METABOLIC AND PROTEOMIC ADAPTATIONS OF METASTATIC PROCESSES IN BREAST CANCER CELLS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Heesoo Jeong

August 2023

METABOLIC AND PROTEOMIC ADAPTATIONS OF METASTATIC PROCESSES IN BREAST CANCER CELLS

Heesoo Jeong, Ph. D.

Cornell University 2023

Developments of targeted therapies improved prognoses and outcomes of patients with localized breast tumors. However, treatment options become limited upon occurrence of metastasis, a spread of tumors to distant organs, as surgical removal of tumors are difficult and tumor cells resist to therapies. Tumor cells need to overcome hurdles such as limited access to oxygen, through hypoxia, and matrix detachment during metastasis. To better understand underlying mechanism enabling tumor cells to overcome the two stresses, metabolic and proteomic changes of breast tumor cells when faced with inhibited oxidative phosphorylation (OXPHOS), a model of hypoxia, and detachment are analyzed.

To explore the influence of nutrients on protein expression and that of protein expression on metabolism, Chapter 2 of this dissertation presents a review of previous findings in studies on vitamin intakes affecting tissue proteomes. Metabolic fluxes reflecting cellular utilization of metabolic pathways are quantified with stable isotope tracing, which needs correction for abundances of naturally occurring isotopes if mass isotopologue distributions are directly used to infer fluxes. Chapter 3 introduces a computational tool, PolyMID, which corrects for abundances of naturally occurring isotopes. Among metabolic fluxes, that through pyruvate carboxylase (PC) is reported

to be altered upon breast cancer-derived lung metastasis. However, work presented in Chapter 4 shows that the widely used method to quantify PC flux is inaccurate, and covers the development of a metabolic flux analysis-based approach with improved performance.

Utilizing the above-mentioned tools, the work presented in Chapter 4 continues with analyses of metabolic and proteomic changes in breast cancer cells upon OXPHOS inhibition and culture in suspension. Protein networks regulating metabolic fluxes are identified. PC flux and its positively correlated protein network upon OXPHOS inhibition are found to be negatively correlated in cells resistant to detachment upon culture in forced suspension. Exposure to hypoxia elevates anchorage-independence of the detachment-resistant breast cancer cells. Taken together, hypoxia and matrix detachment induce metabolic and proteomic changes affecting metastatic potential of breast cancer cells.

BIOGRAPHICAL SKETCH

Heesoo completed her Bachelor and Master of Science in Food and Nutrition at Sookmyung Women's University, Seoul, Korea. Her thesis was on antioxidative and inflammatory mechanisms of galactose-induced aging mouse model. Heesoo entered the doctoral program in Molecular Nutrition in the Division of Nutritional Sciences at Cornell University in 2018. She joined Dr.Nathaniel Vacanti's lab where she applied stable isotope tracing and proteomic analyses to study metabolic and proteomic changes of breast cancer cells. She combined experimental and computational analyses and pursued minors in genomics and computational biology. Heesoo received the Genomics Scholars Awards from the Center for Vertebrate Genomics at Cornell in 2021.

This dissertation is dedicated to my parents, Joon Rae Jeong and Min-Seob Yu, and my brother, Jeehun Jeong.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr.Nathaniel Vacanti, for endless support and guidance. I was able to come this far all because of your mentorship. Thank you for showing and teaching me how to be a good researcher and a good person. It was an honor and pleasure to be your first PhD student.

I would like to express sincere gratitude to my committee members, Drs. Frank Schroeder, Kimberly O'Brien, and Praveen Sethupathy. Your advice changed my academic and professional perspectives. I appreciate having opportunities to learn from you.

I also thank the Vacanti lab members, the Center for Vertebrates Genomics, and the Division of Nutritional Sciences communities at Cornell for providing a great learning environment. I sincerely thank fellow graduate students in the community for being great colleagues and peer mentors.

I would like to thank my mentors from Sookmyung Women's University, Drs. Hyuk-Sook Kim and Mi Kyung Sung, for encouraging and nurturing me in my early academic journey. I am very fortunate to have great mentors at my alma mater.

