph, as some of those displayed points were eliminated by the q-value test, but still included in the graphic. Outliers past the ranges of the x-axis and y-axis are included in the numbers, but were not displayed in order to keep the plot scales uniform. Calorie restriction (CR) versus ad-libitum (AL) is shown across the top row, supplemented protein (CR PROT) versus calorie restriction (CR) is shown in the second row, CR AGE versus calorie restriction (CR) is displayed in the third row, and the bottom row shows CR AGE versus supplemented protein (CR PROT) calorie restriction.
An ontology analysis using the online tool DAVID\textsuperscript{20,21} was carried out on the significantly different protein groups, as described in the Methods. The results are shown in Figure 7, below. Mitochondrial proteins were higher in quantity than the global prediction under calorie restriction, while their turnover rates were lower than globally predicted. This indicates a selective decrease in the degradation of mitochondrial proteins, possibly through the selective slowing of mitophagy. This correlated well with the behavior of mitochondrial proteins in the dietary restriction study of the previous chapter. Ribosomal proteins also displayed higher quantity than the predicted values under calorie restriction, indicating that they may be protected from the global slowdown of protein metabolism under calorie restriction. This also matched well with the ribosomal proteins in the dietary restriction study, though they did not pass all of the cut-offs for significant change in the dietary restriction ontology analysis. As in the dietary restriction study, this suggests that the ribosome and mitochondria are regulated separately from the global slowdown of protein metabolism in the cell under calorie restriction conditions in order to preserve qualities of protein synthesis and energy metabolism that help the cell better adapt to nutritional stress.

The ontology analysis did not reveal any enriched groups of proteins that were significantly different between the calorie-restricted dietary cohorts (CR, CR PROT, and CR AGE) in any of the data sets, as shown in Figure 7, below. Though disappointing, this confirms the similarity between these groups that can be seen in the global measurements above, and reinforces the conclusion that adding a small amount of protein or AGE to a CR diet does little to alter the protein metabolism within the cell. Poor overlap of protein identifications between data sets could also be a contributing factor, and though analysis filters were set conservatively to
ensure high quality data, this may have had the effect of decreasing identifiable protein ontology groups.
Figure 7: Ontology analysis results. Proteins with experimentally measured values that were significantly different from predicted values in the three data sets (see Figure 5 above) were input into the online tool DAVID\textsuperscript{20,21} for ontology analysis. Results after Benjamini-Hochberg correction with values of 0.05 or lower were judged to be significant. These groups are listed in the table above, separated by dietary cohort comparison, measurement type, and direction of change from predicted values. Box plots of the mitochondrial and ribosomal protein data for CR versus AL are shown below the table; note that although “membrane” is listed as a significant ontology category, most of these proteins overlapped with the mitochondrial classification, and were thus subsumed into the mitochondrial category. An analysis using the unique proteins to the mitochondrial and membrane categories showed no significant change in the non-mitochondrial membrane proteins from predicted values (data not shown).
Ribosomal subunits show differential degradation rates in CR, CR plus protein, and CR plus AGE dietary cohorts.

The dietary restriction analysis in the previous chapter revealed differential turnover and quantity of the large versus the small subunit of the ribosome. We adapted the same analysis here (a Mann-Whitney rank order test; see Methods) to see if a similar mechanism is present. As expected, the same phenomenon was detected; interestingly, the difference in turnover and quantity between the subunits appeared slightly exaggerated by the presence of protein or AGE supplementation in the diets. This correlates with the dietary restriction study, in which we see a greater difference between the subunits in the high protein diets compared to the low protein diets. However, the low numbers of measurements in the quant-MS data mean that the comparisons are not necessarily reliable. We have observed relatively large ranges for the turnover and quant values for the ribosomal proteins; therefore the box plots for the quant data cannot be compared at face value. The turnover rate data, however, is solid, and is consistent with the translational modulation through differential ribosomal subunit degradation model proposed in the dietary restriction study in the previous chapter.
Figure 8: Ribosomal subunit data. Turnover rate and quant-MS data are shown for ribosomal proteins, separated by location in either the large subunit or the small subunit. The numbers displayed at the medians of the box plots show the number of measurements for the corresponding box plot (note the low numbers for the quant-MS data). Notches indicate approximate 95% confidence intervals. Horizontal bars and asterisks indicate significance (one-tailed p-values < 0.05).
Mitochondrial, secreted, and endoplasmic reticulum proteins remain largely unperturbed by the addition of protein or AGEs to calorie restriction.

In the previous dietary restriction experiment (Chapter 2), ontology analysis showed that mitochondrial, secreted, and endoplasmic reticulum proteins were significantly changed versus the global regulation. For comparison, these protein ontologies were analyzed for the calorie-restricted groups in this study; the turnover data and quantitative MS data for these proteins are displayed as ratios in Figure 9 below. CR AGE versus CR, as well as CR AGE versus CR PROT, displayed no significant differences from predicted values (the poor match with the CR PROT/CR comparison resulted in low numbers of ratios, and therefore cannot be relied upon). Moreover, these ratios were not significantly different from each other. This implies that the addition of AGEs or proteins does not result in significant perturbations to the regulation of these protein groups. This correlates well with the global measurements from this study, and reinforces the robustness of the calorie restriction regimen to small additions of dietary protein or AGEs.
Figure 9: Mitochondrial, secreted, and endoplasmic reticulum protein rate and quant data. Turnover rate and quant data for the proteins from the significant ontologies found in the previous chapter’s dietary restriction study were extracted from the CR, CR PROT, and CR AGE cohorts. These are shown in the combined box and dot plots above. Colored horizontal lines indicate the ratio of expected values based on regression lines from Figures 2 and 3. Notches indicate approximate 95% confidence intervals.
Discussion

The major goal of this study was to determine the proteomic effects of AGEs on the calorie restriction diet. Given that AGEs are well-documented to elicit inflammatory reactions, and high amounts of dietary protein is shown to be attenuating for calorie restriction benefits, we expected that AGEs should have a measurable effect on the protein metabolism within cells, and that protein turnover rates and quantities would be significantly altered. Indeed, our own data from the high-protein dietary cohorts in the previous chapter showed significant deviations from the low-protein dietary cohort data. However, our results in the current study indicate that calorie restriction is a robust phenotype that is not significantly altered by small additions (0.14% by mass) of dietary protein or AGEs.

The calorie-restricted mice in our study exhibited the reduced overall protein turnover rates we saw in the dietary restriction study (Chapter 2), validating the experimental conditions of the AGE study, but this reduction carried over to the protein- and AGE-modified calorie restriction diets as well. There were also no significant differences in protein quantity or RNA, suggesting that any regulatory changes were minimal, and undetectable to the methods in this study. Significantly upregulated or downregulated ontology groups, identified in both this study as well as the previous dietary restriction study, were not significantly altered among the calorie-restricted cohorts, showing that while selective as well as global regulation perturbations were occurring under CR, they were roughly the same even with the addition of protein or AGEs.

There are several possibilities for the outcomes of this experiment. The first is that factors in the experimental design may have prevented us from properly detecting and measuring the proteomic impact of AGE signaling. The proteome coverage was not large (<10% of the
estimated proteome for turnover rate and quantitative MS measurements), filtering of the data decreased the coverage even further, and overlap between dietary cohort data sets for comparison analyses was often very poor. Future experiments should seek to increase proteome coverage by use of HPLC fractionation, increased replicate numbers, or improved instrumentation and data analysis.

After all of the experimental animals were sacrificed, it was noted that two necessary control groups, ad-libitum with AGEs, and ad-libitum with added protein, were never included in the experimental design. This limited our downstream analysis somewhat (for example, would there have been a change to the ad-libitum animal’s proteomes due to the addition of AGEs, and how would this compare to AGE-mediated changes in calorie restriction?), but for logistical reasons, it was not possible to add these control groups afterwards.

The dosage of additional protein and AGEs was very low (0.14% by mass), which likely muted the effect of the dietary additions. The addition of greater amounts of dietary protein and AGEs to the experiment is expected to magnify any proteomic effects of these additions (e. g. the high-protein cohorts from the previous study showed measurable differences to the low-protein cohorts, and the difference in dietary protein was nearly two-fold), but would complicate the ability to keep the various calorie restriction diets iso-caloric, as well as introduce other variables associated with decreasing the other sources of dietary calories. A nutritional geometry approach, as proposed by Simpson and Raubenheimer²⁵, may be necessary to deconvolute dosage effects of dietary protein and AGEs. This would also be informative as to the mechanism of AGE-mediated attenuation of calorie restriction: are these molecules acting primarily as signals (low-dosage is sufficient, with a plateau at higher levels), as an alternative fuel source (dosage-dependence, with magnified effects at high levels), or other mechanism (perhaps
evidenced by a maximal effect at a moderate concentration)? A metabolic labeling experiment involving isotopically labeled AGEs would also shed light on the mechanism (i.e., the presence of the isotopic label within other molecules in the cell would be dependent on pathways involved in AGE metabolism).

Future studies may also examine AGE signaling on other organs, such as the brain, spleen, kidney, and lung. Our study focused only on the liver, as it has several advantages for such a proteomics study: it is large, relatively homogeneous, and central to organismal metabolism. The lung would be especially interesting, as it is known to have more receptors for AGEs (RAGE) than any other tissue in the body; a comparison study with brain, spleen, and kidney, combined with the liver results would likely demonstrate different ranges of sensitivity and effects from dietary AGEs in various parts of the body. It may even be possible to show in which organ the attenuation of lifespan observed with AGE consumption first begins.