TABLE OF CONTENTS

Biographical Sketch……..…….v
Dedication……...……...vi
Acknowledgements……..……..vii
Table of Contents…….……viii
List of Figures……...……x
List of Tables…….……xi

Chapter 1. Introduction……...…....1
1.1 Stable Isotope Tracing Analysis and Metabolism……..……1
1.1.1 Quantifying Cellular Utilization of Metabolic Pathways……..1
1.1.2 Correcting for Naturally Occurring Isotope Abundances…….….2
1.2 Proteomics and Metabolism……3
1.2.1 Impact of Vitamin Intakes on Tissue Proteomes……..…….4
1.3 Breast Cancer and Metastasis……...…..5
1.3.1 Breast Cancer…….5
1.3.2 Metastasis of Breast Cancer……..….6
1.3.3 Anchorage-Independent Survival of Tumor Cells During Metastasis…...8
1.4 Metabolic Changes Promoting Anchorage-Independent Survival of Breast Tumor Cells…….8
1.5 Proteomic Changes Promoting Anchorage-Independent Survival of Breast Tumor Cells……...……..9
References……...……12

Chapter 2. Systemic Vitamin Intake Impacting Tissue Proteomes……16
Abstract……...…17
Introduction…….……17
Vitamin A……21
Vitamin B_1……...……23
Vitamin B_2……...……25
Vitamin B_3……..……26
Vitamin B_5……...……27
Vitamin B_6…….……..……28
Vitamin B_7…….……..……29
Vitamin B_9…….……..……29
Vitamin B_{12}…….……31
Vitamin C……33
Vitamin D……35
Vitamin E……39
Vitamin K……40
Conclusions……..…41
Reference……46

Chapter 3. Correcting for Naturally Occurring Mass Isotopologue Abundances in Stable-Isotope Tracing Experiments with PolyMID……..……..65
Abstract……..…….66
Introduction…….……66
Results……...……..70
Methods……..……79
Discussion…….……..82
References…….……..85

Chapter 4. Proteomic Regulation of Metabolic Plasticity in Breast Cancer Cells……..88
Abstract…….……..89
Introduction…….……89
Results……...…..92
Discussion…….……....106
Methods……...……..109
References…….……122

Appendix……127
Appendix to Chapter 1……...…127
Abbreviations……...……127
Appendix to Chapter 2……..……128
Abbreviations……...…128
Appendix Table 2.1…….……136
Appendix Table 2.2…….……136
Appendix Table 2.3…….……137
Appendix Table 2.4…….……138
Appendix Table 2.5…….……140
Appendix Table 2.6…….……140
Appendix Table 2.7…….……141
Appendix Table 2.8……..…141
Appendix Table 2.9……..…142
Appendix Table 2.10……...……142
Appendix Table 2.11……...……143
Appendix Table 2.12……...……143
Appendix Table 2.13……...144
Appendix to Chapter 3……..……145
Abbreviations……...…145
Appendix to Chapter 4……...……146
Abbreviations……...…146
Appendix Figure 4.1……...……147
Appendix Figure 4.2……...148
Appendix Figure 4.3……...……149
Appendix Table 4.1…….……150

LIST OF FIGURES

Chapter 2

Figure 2.1 Fat-Soluble Vitamin Structures……..……22
Figure 2.2 Water-Soluble Vitamin Structures……..……24
Figure 2.3 Schematic of Vitamin Involvement in Reactions of Central Carbon Metabolism……43

Chapter 3

Figure 3.1 Principles of Stable-Isotope Tracing……..……67
Figure 3.2 Correction of High- and Low-Resolution Data……..…77
Figure 3.3 Illustrations of Tracer Enrichment and Label Enrichment……..78

Chapter 4

Figure 4.1 PC Flux Quantification……..…93
Figure 4.2 Metabolic Adaptations Upon OXPHOS Inhibition……...96
Figure 4.3 Proteomic Regulation of Metabolic Function Upon OXPHOS Inhibition……...99
Figure 4.4 PC Flux and Proteome Profiles of Anchored and Suspension Cells….104

Appendix

Appendix Figure 4.1 Labeling of TCA Cycle Intermediates by [U-$^{13}C_6$]glucose……147
Appendix Figure 4.2 Proliferation and Mass Isotopologue Distributions…..…..148
Appendix Figure 4.3 Correlations of Flux Responses with Enzyme Abundance Responses Upon Inhibited OXPHOS……..…149

LIST OF TABLES

Chapter 2

Table 2.1 Summary of key findings……..…….38
Table 2.2 Summary of technical depth of orbitrap-, QTOF-, and Triple-TOF-based studies……44

Appendix

Appendix Table 2.1……..……136
Appendix Table 2.2……..……136
Appendix Table 2.3……..……137
Appendix Table 2.4…….…… 138
Appendix Table 2.5……..……140
Appendix Table 2.6……..……140
Appendix Table 2.7……..……141
Appendix Table 2.8……..……141
Appendix Table 2.9……..……142
Appendix Table 2.10……142
Appendix Table 2.11……143
Appendix Table 2.12……143
Appendix Table 2.13……144
Appendix Table 4.1……..……150

Chapter 1. Introduction

1.1 Stable Isotope Tracing and Metabolism

1.1.1 Quantifying Cellular Utilization of Metabolic Pathways

Metabolism is defined as biochemical reactions producing or consuming energy. Cells utilize metabolic pathways to chemically process nutrients to produce energy and synthesize macromolecules (1). Tricarboxylic acid (TCA) cycle is the major metabolic pathway cells utilize to produce reducing equivalents to feed into the electron transport chain (ETC) to make energy. TCA cycle intermediates can also exit mitochondria to synthesize fatty acids, amino acids, hemes, and purines (2). As can be seen from TCA cycle, metabolic pathways contribute to production of energy and macromolecules in the cell. Analyzing how cells utilize metabolic pathways provide information on how cells use available nutrients to survive.

Cellular utilization of metabolic pathways can be quantified by tracing stable isotope-labeled nutrients. Isotopes of an element carry the same number of protons but different numbers of neutrons, resulting in having different masses. One of the commonly utilized isotopes to label nutrients are carbon atoms (3, 4). Carbon atoms whose masses are 13 atomic mass units (amu) which are heavier than the ones found in nature (12 amu) can be used to label nutrients such as glucose or glutamine (5-7). By culturing cells with medium containing a stable isotope-labeled nutrient and analyzing the distribution of the isotopes in metabolites, researchers can get information on how cells utilize the nutrient for fuel.

1.1.2 Correcting for Naturally Occurring Isotope Abundances

i. Direct inspection of mass isotopologue distributions for flux analysis

Researchers can perform a computational quantitation using metabolic flux analysis (MFA) to analyze cellular utilization of metabolic pathways. MFA requires researchers to computationally construct a metabolic reaction network. Information on tracers used in experiments and relative abundances of metabolites whose masses are different due to isotope incorporation are used as inputs for MFA. Fluxes of metabolic pathways in a network are computed by minimizing the sum of squared differences between the measured and model-estimated values. Due to complexity in constructing a network, MFA can be time- and data-consuming to researchers not familiar with computational analysis (3, 8).

A more common approach to study cellular utilization of metabolic pathways is direct interpretation of mass isotopologue distributions (MIDs). When cells are cultured with stable isotope-labeled nutrients, metabolites that only differ in isotope compositions called mass isotopologues are present due to incorporation of isotopes. In a ^{13}C tracing analysis, a metabolite with *n* carbon atoms has isotopologues from M0 (all carbons are ^{12}C) to M*n* (all *n* carbons are ^{13}C). Relative abundances of isotopologues from M0 to M*n* of a metabolite is called the MID. MIDs of metabolites can be directly used to infer how cells use specific metabolic pathways when isotope incorporation in metabolites is constant over time (3). For instance, relative abundances of citrate with two ^{13}C atoms (M2 Cit) are frequently used as a measure of pyruvate dehydrogenase (PDH) activities in the cells cultured with glucose carrying six ^{13}C atoms ([U-$^{13}C_6$]glucose). Pyruvate with three labeled

carbon atoms produced from [U-$^{13}C_6$]glucose is converted to acetyl coenzyme A (AcCoA) with two labeled carbon atoms (M2 AcCoA) by the activity of PDH. If M2 AcCoA condenses with oxaloacetate with unlabeled carbon atoms, M2 citrate is produced, which is commonly used as a readout of PDH activities (5, 6, 9). As direct interpretation of MIDs is more intuitive and requires less time, it is widely used to study cellular utilization of metabolic pathways in various disease models such as cancer cell proliferation (10) and metastasis (7, 11), diabetes (12, 13), and neural tube defects (14).