In summary, this study provided evidence that calorie restriction causes a robust proteome remodeling that is resistant to small amounts of AGEs. The dietary signaling effect of these molecules at low dosage was not sufficient to significantly perturb the protein metabolism of liver cells under CR. Future studies will be necessary to establish a dosage threshold for the AGE effect, and elucidation of the AGE-mediated lifespan-attenuating mechanism.
Supplemental Data

Matched Protein List Analyses

In the analyses of the turnover rate (Figure 2), quant MS (Figure 3), and RNA-Seq data (Figure 5), proteins were directly matched between the two dietary cohorts being compared without regard for matching proteins to other dietary cohorts. This was done to ensure maximum overlap between the two tested cohorts and improve the quality of the direct comparison. The alternative analyses, with matched proteins between all sets, are shown below (Panels A, B, and C); the number of proteins that matched, 59, was judged too small to replace the previous analyses. The alternative analyses are otherwise identical to those in Figures 2, 3, and 5. An ontology listing using DAVID\textsuperscript{20,\textsuperscript{21}} for the 59 proteins was generated (entire mouse proteome used as background, Benjamini-Hochberg $< 0.05$ for inclusion), so that functional clustering/bias could be evaluated for these proteins. 28 of the proteins clustered as mitochondrial, so this group of proteins cannot necessarily be interpreted as representative of the cell as a whole. Despite this, the matched protein analyses, with the exception of the AL quant MS data (Figure B, top panels) resulted in linear regression constants that are very similar to the values calculated using the more inclusive direct comparison analysis method (Figures 2, 3, and 5).
A. Turnover Rate Data
B. Quant MS Data
C. RNA-Seq Data
References


7. Seidell, J. C., Obesity, insulin resistance and diabetes — a worldwide epidemic. *British Journal of Nutrition* 2007, **83** (S1), S5-S8.


“Imaging Regiospecific Lipid Turnover in Mouse Brain with Desorption Electrospray Ionization Mass Spectrometry”

Although this chapter represents a departure from proteomics into lipidomics, with the development of an imaging method using desorption electrospray ionization mass spectrometry (DESI-MS) to make kinetic measurements, the ultimate goal is still to deduce the regulation of proteostasis through the application of combined-omic technologies. In this study, we demonstrate the potential for measuring turnover rates in a two-dimensional tissue imaging technique, allowing for the separate analyses of synthesis and degradation of biomolecules. We began with lipids, as these were the easiest to detect and quantify, but future method development is expected to result in techniques that allow the measurement of proteins. Turnover and quantification measurements of proteins in a regiospecific manner would enable proteostatic analyses similar to those in the preceding chapters, but expand them to tissue- and organ-wide scopes that would also encompass extracellular functions. Data on this scale could then be combined to inform a holistic view of organismal phenotype dependent upon homeostatic controls that vary from tissue to tissue and organ to organ. Ultimately this would have the potential to link broad symptoms of disease to specific shifts in metabolism that can be studied in a systems-wide manner, uncovering information that would escape a cell-centered focus, providing access to previously untapped discovery grounds for novel diagnostic tools and therapies.

This paper was published in the Journal of Lipid Research in 2017. My contributions to the project included mouse work; tissue sample preparation; collaboration in initial method
development; participation in instrument development, operation, and trouble-shooting; identification of lipid targets, and subsequent verification by mass spectrometry fragmentation; and preparation of some figures for the manuscript. The document has been altered slightly to conform to the format of this dissertation, and figures have been renumbered to adjust for the incorporation of supplemental figures in the body of the main text. The paper is otherwise identical to the published version.
**Title:** Imaging Regiospecific Lipid Turnover in Mouse Brain with Desorption Electrospray Ionization Mass Spectrometry

Richard H. Carson*¹, Charlotte R. Lewis*¹, Mercede N. Erickson¹, Anna P. Zagieboylo², Bradley C. Naylor¹, Kelvin W. Li³, Paul B. Farnsworth¹§, John C. Price¹§

¹Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602,
²Current affiliation: Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114
³Current affiliation: Department of Nutritional Sciences and Toxicology, University of California, Berkeley, California 94720

* These authors contributed equally

§ Corresponding Authors:

Email: paul_farnsworth@byu.edu Paul B. Farnsworth Tel.: 801-422-6502 Fax: 801-422-0153

Email: jcprice@chem.byu.edu John C. Price Tel.: 801-422-6040 Fax: 801-422-0153

**Running title:** Imaging Regiospecific Lipid Synthesis with DESI-MS

**Abbreviations:**

Abstract: Compartmentalization of metabolism into specific regions of the cell, tissue and organ, is critical to life for all organisms. Mass spectrometric imaging techniques have been valuable in identifying and quantifying concentrations of metabolites in specific locations of cells and tissues, but a true understanding of metabolism requires measurement of metabolite flux on a spatially resolved basis. Here we utilize desorption electrospray ionization mass spectrometry (DESI-MS) to measure lipid turnover in the brains of mice. We show that anatomically distinct regions of the brain have distinct lipid turnover rates. These turnover measurements, in conjunction with relative concentration, will enable calculation of regiospecific synthesis rates for individual lipid species in vivo. Monitoring spatially dependent changes in metabolism has the potential to significantly facilitate research in many areas, such as brain development, cancer, and neurodegeneration.

Keywords: arachidonic acid, brain lipids, DESI, diet and dietary lipids, diagnostic tools, kinetics, molecular imaging.
**Introduction**

Lipid biosynthesis occurs in every cell and is a critical process for life. Lipids act as structural components, essential nutrients, and signals for processes as diverse as cell proliferation or apoptosis. Many lipids within the body are synthesized locally. Certain classes of lipids, though, cannot be synthesized by mammals, and some cells, like neurons, do not produce enough lipids to serve their needs. Therefore, dietary lipid availability and transport of lipids within the body are critical components of metabolism. Lipid transport is closely linked to multiple diseases, including heart disease and Alzheimer’s disease.

By measuring lipid abundance across a tissue surface, one can create images composed of lipid type and concentration. Imaging lipids within the structure of the brain and other organs has been used extensively to understand response to injury, and to detect cancers. Here we have used the Desorption Electrospray Ionization (DESI) imaging technique. In DESI, a fine spray of charged droplets in a high-velocity gas jet is directed toward a tissue sample. Dissolved analytes from the tissue become ionized, and are pulled into the mass spectrometer (DESI-MS). The mass spectrometer records the mass-to-charge ratio and intensity as the spot where the spray hits the tissue is rastered over the tissue surface. By correlation of the time-dependent mass spectrometric data with the motion of the tissue on a computerized stage, an image can be constructed that displays the types and relative concentrations of analytes at specific positions in the tissue sample.
Figure 1: Desorption electrospray ionization (DESI) is an ambient ionization mass spectrometric imaging technique. Electrospray droplets are formed when high voltage is applied to the solvent. Desorbed analyte is vacuumed into the mass spectrometer inlet, separated by m/z, and detected. After the sample is slowly rastered beneath the DESI source and mass spectrometer inlet, an image of the sample surface can be recreated from the mass spectral data.

Changes in lipid types or concentrations allow boundaries between tissues as well as regions within the tissue to be identified \(^5, 7\). The spectra from modern high-resolution mass spectrometers are exceptionally information rich, but even so, positive identification of all ions in a spectrum is rarely possible, and it is often difficult to unambiguously assign peaks to individual lipids \(^9\). As a result, many studies have been published using changes in unidentified features to differentiate between areas of an image \(^5b, 5c\). Even with this limitation, DESI-MS imaging has recently been shown to improve the efficacy of cancer surgery by highlighting the boundaries between tumor and healthy tissue \(^5, 7\).

An advantage of identifying and focusing on discrete lipid species when they can be positively identified is that we can assess the activity of biological pathways that produce those lipids. This is particularly useful in developing a rational intervention to prevent developmental defects, and
treat cancers or degenerative diseases. Trying to evaluate enzyme activity using only \textit{in vivo} concentration is problematic because the body modifies synthesis and degradation rates to resist changes in concentration. Therefore measurement of \textit{in vivo} rates is the most sensitive method for detecting changes in metabolism\textsuperscript{10} and may allow prognostic diagnosis of developing problems.

Here we show that using DESI in conjunction with metabolic labeling, we can measure turnover of individual lipid species in the mouse brain (Fig. 2). Further, we show that the turnover rates vary between physiologically recognizable structures. Although a large number of compounds can be detected in the mass spectra, in this proof of principle study we limited the analysis to four identified lipids. In principle this analysis can be conducted on large numbers of lipids simultaneously, and will be particularly useful in detecting fast growing tissues like cancer within the context of healthy tissues.
Materials and Methods

Figure 2: The workflow is shown here in chronological order. First, mice are fed 8% deuterated water over a course of different time periods to increase their deuterium concentration in the lipids within their bodies. Then the mice are euthanized, their organs are sliced to 50 μm thick, and the slices are thaw mounted on glass slides. The desorption electrospray ionization (DESI) source is rastered over each tissue slice over a period of approximately two hours. The mass spectral data files are converted initially from .d files to a MATLAB data cube. Individual m/z peaks are selected to see the spatial and regioselective nature of individual lipids. Isotope ratio images reveal differing lipid turnover rates within the organ structure for individual lipids, which are measured using incorporation curves.