ii. Naturally occurring isotope abundances confounding MID interpretation

However, mass shifts of metabolites can also arise from incorporation of naturally occurring isotopes in stable isotope tracing analyses. Specifically, presence of naturally occurring carbon atoms whose masses are 13 amu are 1.11%. If researchers take MIDs to draw conclusions on cellular utilization of metabolic pathways without correcting for the 1.11%, results can be misleading in ^{13}C tracing analyses (15). In addition to carbon atoms, naturally occurring isotopes exist in other chemical atoms such as ^{2}H (relative abundances of 0.0115%), ^{15}N (relative abundances of 0.368%), ^{17}O (relative abundances of 0.038%), and ^{18}O (relative abundances of 0.205%) (16). Thus MIDs should be corrected for abundances of naturally occurring isotopes to be accurately used as readouts of metabolic pathway utilization.

1.2 Proteomics and Metabolism

Cellular metabolism can be analyzed by inspecting abundances of substrates and

products of metabolic pathways by metabolomics. However, abundances of metabolites do not always correspond to cellular utilization of metabolic pathways. Specifically, accumulation of metabolites does not necessarily imply increased cellular utilization of metabolic pathways producing the metabolites as pathways consuming the metabolites may be inhibited (17). Quantifying activities and abundances of enzymes catalyzing metabolic reactions can aid in better understanding of how cells utilize metabolic pathways. Analyzing cellular proteomes enables researchers to inspect protein abundances of multiple metabolic enzymes simultaneously. Researchers can also examine correlations of protein abundances between metabolic enzymes and how they are correlated with proteins other than metabolic enzymes such as those involved in cell cycle or deoxyribonucleic acid (DNA) repair. Thus proteomics can be a useful tool to study correlations and regulations of metabolic enzymes.

1.2.1 Impact of Vitamin Intakes on Tissue Proteomes

i. Vitamins as cofactors in metabolic reactions

Vitamins are cofactors in majority of metabolic reactions. Vitamin B_3 works as a cofactor in the pentose phosphate pathway, pyruvate oxidation catalyzed by PDH, *de novo* serine synthesis, and TCA cycle (18). Vitamin B_2 is needed for activities of ETC Complex I and II which take reducing equivalents generated from TCA cycle to produce energy (19). In addition to the two vitamins mentioned above, other vitamins participate as cofactors in various metabolic pathways (18, 20). As metabolic enzymes need cofactors to be activated, availability of vitamins in our body may impact abundances of metabolic enzymes, and consequently, cellular proteomes.

ii. Vitamins as regulators of nuclear transcription factors

Vitamins also affect proteomes by regulating transcription of genes which consequently impacts protein abundances of the genes. Vitamin A can activate the retinoic acid receptor (RAR)-retinoid X receptor (RXR) heterodimers which regulate transcription by interacting with retinoic acid-responsive elements in the promoters of genes (21). Vitamin D can bind to vitamin D receptors which regulate transcription together with RXR by interacting with vitamin D receptor elements in the promoter regions of genes (22). Considering their roles in regulation of gene transcription, vitamins can affect abundances of proteins which are the products of transcription and translation and ultimately carry physiological functions in our body.

1.3 Breast Cancer and Metastasis

1.3.1 Breast Cancer

Breast cancer is estimated to have the highest number of new cases among all cancers in 2023 in the United States (US). Approximately 13% of women is estimated to be diagnosed with breast cancer during lifetime (23) and death rate is reported to be 19.6% per 100,000 women in the US (24). Breast tumors are heterogeneous and have been categorized based on immunohistochemical measurements of the estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) in clinical settings (25). In 2000, Perou and colleagues classified breast tumors based on mRNA expressions and identified five molecular subtypes; luminal A and B, HER2, basal, and normal-like subtypes (26).

Among breast cancer-subtypes, patients with triple-negative (TN) breast tumors lack expressions of the ER, PR, and HER2 and account for 15% in the US breast cancer population. Though 70-80% of TN breast tumors are reported to be classified as basal-like molecular subtype, TN breast tumors that are not basal-like subtype exist, indicating the two terms are not synonymous (27, 28). Development of targeted therapies such as ER-targeting endocrine agents and HER2-targeting antibodies improved prognoses and outcomes of patients with hormone receptors-positive breast cancer patients (patients expressing ER, PR, or HER2). However, no targeted therapies are currently available for patients with TN breast tumors which may be associated with higher rates of mortality and earlier relapse compared with patients with other breast cancer-subtypes. Due to lack of targeted therapies, only options available for patients with TN breast tumors are surgery, radiation, and chemotherapy, all of which have severe side effects (25).

1.3.2 Metastasis of Breast Cancer

i. Metastasis as the cause of majority of deaths in breast cancer patients

Metastasis, a spread of tumors into distant organs, occurs through a series of sequential steps. Tumors cells invade and migrate through extracellular matrix (ECM) in primary tumors, move into vasculature or lymphatic system, survive during circulation, exit the circulation, and enter/colonize secondary organs (29). Upon metastasis, removal of tumors with surgery becomes difficult and tumor cells can resist to targeted therapies. Due to these challenges in treatment of metastatic tumors, 90% of cancer deaths are reported to be associated with metastasis (30).

Majority of breast cancer-associated deaths are also due to metastasis rather than

localized tumors themselves (31). According to the US cancer statistics from 2012 to 2018, five-year survival rates of breast cancer patients with metastasis were only one-third of the survival rates of patients with primary breast tumors (23). Gogate and colleagues estimated the number of breast cancer patients with metastasis to increase by 50% from 2015 to 2030 (32), indicating survival of more breast cancer patients may be at risk. Though the development of targeted therapies, early screening, and identification of breast cancer molecular subtypes contributed to increased survival of breast cancer patients with primary tumors, further research is needed to prevent or treat breast cancer metastasis.

ii. Distinct patterns of metastasis among breast cancer-subtypes

Timing and sites of metastasis are reported to differ according to the breast cancer subtypes. Kennecke and colleagues reported patients with TN breast tumors have higher probabilities of developing metastasis within 15 years since diagnosis. Duration of survival after diagnosis of metastasis is reported to be shorter in patients with TN breast tumors compared with those carrying the other breast cancer-subtypes (33).

For patients with hormone receptors-positive breast tumors, bone is reported to be the common site of metastasis. For patients with TN breast tumors, bone and lung are reported to be the two most common sites of metastasis. According to a multivariate analysis comparing preferential metastatic sites across patients with different breast cancer-subtypes, patients with TN breast tumors have higher rates of distant nodal metastasis and lower rates of bone metastasis compared with those with the other breast cancer subtypes. Interestingly, patients with HER2-positive but

ER/PR-negative breast tumors have higher rates of brain metastasis compared with those with the other breast cancer-subtypes (33).

1.3.3 Anchorage-Independent Survival of Tumor Cells During Metastasis

Tumor cells need to overcome multiple hurdles to spread to distant organs and failure to overcome any of the hurdles can halt metastatic cascade. Only those tumor cells that acquired mechanisms to overcome all the hurdles can land at secondary sites and interact with host microenvironments (34). The first barrier tumor cells need to overcome is to maintain survival upon detachment from ECM at primary tumors. Non-malignant or tumor cells without metastatic potential undergo cell death process called anoikis upon detachment from ECM. However, certain tumor cells develop mechanisms to resist anoikis by regulating the tumor necrosis factor receptors superfamily called death receptors or mitochondrial caspases (35), enabling them to survive under anchorage-independent conditions. Cellular mechanism supporting anchorage-independent survival affects viability of tumor cells going through other steps of the metastatic cascade as tumor cells in circulation or at secondary sites (before colonization) are also under anchorage-independent condition (36). Thus, maintaining cellular survival in anchorage-independent condition serves a gatekeeper to successful metastasis.