Metabolic labeling

All experiments were performed under the approval of the Institutional Animal Care and Use Committees of Brigham Young University in conformity with the Public Health Service Policy on Humane Care and Use of Laboratory Animals. Female mice (C57/Bl6:Spt2) between the ages of 13-16 months were housed in groups of five. To initiate the experiment all mice received an intra-peritoneal (IP) injection of saline deuterium oxide (0.9% w/v NaCl, 99.8% D2O), 15 μL saline per gram body weight. After the IP bolus, mice were provided free access to food and deuterium enriched drinking water (8% molar enrichment) for the remainder of the experiment (Fig 2). This method caused an initial jump in deuterium enrichment up to 2 molar percent
excess (MPE) to initiate the experiment, followed by a rise to 5 MPE over the next 10 days. Urine was collected up to 10 days during the labeling period to build a profile of deuterium enrichment in each animal (Fig. 3). At the designated time points (0, 1, 5, 10, 20 and 40 days post IP), mice were anesthetized with carbon dioxide, and then euthanized via cardiac puncture to collect blood. Brain tissue was immediately collected. Cells were removed from the blood plasma via centrifugation (10 minutes at 800xG at 4°C). Brain tissue and blood plasma were flash frozen on dry ice and stored at -80 °C until needed.

Figure 3: Urine and blood deuterium enrichments for the mice used in this study were combined to measure an average body water turnover for these mice. There was an initial bolus to 2.1 molar percent excess (MPE), followed by an increase to 5.3 MPE. Body water turnover rate was 0.27 per day, resulting in a body water half-life of 2.5 days in these mice. This body water turnover rate was used to define a curve of deuterium enrichment for calculating turnover of the individual lipids.

Measurement of deuterium enrichment

Blood plasma and urine isotopic enrichments were measured in duplicate for all time points. In preparation for this measurement, aliquots of each sample were diluted in duplicate (1:300 in 18 MΩ water) and placed into the caps of inverted sealed screw-capped vials for overnight
distillation at 80°C. Similar to previous studies, the MPE deuterium in the distilled water was measured against a D₂O standard curve using a cavity ring-down water isotope analyzer (Los Gatos Research [LGR], Los Gatos, CA, USA) according to the published method ¹¹.

**Cryosectioning**

Brains were cut coronally immediately after dissection and flash-frozen on dry ice. To obtain slices for DESI, frozen forebrains were adhered using VWR Clear Frozen Section Compound to the chuck of a Cryostat Thermo Scientific Microm HM550, and flash-frozen in liquid nitrogen. 50 μm sections were sliced at -15 °C. Sections were thaw-mounted onto VWR Superfrost Plus glass slides, flash-frozen on dry ice, and stored at -80 °C. Prior to DESI imaging, the glass slides were dried at slight negative pressure for approximately 20 min at room temperature, as previously described ¹².

**DESI-MS data acquisition**

Sample images were gathered using a MicrOTOF II (Bruker Daltonics, Billerica, MA, USA) and a laboratory constructed DESI source. DESI parameters are presented in Table 1. Together, those parameters created a solvent spot that desorbed lipids and fatty acids from the brain tissue, releasing them into the atmosphere, where they were sampled by an extension to the capillary inlet of the mass spectrometer.
### Table 1. Instrumental parameters for DESI image acquisition

**DESI Source**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emitter voltage</td>
<td>-5kV</td>
</tr>
<tr>
<td>Solvent</td>
<td>100% methanol</td>
</tr>
<tr>
<td>Solvent flow rate</td>
<td>2.5-3.0 µl min⁻¹</td>
</tr>
<tr>
<td>N₂ gas pressure</td>
<td>160 psi</td>
</tr>
<tr>
<td>Emitter angle</td>
<td>55 degrees</td>
</tr>
<tr>
<td>Emitter tip to surface distance</td>
<td>1 mm</td>
</tr>
<tr>
<td>Emitter tip to inlet distance</td>
<td>4 mm</td>
</tr>
</tbody>
</table>

**Stage Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>x-axis scan rate</td>
<td>75 µm sec⁻¹</td>
</tr>
<tr>
<td>y-axis row spacing</td>
<td>150 µm</td>
</tr>
</tbody>
</table>

**Mass spectrometer parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ion mode</td>
<td>negative</td>
</tr>
<tr>
<td>Mass range</td>
<td>200 m/z to 900 m/z</td>
</tr>
<tr>
<td>Acquisition rate</td>
<td>1 spectrum sec⁻¹</td>
</tr>
<tr>
<td>Averaging</td>
<td>2 spectra</td>
</tr>
<tr>
<td>Inlet voltage</td>
<td>500 V</td>
</tr>
</tbody>
</table>
DESI image quality considerations

Acquisition of high-quality DESI images requires careful attention to a large number of experimental parameters and inevitably requires a compromise between sensitivity and spatial resolution. The experimental parameters presented in Table 1 represent such a compromise. The nominal pixel size was 75 μm × 150 μm. Additional important considerations include the complete elimination of gases from the solvent, the shape of the mass spectrometer inlet extension, and the age of the sample tissue. Since no preventive measures were taken to keep the brain cells intact at the surface of the tissue slices during cryosectioning, the DESI-MS data represent an average of all membrane and cytosolic lipid populations.

Analysis workflow (Image Inspector)

Acquisition files from the Bruker MicroToF were converted from .d files to .imzML using ProteoWizard MSconvert. The imzML was then converted to a data cube using a MatLab script adapted from the read_imzML script that was included as part of the omniSpect package published by the Fernandez group at Georgia Tech. It is worth noting that, despite attempts to ensure that imzML is a universal format, the imaging files derived from different proprietary manufacturers’ formats have idiosyncrasies that require attention when reading them into a custom software package. For example, the number of spectra in chronograms acquired by the Bruker instrument varies by one because of a lack of synchronization between an external start signal and the internal clocking of the data acquisition. We provide the imzML-to-cube script as a template for other users, but it has only been tested with Bruker data.

After creation, the brain data cube was then assembled into an image using a MatLab script we developed called Image Inspector (Fig. 2). Image Inspector allows flexible manipulation of the
data cube to examine and export mass spectra from individual pixels or regions of interest, to view and export concentration maps for selected masses, and to view and export maps of neutromer ratios. Portions of Image Inspector were also adapted from the omniSpect library. Areas defined by a region of interest tool were summed for export and calculation of isotope labeling. Turnover kinetics of individual brain regions in each image were calculated from the time-dependent ratio of labeled to unlabeled lipids. The code and readme for installation and usage of the Image Inspector tool are provided as supplementary information.

**Identification of individual lipids**

In order to verify the lipid identities, purified lipids were ordered from Sigma-Aldrich and used as standards: arachidonic acid (A3611-10MG); cis-4,7,10,1,16,19-docosahexanoic acid (53171-10MG); 1,2-diacyl-sn-glycero-3-phospho-L-serine (P7769-5MG); L-alpha-phosphatidylinositol ammonium salt solution (P2517-5MG). While the AA, DHA, and PI standards were exact matches for the lipid identifications we tested, it is worth noting that the PS standard turned out to be a different phosphatidylserine species, with a lower molecular weight than the one we observed in our DESI brain scans. For DESI scanning, a few drops of standard in solution (~10mg/mL in chloroform) were placed onto VWR glass slides via pipette, and fragmentation was accomplished as described below for the brain slices.

MS/MS fragmentation spectra (Fig. 4) were collected in negative ion mode for ions at 303.25, 327.25, 834.5 and 885.6 m/z, using an Agilent QToF mass spectrometer equipped with a DESI source. A metabolically unlabeled 50 μm thick brain slice was scanned in targeted MS/MS mode three times at each of three different collision-induced dissociation (CID) fragmentation energies (10, 20, and 40 EV). Nitrogen was the collision gas. Each scan lasted approximately 3 minutes;
during the scans, the brain slice was manually moved under the DESI spray (100% methanol at 3 mL/min.) in order to maximize total ion counts. The capillary voltage was set at 1000 V, and MS1 and MS2 data were collected at a maximum rate of 4 spectra/second (other settings: 250 MS/spectrum, narrow isolation width ~1.3 m/z). Fragmentation data at each collision energy were averaged to collect the MSMS spectra (supplemental information); experimentally observed fragments for tissue and standards were initially compared against the literature \(^8b\), then further analyzed using online tools from LipidMaps.org \(^{16}\) and Metlin.Scripps.edu \(^{17}\) to identify each of the lipid species.
Figure 4, A-D: Lipid fragmentation from brain samples compared to standards. Fragmentation spectra for lipids at 303.3, 327.3, 834.6, and 885.6 m/z at 40 eV collision energy (top spectra, each panel) matched with fragmentation spectra for the corresponding commercial lipid standards (bottom spectra, each panel, inverted) under identical DESI conditions.

Figure 4A: Fragmentation of 303.25 m/z peak and arachidonic acid standard. Despite not being able to identify structures to accompany all fragments, good agreement between sample and standard spectra allow us to confidently identify the lipid at 303.25 m/z as arachidonic acid.
Figure 4B: Fragmentation of 327.3 m/z peak and docosahexaenoic acid standard. Despite not being able to identify structures to accompany all fragments, good agreement between sample and standard spectra allow us to confidently identify the lipid at 327.3 m/z as docosahexaenoic acid.
Figure 4C: Fragmentation of 834.6 m/z peak and phosphatidylserine standard. Although the PS standard was not an exact match for the target in the brain samples (parent ion m/z of 788.5, see Methods), the lipid standard enabled us to confirm that the species at 834.6 m/z is a phosphatidyserine.
Figure 4D: Fragmentation of 885.6 m/z peak and phosphatidylinositol standard. Despite not being able to identify structures to accompany all fragments, good agreement between sample and standard spectra allow us to confidently classify the lipid at 885.6 m/z as a phosphatidylinositol.