1.4 Metabolic Changes Promoting Anchorage-Independent Survival of Breast Tumor Cells

Tumor cells are reported to change phenotypes to maintain viability during metastatic cascade. For instance, tumor cells change from epithelial to mesenchymal morphology to detach from ECM and move to circulatory system. Throughout the metastatic cascade, anchorage-independent tumor cells are exposed to different types and amounts of metabolites upon losing contacts with ECM (37). Thus, tumor cells alter utilization of nutrients and metabolic pathways to meet metabolic requirements during metastasis.

A previous study reported decreased utilization of glucose in glycolysis, pentose phosphate pathway, and TCA cycle upon detachment from ECM in a mammary epithelial cell line (38). In an anchorage-independent mammary epithelial cell line, HER2-overexpression increased glucose utilization in glycolysis, pentose phosphate pathway, PDH activities, and TCA cycle , indicating metabolic changes upon ECM detachment may be different according to distinct molecular signatures in mammary epithelial cells (39).

In a TN breast cancer cell line, anchorage-independent cells are reported to depend more on glutamine than glucose to maintain viability (40). Fendt and colleagues reported increased dependence on pyruvate upon detachment from ECM in a breast cancer cell line (41). These previous findings suggest breast cancer cells may rewire metabolism to survive in anchorage-independent condition.

1.5 Proteomic Changes Promoting Anchorage-Independent Survival of Breast Tumor Cells

i. Proteomes reflecting survival and metastasis occurrence of breast cancer patients

Proteomic signatures are reported to stratify breast tumors into survival outcomes according to a quantitative proteomics on breast tumors of 178 patients (42), indicating analyzing proteomes may provide information directly related to patient survival. As proteins carry physiological functions in our body, proteomes may contain mechanisms crucial to survival of breast cancer patients. As metastasis is the cause of majority of cancer-associated deaths (30), analyzing proteomes of patients with metastasis may enable researchers to identify molecular mechanisms associated with patient survival. A previous study on plasma proteomes of 48 TN breast cancer patients reported cell adhesion and migration as pathways proteins whose abundances were differentially expressed in patients with metastasis compared with those without metastasis (43). The primary purpose of tumor cell adhesion and migration is to support anchorage-independent survival in metastasis (44). Thus analyzing proteomic changes associated with anchorage-independent survival of breast tumor cells can provide mechanisms critical to metastasis occurrence.

ii. Reported proteomic changes upon detachment from ECM in breast cancer cells

A previous study on the proteome of a hormone receptor-positive breast cancer cell line reported increased abundances of proteins regulating cell-cell or cell-ECM interactions upon detachment from ECM (45). In two hormone receptor-positive breast cancer cell lines, detachment from ECM increased abundances of proteins involved in degradations of fatty acids and ketone bodies, an oxidative branch of pentose phosphate pathway, and pyruvate carboxylase-catalyzed glucose anaplerosis (46). In addition to the abundance changes in the metabolic enzymes,

decreased abundances of proteins in phosphoinositide 3-kinases (PI3K)-AKT serine/threonine kinases (AKT) signaling and an integrin signaling pathways, all of which regulate cell cycle, are reported in the two anchorage-independent hormone receptor-positive breast cancer cell lines (47).

iii. Associations of metabolomes and proteomes in breast tumors

A proteomic analysis on breast tumors of 45 patients reported proteomic signatures are associated with glycolytic properties of breast tumors (48), suggesting a close association between metabolomes and proteomes in breast tumors. As metabolic rewiring and proteome alterations are both critical to development of metastasis, analyzing metabolic and proteomic changes simultaneously may enable identification of more efficient molecular targets that affect both.

REFERENCES

1. DeBerardinis RJ, Thompson CB. Cellular metabolism and disease: what do metabolic outliers teach us? Cell. 2012;148(6):1132-44.

2. Owen OE, Kalhan SC, Hanson RW. The key role of anaplerosis and cataplerosis for citric acid cycle function. J Biol Chem. 2002;277(34):30409-12.

3. Buescher JM, Antoniewicz MR, Boros LG, Burgess SC, Brunengraber H, Clish CB, et al. A roadmap for interpreting (13)C metabolite labeling patterns from cells. Curr Opin Biotechnol. 2015;34:189-201.