Calculation of the number of deuterium sites (n)

As we\textsuperscript{18} and others\textsuperscript{19} have previously described for peptides, independent turnover rates can be measured for multiple analytes simultaneously using mass spectrometry and D\textsubscript{2}O labeling. The three main variables to consider when calculating the turnover rate (k), are the number of sites in each lipid that can incorporate deuterium from water (n), the D\textsubscript{2}O enrichment of the body water at the time the lipid was labeled (p), and the time-dependent labeled fraction (f) of the lipid being measured (32).
Table 2. The common name, observed ionic mass, and elemental composition of the 4 major singly charged ions observed in our DESI-MS spectra. The asterisk (*) next to the name indicates a statistically significant difference in the turnover rate (k) of lipids between the caudoputamen and cortex. The number of deuteriums (n) was the same regardless of the area of the brain. There were large differences in the amount of biosynthesized versus dietary lipid (%) for different lipid species.

<table>
<thead>
<tr>
<th>Lipid</th>
<th>m/z</th>
<th>Formula</th>
<th>n</th>
<th>k in CA</th>
<th>k in CO</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>arachidonic acid, *</td>
<td>303.25</td>
<td>C_{20}H_{32}O_{2}</td>
<td>6</td>
<td>0.046±0.005</td>
<td>0.055±0.007</td>
<td>40</td>
</tr>
<tr>
<td>DHA,</td>
<td>327.25</td>
<td>C_{22}H_{32}O_{2}</td>
<td>~8</td>
<td>0.057±0.019</td>
<td>0.045±0.010</td>
<td>32</td>
</tr>
<tr>
<td>phosphatidylserine, *</td>
<td>834.6</td>
<td>C_{46}H_{77}NO_{10}P</td>
<td>21</td>
<td>0.041±0.003</td>
<td>0.053±0.001</td>
<td>100</td>
</tr>
<tr>
<td>phosphatidylinositol</td>
<td>885.6</td>
<td>C_{47}H_{82}O_{13}P</td>
<td>27</td>
<td>0.072±0.005</td>
<td>0.081±0.005</td>
<td>73</td>
</tr>
</tbody>
</table>

The number of covalent deuterium sites (n) is unique for each lipid (Table 2). This value was calculated using the isotope pattern for the metabolically labeled form of each identified lipid and the known D_{2}O enrichment (Fig. 5). Highly labeled spectra for the individual lipids were collected by imaging brain tissue from mice collected after 40 days. We subtracted the unlabeled isotope pattern from the labeled spectra to isolate the specific changes due to the newly synthesized lipid. We found that using an experimentally-determined unlabeled background improved the calculations. We then calculated the deviation assuming a range of n values for the lipid. Three of the four lipids had a single best n which minimized the deviation between theory and experiment for each of the isotope peaks within the isotope pattern (Fig. 5) as previously described^{18,20}. DHA had a very small overall change in the isotope pattern that is affected by interference from nearby molecules. We were not able to find a unique best n value for DHA.

**Calculating in vivo lipid turnover**

Every time a new lipid is synthesized it will incorporate protons and deuterons from the body
water pool. The relative percentage of deuterium incorporated is dependent on the enrichment of the water at that time. The changing deuterium enrichment of the body water in these mice required that the effective deuterium enrichment be calculated based on the half-life of the analyte. Because human studies usually have an increasing non-linear body water enrichment, this study provided a good proof-of-principle test for the calculations that would be required for analyzing metabolically D₂O-labeled human tissue samples. The enrichment of the body water for this calculation was defined as the average of the urine and plasma values for all mice at a time point (Fig. 3). This body water curve provided the time-dependent precursor enrichment for the calculation of the $f$ values. The time-dependent change in normalized intensity $\Delta I_x$ for each mass in the neutromer pattern was calculated according to equation 1, for both experimental and theoretical spectra. As the initial point for the Bayesian calculation of $k$, we calculate the theoretical $\Delta I_x$ for the labeled molecule using the body water measured that day. The daily $f$ values were then calculated using the single pool rise to plateau kinetic shown in equation 2. The half-life from this initial $k$ value was then used to calculate the effective body water at each time point as the integral over that time period, which was used to calculate new theoretical $\Delta I_x$ and a new $f$ as described previously for protein. These $f$ values were used to fit a new $k$, which was used to calculate the new effective deuterium enrichment for the next round. This recursive optimization was conducted until the lipid half-life changed by less than 0.5 days.

Equation 1:

$$\Delta I_x = \frac{M_x \text{ signal}}{\sum M_i \text{ signal}}_t - \frac{M_x \text{ signal}}{\sum M_i \text{ signal}}_{t=0}$$

Equation 2:

$$f_x = \frac{\text{Experimental } \Delta I_x}{\text{Theoretical } \Delta I_x} = A e^{-k t}$$

We compared $k$ values in the cortex and caudoputamen of the brain by averaging the spectra in
those regions using Image Inspector, then calculating $k$ using the scipy package in python for the recursive, non-linear regression fit.

**Results**

**Establishment of lipid identities and n values**

In the DESI spectra, four singly-charged ions were consistently observed with significant signal intensities (Fig. 5).

![Graphs showing comparison of experimental and optimized simulations of labeled and unlabeled spectra](image)

**Figure 5:** Comparison of experimental and optimized simulations of labeled and unlabeled spectra(A). The number of deuterium sites was derived by minimizing the deviation between the simulated and experimental spectra when the number of covalent deuterium sites was allowed to vary (B).
We compared the parent m/z and MS/MS fragmentation data (Fig. 4) observed here against a literature source and tentatively identified them as arachidonic acid (AA, C20:4 \( \Delta 5,8,11,14 \), m/z = 303.25, C\(_{20}\)H\(_{32}\)O\(_2\)), docosahexaenoic acid (DHA, C22:6 \( \Delta 4,7,10,13,16,19 \), m/z=327.3, C\(_{22}\)H\(_{32}\)O\(_2\)), phosphatidylserine (PS, C40:6, m/z=834.6, C\(_{46}\)H\(_{77}\)NO\(_{10}\)P), and phosphatidylinositol (PI, C38:4, m/z=885.6, C\(_{47}\)H\(_{82}\)O\(_{13}\)P). Subsequent MSMS fragmentation of purified lipid standards verified the identification of AA, DHA, PS, and PI (Fig. 4).

The calculated \( n \) value for each of the lipids (Fig. 5) is consistent with their individual biosynthetic pathways. AA and DHA are both conditionally essential fatty acids. They are synthesized from the essential fatty acids linoleic (C18:2 D9,12) and linolenic acid (C18:3 D9,12, 15) respectively, but can also be obtained through the diet. Our results suggest that the addition of the C2 unit to make AA from linoleic acid includes the covalent addition of 6 protons from water (Fig. 6). During deuterium enrichment these 6 positions can incorporate deuterium leading to the \( n=6 \) value for this lipid (Fig. 5). DHA was less clear because the spectra were noisier, and the experimental \( \Delta I \) was very small. The change at M2 for DHA was confounded in many spectra by overlapping peaks from unidentified species. Although there were multiple \( n \) values mathematically (Fig. 5), the similarity to the AA biological pathway suggested that an \( n=8 \) value would be most appropriate (Fig. 6). It is highly likely that the PS and PI spectra represent two families of related structures where the unsaturated bonds are at multiple different positions in the fatty acyl tails. Our DESI fragmentation method was not capable of distinguishing between specific isomers that differ only in the placement of double bonds. However, in spite of this potential structural variation, we were able to calculate an \( n \) value of 21 for PS and 27 for PI (Fig.3). The glycerol backbone and the head groups of these lipids are synthesized from glucose, which would allow each of the carbon hydrogen bonds to incorporate...
a deuterium \(^2\text{H}\). For PS, this means that up to 8 deuterium positions are within the glycerol and serine. The 13 other deuterium sites are presumably distributed between the two fatty acyl tails. The inositol head group in PI and two fewer desaturations in the fatty acyl tails could easily account for 6 more deuterium positions observed in this molecule (Fig. 6).

**Figure 6:** **Stable carbon hydrogen bonds which could incorporate deuterium during metabolic labeling** are indicated with asterisks. The biologically predicted number of deuterium sites are close to the theoretical expectations based on the biosynthesis of AA and DHA. Because of the isobaric diversity that is likely present for the PS and PI lipids, it is not possible to specify the deuterium incorporation sites for these lipids in the hydrophobic tails.

**Spatially distinct changes in isotope ratio are specific to individual lipids and independent of concentration**

In the DESI images, signal intensity at a given m/z value is a measure of the relative
concentration of a particular compound (the neutromer distribution is a measure of the turnover of that compound). In addition to depending on the concentration of a compound in the tissue, DESI signals can be affected by a range of factors, including the tissue type, microscopic surface roughness, and changes in ambient conditions during image acquisition. Even with this inherent variability, the spatial distribution of lipid signal intensity was distinct for each targeted lipid and followed boundaries of the classical brain structures in a manner that is consistent with changes in concentration. Measurement of PS intensity in tissues collected throughout the metabolic labeling time course (Fig. 7, right column) showed that its concentration was consistently relatively high in the cortex and the caudoputamen, and lower in the corpus callosum.
Figure 7: Concentration/isotope ratio images for phosphatidylserine. A representative stained tissue slice is shown (top left), with the areas of the cortex and caudoputamen (CA) highlighted in the cartoon (top right).

Comparison of the change in intensity with change in isotope ratio for the four test compounds
across an image, shows that they are not necessarily linked for the different lipid species (Fig. 7, Fig. 8). Using PS as an example, if the change in isotope pattern is scaled linearly over the 40 days of the experiment (Fig. 7, global scaling), we see similar changes across the entire brain. This similarity suggests that newly synthesized lipid is present across the entire organ. Changing the scaling of the image to emphasize the range within a single image (independent scaling) shows that there is a small, but measurably faster rate in a subsection of the cerebral cortex and the septal nucleus. Importantly, the symmetry of the image is in agreement with the structure of the brain, reducing the possibility that this is a transient fluctuation in the image. The turnover measurement of each lipid is much more resistant to variations in ionization efficiency, because the heavy and light versions of the lipid are chemically identical and change similarly, preserving the ratio between them. This means that the turnover measurement is in principle more robust across an image than simple concentration measurements.

Figure 8, A-C: Concentration as the relative ion intensity (right hand column), and isotope ratio images scaled either within the timepoint (right column) or across all the timepoints (middle column) for arachidonic acid (AA, S4A), docosahexaenoic acid (DHA, S4B), and phosphatidyl inositol (PI, S4C). A representative stained tissue slice is shown (top left), with the areas of the cortex and caudoputamen (CA) highlighted in the cartoon (top right).
Figure 8A: Concentration/Isotope maps for AA
Figure 8B: Concentration/Isotope maps for DHA
The other lipids all showed similar, yet unique distributions in both concentration and turnover rate. AA, for example, is found at relatively high concentrations everywhere in the brain except the caudoputamen area (Fig. 8A). The rate of turnover was also slower in the caudoputamen. This is consistent with the idea that AA is synthesized and degraded locally. DHA is found at relatively high concentrations in the cortex, but there is a very small and somewhat noisy change in the isotope ratio, suggesting that turnover is minimal anywhere in the brain (Fig. 8B). This is not surprising, given that DHA is conditionally essential in the diet,
but these results suggest that 65% of the DHA in the brain is sourced directly from the diet. PI has a similarly wide distribution of relatively high concentrations in the brain, except in the approximate area of the corpus callosum (Fig. 8C). Interestingly, the turnover of this lipid is fastest in the corpus callosum where the concentration is very low. This suggests that the lipid is either synthesized and immediately degraded in the area of the corpus callosum, or that it is synthesized and distributed to the rest of the brain from the corpus callosum.

**Turnover calculations highlight differences in metabolism between lipids**

Using the lipid-specific $n$ value and the measured body water, we can calculate the time-dependent percentage of the molecules that have been labeled. This allows us to directly compare metabolic rates between the different lipids. We found that using the average body water curve and combining all the time dependent measurements from the individual mice to constrain the turnover calculation dramatically improved the confidence in the fit. This is consistent with the fact that these mice were all genetically identical and of a similar age. Therefore, they can be considered biological replicates for these metabolic measurements.
Figure 9: Deuterium enrichment of all four lipids included was compared in different areas of the brain. The distribution of enrichment in individual pixels in the Cortex (gray inset) versus the Caudoputamen (black inset) varied according to lipid and area of the brain (Day 20 shown). Turnover rates were calculated using kinetic curves composed of 3 separate images from 2 mice at each time point.

As suggested by the images (Fig. 7, Fig. 8), we found that there was a small but measureable difference in turnover rate between different regions of the brain for the lipids AA and PS (Fig.
9, cortex vs caudoputamen). PI turned over more rapidly than the others and did not have a significant difference between the cortex and caudoputamen (Table 2). Interestingly, the turnover kinetics also highlighted differences in the source of the lipids. For example, the incorporation curve for PS had a rate of approximately 5% per day and increased over the 40 day experiment such that ~100% of the lipid was new by the final measurements. PS is evidently formed from fully biosynthesized precursors, while the majority of AA and DHA are sourced from the diet. PS, AA, and DHA curves had similar rates of incorporation (~5% per day), but the amount of labeled AA and DHA saturated at much lower amounts than PS (Table 2). The saturation percentage reflects the relative contributions of biosynthesis and dietary sources for each lipid species. *De novo* synthesis and dietary sources of lipids are known to both be critical for the development and health of the brain. The uniform kinetic plateau across the image for each lipid also suggests there is no spatial differentiation in the source, i.e. diet versus de novo synthesis.

**Discussion**

We have shown that DESI can be used to image spatial regulation of metabolism in the mouse brain. Although any imaging technique could potentially be used to monitor metabolism using these methods. DESI is convenient as an imaging system because the sample preparation is fast and inexpensive, with simple atmospheric pressure ionization. A disadvantage of DESI is the variation in ionization due to chemical and/or physical variability in the tissue samples. Because the sniffer (Fig. 1) is held so close to the tissue surface, we have seen that if the 50 μm tissue slice has ragged edges or variable thickness, the sniffer position relative to the sample surface changes during a scan, causing biases in the image.
Imaging lipids is convenient because they ionize well in DESI, and because lipid metabolism lies at the core of multiple important diseases such as neurodegeneration \(^{2b, 3a, 4, 21}\) and cancer \(^{5b, 10, 22}\). Improved identification of lipids is an important direction for future development in biochemistry because of their roles in metabolism and signaling. Identification of lipids is difficult; grouping mass spectral features into (sometimes broad) lipid classes is the current state of the art \(^9\). For this proof-of-principle study we monitored the turnover of four previously identified lipids, but the techniques can be applied to any identifiable molecule.

There have been many excellent studies investigating metabolism of lipids in the brain \(^{23}\). To our knowledge we are the first to image metabolism of specific lipids. By identifying the individual lipids, we unlock the ability to investigate the metabolic pathways of each lipid. The \textit{in vivo} turnover reports the rate-limiting step for metabolism. Interestingly, use of radiolabeled AA and DHA has shown that phospholipid incorporation and turnover occurs over a range of about 2-8\% per day \(^{24}\), in agreement with our observations. The use of radioactivity also precludes assessment of the percent of the total pool that is synthesized. Using stable isotopes, the experiment measures turnover rate, as well as the relative contribution of dietary sources to the metabolism of the tissue. We observed that the PI turnover is faster than the PS, but that there is a portion of the PI pool within the brain that does not display turnover. This could be due to a dietary supply of PI, or to brain cells which are preserving the PI in a specific structure. We favor the dietary supply hypothesis because this 80\% limit was uniform across the brain, and both AA and DHA, which are known to be sourced from the diet to supplement biosynthesis \(^{3a}\), show similar trends (Fig. 8).

Both biological and instrumental factors cause variations between images, and in the symmetry of individual images. As mentioned above, changing ionization efficiency in the DESI source
introduces variations that are particularly evident in the concentration images. The section of the brain that is imaged varies somewhat as well because it can be difficult to match the exact location of the tissue slice within the brain for each time point. The isotope analysis feature of Image Inspector is valuable for analysis of both biological and instrumental noise. In unlabeled samples it provides the analyst with an accuracy metric that can be used to evaluate image quality. Overlapping isotopic signatures can be identified by looking at the intensity patterns across the brain. For example, there was a significant change in the observed distribution of the 888 m/z neutrometer of PI across the brain relative to the monoisotopic 885 m/z (Fig. 10). This was indicative of an overlapping lipid species, i.e. the low purity of the mass spectrum. Therefore, the kinetics calculations didn’t include the 888 or heavier m/z neutromers. We also found that by monitoring bias in the neutrometer pattern we have rational criteria for setting signal thresholds for each analyte in each image. This is a continuous quality analysis which can be used to rapidly identify noise in the instrument or bias from closely overlapping signals, which can vary with time or tissue composition.
Figure 10: Signal purity tested by spatial distribution of neutromers. The isotope pattern of the phosphotidyl inositol (PI) overlays a second unidentified species at 888.7. The M0-M2 neutromers of PI (885.6, 886.6, 887.6) have the same distribution (left image). Moving from the left hand side to the right hand side of the 888 peak produces a very different intensity distribution.

The next step in developing this kinetic imaging method will be to focus on improved quantitation in combination with the turnover of the individual lipids. This would allow investigators to distinguish between changes in synthesis, transport, or degradation of lipids. Expanding the imaging to a wide variety of other analytes, including other metabolites and proteins, would dramatically expand the utility of this method. This allows us to consider future
experiments looking into the effect of diet on brain development and health, or investigating the effect of cancer therapies that target lipid metabolism.

Acknowledgements

The authors would like to thank members of the Facundo Fernandez group at Georgia Institute of Technology, particularly Rachel Bennett Stryffeller and Martin Payne, for their assistance in designing and setting up the DESI imaging system. We would also like to thank Dr. Benjamin Bikman of the Physiology and Developmental Biology Department at Brigham Young University for donating the mice used in this study. We thank Brigham Young University and the Fritz B. Burns Foundation for financial support.

References


CHAPTER 5

Conclusion

Addressing age-related diseases is one of the most urgent healthcare needs of our time. Without a comprehensive understanding of the etiology of aging and its biochemical pathways, our ability to deal with these illnesses will be limited. In this dissertation, I addressed one of the known hallmarks of increasing age, loss of proteostasis, with a novel combined omics methodology that included an assessment of protein flux within the cell. Combined with traditional quantitative proteomics and transcriptomics, we were able to gain insight into global proteostasis, separating the processes of protein synthesis and protein degradation for thousands of proteins. We applied this methodology to calorie restriction, a model system for aging studies, and were able to recapitulate many of the known findings regarding calorie restriction, validating our approach. Global translation was seen to slow, with exceptions for select protein groups in the mitochondria, endoplasmic reticulum, and secretome. Increasing dietary protein altered the turnover and quantitative levels of the proteome, demonstrating that calorie restriction was a result of integrated dietary signaling, with interplay among multiple biochemical pathways. Transcriptomics enabled us to narrow these changes to post-transcriptional regulation, and a subsequent focus on protein translation led us to hypothesize a novel model in which differential degradation of ribosomal subunits (likely resulting from sequestration of the 43S pre-initiation complex into stress granules) contributes to decreased protein synthesis in the cell, enabling improved proteostasis.

Future experiments will be necessary to test whether this mechanism is correct and contributes to calorie restriction benefits. While we did gather data on rRNA turnover rates, we
would need to measure turnover rates for each of the four rRNA molecules in the ribosome to provide crucial orthogonal data for testing the differential degradation hypothesis. Ribo-SEQ experiments would provide additional measurements of translational speed, allowing corroboration of our kinetic proteomics method. Additional longevity studies under identical conditions would test how well the omic measurements translate to the traditional phenotypic readouts of calorie restriction regimens. In particular, based on the literature, the addition of dietary protein and advanced glycation end-products should attenuate calorie restriction induced longevity; a nutritional geometry-based approach would be expected to be fruitful, especially for establishing dose dependency. Other organs in the body, such as the brain, kidney, lung, and spleen, likely react differently to calorie restriction; a comparative analysis between the liver and these others may uncover alternative pathways of proteostasis modulation under calorie restriction.

However, further optimization of the current method would probably be necessary to increase proteome coverage and improve signal-to-noise ratios. In order to ensure high quality in our data sets, we were forced to discard relatively large numbers of measurements due to uncertainty and noise. Combined with incomplete proteome coverage, this significantly decreased overlap between cohort protein data sets, and undermined our ability to make large numbers of comparisons between individual proteins. Increasing the number and types of comparative analyses will require much greater overlap in proteome coverage in order to be meaningful.

But we will eventually want to expand the methodology to incorporate other omics technologies, such as metabolomics and lipidomics, because it has so much potential. Kinetic metabolomics would enable the dynamic study of smaller molecules within the cell such as
amino acids and nucleotides, two species vital for anabolic processes and energy metabolism within the cell. The expansion of scope to incorporate these smaller, shorter-lived species would greatly inform our understanding of proteostasis within the cell, and the adaptation of the methodology would be very straightforward. Indeed, by using deuterated water, we were able to label multiple molecular species within the cell, and measure the turnover of DNA and lipids in addition to proteins. We have already shown the feasibility of dynamic lipidomics in our DESI-MS study (Chapter 4), opening a new direction in the study of lipid metabolism in brain tissue. This study also pioneered using kinetics in two-dimensions, which holds great promise for even greater detailed understanding of the regional differences in the dynamics of metabolism in tissues.

Ultimately, the goal is to progress to studying aging and age-related diseases in humans in order to advance clinical diagnostic and therapeutic practices. Since deuterated water is non-toxic in small amounts, the barrier to transfer of the methodology to human studies is low, and our lab has already begun kinetic proteomic studies in humans. In addition to applications related to the etiology of aging and its associated diseases, the next step could be studying the effects of therapeutics designed to delay or reverse aspects of aging. For example, metformin, resveratrol, aspirin, and rapamycin have all been studied, to one degree or another; an unbiased combined omics approach including kinetic proteomics may reveal much about pathways of aging affected, and suggest new therapies. Future application of the methodology in humans may also make possible personalization of anti-aging therapies to account for differences in genotype, health, and metabolism.

Clearly, realization of this goal will take time; for now, we still need methodology development, and more research into the phenomenon of aging. In this dissertation, I have
described a combined omics methodology that provides a systems-wide view of metabolism within the cell, with the ability to separate protein synthesis from protein degradation rates, and discern broad proteomic regulatory levels in just three measurements. This has enabled me to characterize a previously unreported phenomenon in calorie restriction, and hypothesize a mechanism that may later become exploitable for therapeutic agents. Combining other omics technologies with kinetic proteomics has the potential to become foundational for the study of cellular metabolism, and holds great promise for yielding future scientific insights that lead to improved human health and lifespan.
R Scripts for Data Analysis

Listed below are scripts written in R used to analyze and present the data described in the preceding chapters. More information on the analyses can be found in the Methods sections for Chapter 2 and Chapter 3.

RNA-Seq data analysis code

The following code was used to obtain gene counts from RNA-Seq data.

R Code:

# Import Bam files and Assign to features

library(Rsubread)

# This is set up to find all bam files in all subdirectories of set directory

p <- "/Users/richardcarson/Desktop/BYU/Price Lab/RNA Sequencing/Bamfiles/"

f <- list.files(path = p,

    pattern = ".bam$",

    recursive = T)

f <- paste0(p, f)

rsub <- featureCounts(f,
annot.ext = "Mus_musculus.GRCm38.86.gtf",

isGTFAnnotationFile = TRUE,

nthreads = 3,

minMQS = 10)

rsubdat <- rsub$counts #extracts the count table from the featurecount object

#name columns

                        "B3", "G1", "d2.1", "D1", "A3", "h2.2", "E1", "G2", "A1", "h2.3",
                        "e2.1", "E2", "C1", "H3", "H1", "B2", "F2", "E3", "A2", "e2.2")

## Get normalized values and filter for diff expressed gens

library(DESeq2)

dds <- DESeqDataSetFromMatrix(countData = rsubdat,
                               colData = data.frame(Diet=gsub("\d.*", "", colnames(rsubdat))),
                               design = ~ Diet)

dds <- DESeq(dds)

# ANOVA-like test

ddsLRT <- nbinomLRT(dds, reduced = ~1)
rlogvals <- assay(rlog(ddsLRT)) #gets rlog normalized values

r_all_logvals <- rlogvals #this is the data set without filtering for diff-expressed genes

r_all_logvals_filtered <- r_all_logvals[(!is.na(results(ddsLRT)$padj)), ] #filters out NAs

rlogvals <- rlogvals[(!is.na(results(ddsLRT)$padj)) &
                      results(ddsLRT)$padj < .05,
                      ] #ANOVA filter applied

colnames(rlogvals) <- colnames(rsubdat)

nrow(rlogvals)

**Confidence interval calculator function**

Since the majority of the data presented in the previous chapters consists of ratios, either
due to normalization or as comparisons between two different experimental groupings, and the
computation of confidence intervals for ratios of empirical values is non-trivial, I wrote this
function in R, and applied it to all instances when I needed to calculate the 95% confidence
interval of a ratio. I used the following URL’s as references:

Motulsky).

**R Code:**
calculate_CI_Q <- function(x_mean, y_mean, x_SD, y_SD, x_num_meas, y_num_meas) {
    # Set the t-ratio based on total degrees of freedom in x and y (for 2 replicates in Quant, this will be 2 DF), as well as desired confidence (95% standard)
    if(!is.na(x_num_meas) && !is.na(y_num_meas)) {
        if((x_num_meas+y_num_meas)>2) {
            t <- qt(0.975, df=(x_num_meas+y_num_meas-2))

            # Calculate the standard errors
            if(!is.na(x_SD) && !is.na(y_SD) && !is.na(y_mean)) {
                x_SE <- x_SD/(x_num_meas^0.5)
                y_SE <- y_SD/(y_num_meas^0.5)

                # Calculate an intermediate variable, g
                g <- (t*(y_SE/y_mean))^2
            }
            else
                g <- 2
        }
        else
            g <- 2
    }
    else
        g <- 2
}
# Set g to an arbitrary number greater than 1, since df came out to be less than 1

g <- 2

# Check to see if the CI of the ratio can be calculated - this is derived from url references

if(g > 1) {

    Outcome <- "Could not be calculated: g > 1"

    Q_upper <- "N/A"

    Q_lower <- "N/A"

    Q_CI_margin <- "N/A"

} else {

    # Calculate the standard error of the quotient

    Q_SE <- (((x_mean/y_mean)/(1-g))*((((1-g)*x_SE^2/x_mean^2)+(y_SE^2)/(y_mean^2))^0.5))

    # Calculate the confidence interval of the quotient

    Q_upper <- ((x_mean/y_mean)/(1-g))+(t*Q_SE)

    Q_lower <- ((x_mean/y_mean)/(1-g))-(t*Q_SE)

}
Q_CI_margin <- (Q_upper-Q_lower)/2

Outcome <- "Success: g < 1"

else {
    Outcome <- "Could not be calculated: missing values"
    Q_upper <- "N/A"
    Q_lower <- "N/A"
    Q_CI_margin <- "N/A"
}

# Enter the interval into an output construct

ratio <- x_mean/y_mean

CI_function_output <- list(Outcome, Q_upper, Q_lower, Q_CI_margin, ratio)

return(CI_function_output)

---

P-value and Q-value computation for discerning between selective and global regulation
In order to separate protein data (rate, quantitative MS, RNA-Seq, k-synth, and k-deg) that was significantly different from the global regulatory shifts observed in Chapter 2 and Chapter 3, a paired t-test between predicted and measured values was run using the R code below. In order to correct for false positives arising from multiple comparisons, a q-value test was also run at a 1% false discovery rate. P-values and q-values were then output to .csv values for further analyses.

**R Code:**

```r
# Set up a for-loop to go through and do a t-test by row

library(BSDA)

library(rlang)

library(qvalue)

ttest_data_set_pvalues <- data.frame()

for(i in 1:nrow(ttest_data_set)) {

    x_mean <- as.numeric(ttest_data_set[i,2]) #Mean predicted value

    s_x <- as.numeric(ttest_data_set[i,3]) #Standard deviation of predicted value

    n_x <- as.numeric(ttest_data_set[i,4]) #Number of measurements of the predicted value

    y_mean <- as.numeric(ttest_data_set[i,5]) #Mean measured value

    s_y <- as.numeric(ttest_data_set[i,6]) #Standard deviation of measured value

    n_y <- as.numeric(ttest_data_set[i,7]) #Number of measurements of the measured value

    # Perform paired t-test

    t <- (x_mean - y_mean) / sqrt((s_x^2 / n_x) + (s_y^2 / n_y))

    df <- n_x + n_y - 2

    se <- sqrt((s_x^2 / n_x) + (s_y^2 / n_y))

    p <- 2 * (1 - pt(abs(t), df))

    # Correct for multiple comparisons

    q <- qvalue(p)

    # Add p-values and q-values to data frame

    ttest_data_set_pvalues <- rbind(ttest_data_set_pvalues, data.frame(i, x_mean, y_mean, t, se, p, q))
}
```

197
new_pvalue_test <- tsum.test(x_mean, s_x, n_x, y_mean, s_y, n_y, alternative = "two.sided", mu=0, var.equal = FALSE, conf.level = 0.95)

new_pvalue <- new_pvalue_test$p.value

if (!is.na(new_pvalue)) {
    new_row <- c((ttest_data_set[i,]), new_pvalue)
    ttest_data_set_pvalues <- rbind(ttest_data_set_pvalues, t(new_row))
}

print(i)

}

colnames(ttest_data_set_pvalues) <- c(colnames(ttest_data_set), "t_test_p_values")

# Run the q-value test
p_values <- as.list(as.matrix(ttest_data_set_pvalues$t_test_p_values))

# Note: if errors, set pi0=1 for the Benjamini-Hochberg procedure instead
qobj <- qvalue(p = as.vector(p_values, mode = "numeric"), fdr.level = 0.01)

**Synthesis and degradation rate constant ontology analyses**

Below is the R script used for analyzing the k-synth and k-deg ontology data. While the computation of k-synth and k-deg values and their confidence intervals is straightforward, the
ontology analysis data from DAVID required further processing. The following code outputs
gene ontologies that are significantly different between experimental groups, and pairs them with
the k-synth and k-deg values for the lists of genes unique to each category.

**R Code:**

```r
# R ontology data cleanup

# Import the .csv

k_ontology_data <- read.csv("/Users/richardcarson/Desktop/BYU/Price Lab/Lab Papers/Super Paper/NIH31 vs Teklad8604/Ksynth_and_Kdeg_for_JC_grant/Annotations_for_ksynth_and_kdeg_R_import.csv", stringsAsFactors = FALSE) #Contains ontology data output from DAVID

library(tidyr)

library(plyr)

library(dplyr)

# Initialize global variables

temp_list <- matrix()

cluster_list <- character()

output_ontology_clusters <- matrix()
```
# Loop through the data and pull out the unique genes for each annotation cluster

for (i in 1:length(rownames(k_ontology_data))) {

    current_gene_list <- k_ontology_data$Genes[i]  # Pulls out the unedited list of genes

    current_gene_list <- gsub(" ", "", current_gene_list)  # Removes white space

    current_gene_list <- strsplit(current_gene_list, ",")  # Splits the list based on comma delimiter

    current_gene_list <- do.call(rbind, current_gene_list)  # Transforms the list into a matrix

    temp_list <- cbind(current_gene_list, temp_list)  # Adds the list to the previous iteration's list

    temp_list <- unique(temp_list)  # Removes duplicates

    current_cluster <- k_ontology_data$Annotation_cluster[i]

    print(current_cluster)

    if (i==length(rownames(k_ontology_data))) {  # If at the end of the annotation list, immediately put the gene list into the output

        output_ontology_clusters <- rbind.fill.matrix(output_ontology_clusters, temp_list)  # Attach the list to the output matrix

        cluster_list <- c(cluster_list, current_cluster)

    }

}
else {  # If not at the end of the annotation list, continue

    next_cluster <- k_ontology_data$Annotation_cluster[i+1]

    if(current_cluster != next_cluster) { # Checks to see if the next gene list goes with
        the previous one, i.e. has the same annotation cluster designation

        output_ontology_clusters <- rbind.fill.matrix(output_ontology_clusters,
            temp_list) # Attach the list to the output matrix

        temp_list <- matrix() # Resets the list for the next cluster

        cluster_list <- c(cluster_list, current_cluster)
    }
}

# Set the output matrix row names to the cluster_list

output_ontology_clusters <- output_ontology_clusters[-1,] # Deletes the first row of the matrix, which is all NA

rownames(output_ontology_clusters) <- cluster_list # Assigns cluster numbers to the rows
# Delete duplicates in the ontology cluster lists

test_output_ontology_clusters <- matrix()

for (i in 1:length(rownames(output_ontology_clusters))) {

    temp_row <- list()

    temp_row <- output_ontology_clusters[i,]

    temp_row <- unique(temp_row)

    test_output_ontology_clusters <- rbind.fill.matrix(test_output_ontology_clusters, t(temp_row))
}

test_output_ontology_clusters <- test_output_ontology_clusters[2:45,]

rownames(test_output_ontology_clusters) <- cluster_list

output_ontology_clusters <- test_output_ontology_clusters

### Pair k-synth and k-deg values with gene names (reference: https://stackoverflow.com/questions/27171505/loop-for-multiple-vlookup-in-r)

# Input k-synth and k-deg data
# Match the data to the genes:

```r
cluster_ksynth_kdeg_matched <- as.data.frame(t(output_ontology_clusters)) # Transpose the output and make it a dataframe

colnames(cluster_ksynth_kdeg_matched) <- rownames(output_ontology_clusters)

cluster_list <- names(cluster_ksynth_kdeg_matched)
```

# Diet A (LP AL):

```r
cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "A_ksynth", sep = ".")]
<- k_synth_and_deg$A_k_synth[match(unlist(cluster_ksynth_kdeg_matched[,1:44]),
k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup for A_k_synth values

cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "A_kdeg", sep = ".")]
<- k_synth_and_deg$A_k_deg[match(unlist(cluster_ksynth_kdeg_matched[,1:44]),
k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup for A_k_deg values
```
# Diet C (LP DR):

```
cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "C_ksynth", sep = ",")][-]
k_synth_and_deg$C_k_synth[match(unlist(cluster_ksynth_kdeg_matched[,1:44]),
k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup
for C_k_synth values
```

```
cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "C_kdeg", sep = ",")][-]
k_synth_and_deg$C_k_deg[match(unlist(cluster_ksynth_kdeg_matched[,1:44]),
k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup
for C_k_deg values
```

# Diet F (HP AL):

```
cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "F_ksynth", sep = ",")][-]
k_synth_and_deg$F_k_synth[match(unlist(cluster_ksynth_kdeg_matched[,1:44]),
k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup
for F_k_synth values
```

```
cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "F_kdeg", sep = ",")][-]
k_synth_and_deg$F_k_deg[match(unlist(cluster_ksynth_kdeg_matched[,1:44]),
k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup
for F_k_deg values
```

# Diet G (HP_DR):
cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "G_ksynth", sep = ",")]<- k_synth_and_deg$G_k_synth[match(unlist(cluster_ksynth_kdeg_matched[1:44]), k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup for G_k_synth values

cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "G_kdeg", sep = ",")]<- k_synth_and_deg$G_k_deg[match(unlist(cluster_ksynth_kdeg_matched[1:44]), k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup for G_k_deg values

# Ratio of C/A (LP DR/AL):

cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "CvA_ksynth", sep = ",")]<- k_synth_and_deg$CvA_ratio_ksynth[match(unlist(cluster_ksynth_kdeg_matched[1:44]), k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup for C/A_k_synth values

cluster_ksynth_kdeg_matched[paste("Cluster", cluster_list, "CvA_kdeg", sep = ",")]<- k_synth_and_deg$CvA_ratio_kdeg[match(unlist(cluster_ksynth_kdeg_matched[1:44]), k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup for C/A_k_deg values

# Ratio of G/F (HP DR/AL):
cluster_ksynth_kdeg_matched <- k_synth_and_deg$GvF_ksynth[match(unlist(cluster_ksynth_kdeg_matched[,1:44]), k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup for G/F_k_synth values

cluster_ksynth_kdeg_matched <- k_synth_and_deg$GvF_kdeg[match(unlist(cluster_ksynth_kdeg_matched[,1:44]), k_synth_and_deg$ProteinID)] # Makes new columns, one for each cluster, and does a v-lookup for G/F_k_deg values

### Pull out ontology clusters that are significantly different from each other

# ANOVA test, or Kruskal-Wallis non-parametric ANOVA (kruskal.test()) to start

# For kruskal-wallis test, need to convert to a list of lists; run a FOR loop

data_to_analyze <- select(cluster_ksynth_kdeg_matched, contains("_GvF_kdeg"))

list_to_analyze <- list()

for (i in 1:length(colnames(data_to_analyze))) {
  temp_list <- data_to_analyze[,i]
  temp_list <- temp_list[!is.na(temp_list)]
  list_to_analyze[[i]] <- temp_list
kruskal.test(list_to_analyze) # This works if test is a list of lists with more than 1 element

# Note: CvA-ksynth p-value = 0.02994, CvA-kdeg p-value = 0.04766, GvF-ksynth p-value = 2.538e-10, GvF-kdeg p-value = 4.931e-15.

# To get which clusters are different, do a paired t-test: see t.test (library(stats)) or non-parametric: wilcox.test. Question to answer: which clusters are statistically different between diet comparisons?

# Set global variables - diets to compare and values (x/y)

dietx <- "G"

diety <- "F"

value_to_compare <- "synth"

# Pull data out of the file - make 2 objects, each with the same number and order of clusters

x <- select(cluster_ksynth_kdeg_matched, contains(paste("_", dietx, "_k", value_to_compare, sep = "")))

y <- select(cluster_ksynth_kdeg_matched, contains(paste("_", diety, "_k", value_to_compare, sep = "")))
# Loop through the columns and do the wilcox.test and t.test for each; append p-values to an
object

ttest_pvalue_results <- data.frame()

cluster_list <- sapply(strsplit(colnames(x), split=paste("_", dietx, sep=""), fixed=TRUE),
function(x) (x[1])) # Pulls out just the cluster numbers for labeling

for (i in 1:length(colnames(x))) {
    result <- wilcox.test(x[,i], y[,i], paired = TRUE, conf.level = 0.95, na.rm = TRUE) #
    Wilcoxon non-parametric paired t-test

    ttest_pvalue_results[i,1] <- cluster_list[i]

    ttest_pvalue_results[i,2] <- result$p.value # Get the p-value only
}

colnames(ttest_pvalue_results) <- c("Annotation cluster", "Wilcoxon test p-value")

**Linear Regression Statistics**

The following tables supply the 95% confidence interval (CI) and p-value statistics for
the linear regression analyses in Chapter 2 (Figure 4 and Figure 5) and Chapter 3 (Figure 2,
Figure 3, and Figure 5). These values were obtained using the summary() and confint()
functions in R.
**Table S1: Regression value statistics for Chapter 2.**

<table>
<thead>
<tr>
<th>Turnover Rates</th>
<th></th>
<th></th>
<th>Slope Lower CI</th>
<th>Slope Upper CI</th>
<th>Slope p value</th>
<th>Regression Intercept</th>
<th>Intercept Lower CI</th>
<th>Intercept Upper CI</th>
<th>Intercept p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>Regression slope</td>
<td>Slope p value</td>
<td>Intercept Lower CI</td>
<td>Intercept Upper CI</td>
<td>Intercept p value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LP DR vs. LP AL</td>
<td>0.7796</td>
<td>3.64E-179</td>
<td>0.018355</td>
<td>0.024114</td>
<td>8.18E-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP DR vs. HP AL</td>
<td>0.83055</td>
<td>1.00E-155</td>
<td>0.022875</td>
<td>0.030251</td>
<td>2.02E-09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quantitative MS</th>
<th></th>
<th></th>
<th>Slope Lower CI</th>
<th>Slope Upper CI</th>
<th>Slope p value</th>
<th>Regression Intercept</th>
<th>Intercept Lower CI</th>
<th>Intercept Upper CI</th>
<th>Intercept p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>Regression slope</td>
<td>Slope p value</td>
<td>Intercept Lower CI</td>
<td>Intercept Upper CI</td>
<td>Intercept p value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LP DR vs. LP AL</td>
<td>0.89373</td>
<td>1.49E-273</td>
<td>-0.077118</td>
<td>-0.058345</td>
<td>2.25E-15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP DR vs. HP AL</td>
<td>0.99355</td>
<td>4.97E-279</td>
<td>-0.11172</td>
<td>-0.082765</td>
<td>8.69E-14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RNA-Seq</th>
<th></th>
<th></th>
<th>Slope Lower CI</th>
<th>Slope Upper CI</th>
<th>Slope p value</th>
<th>Regression Intercept</th>
<th>Intercept Lower CI</th>
<th>Intercept Upper CI</th>
<th>Intercept p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>Regression slope</td>
<td>Slope p value</td>
<td>Intercept Lower CI</td>
<td>Intercept Upper CI</td>
<td>Intercept p value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LP DR vs. LP AL</td>
<td>0.99753</td>
<td>0.00E+00</td>
<td>0.018802</td>
<td>0.030195</td>
<td>1.22E-03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP DR vs. HP AL</td>
<td>0.98578</td>
<td>0.00E+00</td>
<td>0.099155</td>
<td>0.11283</td>
<td>1.74E-45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table S2: Regression value statistics for Chapter 3.

<table>
<thead>
<tr>
<th>Turnover Rates</th>
<th>Comparison</th>
<th>Regression slope</th>
<th>Slope Lower CI</th>
<th>Slope Upper CI</th>
<th>Slope p value</th>
<th>Regression intercept</th>
<th>Intercept Lower CI</th>
<th>Intercept Upper CI</th>
<th>Intercept p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CR vs. AL</td>
<td>0.6408</td>
<td>0.60456</td>
<td>0.67704</td>
<td>2.22E-134</td>
<td>0.043045</td>
<td>0.036707</td>
<td>0.049383</td>
<td>7.06E-35</td>
</tr>
<tr>
<td></td>
<td>CR PROT vs. AL</td>
<td>0.78865</td>
<td>0.74681</td>
<td>0.83049</td>
<td>3.54E-139</td>
<td>0.026124</td>
<td>0.019076</td>
<td>0.033172</td>
<td>1.44E-12</td>
</tr>
<tr>
<td></td>
<td>CR AGE vs. AL</td>
<td>0.78865</td>
<td>0.74681</td>
<td>0.83049</td>
<td>2.32E-151</td>
<td>0.026124</td>
<td>0.019076</td>
<td>0.033172</td>
<td>5.18E-01</td>
</tr>
<tr>
<td></td>
<td>CR PROT vs. CR</td>
<td>0.85399</td>
<td>0.80946</td>
<td>0.89851</td>
<td>1.59E-152</td>
<td>0.025352</td>
<td>0.017791</td>
<td>0.032913</td>
<td>1.08E-10</td>
</tr>
<tr>
<td></td>
<td>CR AGE vs. CR</td>
<td>1.0045</td>
<td>0.96561</td>
<td>1.0434</td>
<td>8.52E-205</td>
<td>-</td>
<td>-0.007046</td>
<td>0.0057537</td>
<td>8.43E-01</td>
</tr>
<tr>
<td></td>
<td>CR AGE vs. CR PROT</td>
<td>0.93759</td>
<td>0.90708</td>
<td>0.9681</td>
<td>2.76E-236</td>
<td>0.0039489</td>
<td>-</td>
<td>0.0012656</td>
<td>1.37E-01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quantitative MS</th>
<th>Comparison</th>
<th>Regression slope</th>
<th>Slope Lower CI</th>
<th>Slope Upper CI</th>
<th>Slope p value</th>
<th>Regression intercept</th>
<th>Intercept Lower CI</th>
<th>Intercept Upper CI</th>
<th>Intercept p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CR vs. AL</td>
<td>0.39729</td>
<td>0.0023333</td>
<td>0.79224</td>
<td>4.87E-02</td>
<td>0.22072</td>
<td>0.045159</td>
<td>0.39629</td>
<td>1.39E-02</td>
</tr>
<tr>
<td></td>
<td>CR PROT vs. AL</td>
<td>0.45092</td>
<td>0.069404</td>
<td>0.83245</td>
<td>2.07E-02</td>
<td>0.054955</td>
<td>-0.11977</td>
<td>0.22968</td>
<td>5.36E-01</td>
</tr>
<tr>
<td></td>
<td>CR AGE vs. AL</td>
<td>0.45189</td>
<td>0.11187</td>
<td>0.7919</td>
<td>9.35E-03</td>
<td>0.033405</td>
<td>-0.12265</td>
<td>0.18946</td>
<td>6.74E-01</td>
</tr>
<tr>
<td></td>
<td>CR PROT</td>
<td>0.94615</td>
<td>0.91766</td>
<td>0.97464</td>
<td>5.26E-</td>
<td>-0.14488</td>
<td>-0.17831</td>
<td>-0.11145</td>
<td>9.12E-</td>
</tr>
<tr>
<td>Comparison</td>
<td>Regression slope</td>
<td>Slope Lower CI</td>
<td>Slope Upper CI</td>
<td>Slope p value</td>
<td>Regression intercept</td>
<td>Intercept Lower CI</td>
<td>Intercept Upper CI</td>
<td>Intercept p value</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>---------------</td>
<td>----------------------</td>
<td>--------------------</td>
<td>--------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>CR vs. AL</td>
<td>1.0016</td>
<td>1.0002</td>
<td>1.003</td>
<td>0.00E+00</td>
<td>-0.010582</td>
<td>-0.021673</td>
<td>0.00050949</td>
<td>6.15E-02</td>
<td></td>
</tr>
<tr>
<td>CR PROT vs. AL</td>
<td>1.0028</td>
<td>1.0014</td>
<td>1.0042</td>
<td>0.00E+00</td>
<td>-0.014796</td>
<td>-0.025781</td>
<td>-0.0038101</td>
<td>8.30E-03</td>
<td></td>
</tr>
<tr>
<td>CR AGE vs. AL</td>
<td>1.0019</td>
<td>1.0006</td>
<td>1.0032</td>
<td>0.00E+00</td>
<td>-0.011121</td>
<td>-0.021269</td>
<td>-0.0009729</td>
<td>3.17E-02</td>
<td></td>
</tr>
<tr>
<td>CR PROT vs. CR</td>
<td>1.0007</td>
<td>1.0002</td>
<td>1.0012</td>
<td>0.00E+00</td>
<td>-0.00052267</td>
<td>-0.0047415</td>
<td>0.0036962</td>
<td>8.08E-01</td>
<td></td>
</tr>
<tr>
<td>CR AGE vs. CR</td>
<td>0.99931</td>
<td>0.99875</td>
<td>0.99987</td>
<td>0.00E+00</td>
<td>0.0067089</td>
<td>0.0022161</td>
<td>0.011202</td>
<td>3.43E-03</td>
<td></td>
</tr>
<tr>
<td>CR AGE vs. CR PROT</td>
<td>0.99817</td>
<td>0.99759</td>
<td>0.99874</td>
<td>0.00E+00</td>
<td>0.010607</td>
<td>0.0060107</td>
<td>0.015203</td>
<td>6.14E-06</td>
<td></td>
</tr>
</tbody>
</table>