4. Zamboni N, Fendt SM, Rühl M, Sauer U. (13)C-based metabolic flux analysis. Nat Protoc. 2009;4(6):878-92.

5. Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, et al. Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature. 2011;481(7381):380-4.

6. Pachnis P, Wu Z, Faubert B, Tasdogan A, Gu W, Shelton S, et al. In vivo isotope tracing reveals a requirement for the electron transport chain in glucose and glutamine metabolism by tumors. Sci Adv. 2022;8(35):eabn9550.

7. Labuschagne CF, Cheung EC, Blagih J, Domart MC, Vousden KH. Cell Clustering Promotes a Metabolic Switch that Supports Metastatic Colonization. Cell Metab. 2019;30(4):720-34.e5.

8. Antoniewicz MR. A guide to (13)C metabolic flux analysis for the cancer biologist. Exp Mol Med. 2018;50(4):1-13.

9. Vacanti NM, Divakaruni AS, Green CR, Parker SJ, Henry RR, Ciaraldi TP, et al. Regulation of substrate utilization by the mitochondrial pyruvate carrier. Mol Cell. 2014;56(3):425-35.

10. Faubert B, Li KY, Cai L, Hensley CT, Kim J, Zacharias LG, et al. Lactate Metabolism in Human Lung Tumors. Cell. 2017;171(2):358-71.e9.

11. Bartman CR, Weilandt DR, Shen Y, Lee WD, Han Y, TeSlaa T, et al. Slow TCA flux and ATP production in primary solid tumours but not metastases. Nature. 2023;614(7947):349-57.

12. Krycer JR, Yugi K, Hirayama A, Fazakerley DJ, Quek LE, Scalzo R, et al. Dynamic Metabolomics Reveals that Insulin Primes the Adipocyte for Glucose Metabolism. Cell Rep. 2017;21(12):3536-47.

13. Wortham M, Benthuysen JR, Wallace M, Savas JN, Mulas F, Divakaruni AS, et al. Integrated In Vivo Quantitative Proteomics and Nutrient Tracing Reveals Age-Related Metabolic Rewiring of Pancreatic β Cell Function. Cell Rep. 2018;25(10):2904-18.e8.

14. Leung KY, Pai YJ, Chen Q, Santos C, Calvani E, Sudiwala S, et al. Partitioning of One-Carbon Units in Folate and Methionine Metabolism Is Essential for Neural Tube Closure. Cell Rep. 2017;21(7):1795-808.

15. Fernandez CA, Des Rosiers C, Previs SF, David F, Brunengraber H. Correction of 13C mass isotopomer distributions for natural stable isotope abundance. J Mass Spectrom. 1996;31(3):255-62.

16. van Winden WA, Wittmann C, Heinzle E, Heijnen JJ. Correcting mass isotopomer distributions for naturally occurring isotopes. Biotechnol Bioeng. 2002;80(4):477-9.

17. Jang C, Chen L, Rabinowitz JD. Metabolomics and Isotope Tracing. Cell. 2018;173(4):822-37.

18. Peterson CT, Rodionov DA, Osterman AL, Peterson SN. B Vitamins and Their Role in Immune Regulation and Cancer. Nutrients. 2020;12(11).

19. Powers HJ, Corfe BM, Nakano E. Riboflavin in development and cell fate. Subcell Biochem. 2012;56:229-45.

20. Rebouche CJ. Ascorbic acid and carnitine biosynthesis. Am J Clin Nutr. 1991;54(6 Suppl):1147s-52s.

21. Minucci S, Leid M, Toyama R, Saint-Jeannet JP, Peterson VJ, Horn V, et al. Retinoid X receptor (RXR) within the RXR-retinoic acid receptor heterodimer binds its ligand and enhances retinoid-dependent gene expression. Mol Cell Biol. 1997;17(2):644-55.

22. Carlberg C, Campbell MJ. Vitamin D receptor signaling mechanisms: integrated actions of a well-defined transcription factor. Steroids. 2013;78(2):127-36.

23. Surveillance E, and End Results (SEER) Program (www.seer.cancer.gov). SEER*Stat Database: Incidence - SEER Research Data, 8 Registries, Nov 2021 Sub (1975-2020) - Linked To County Attributes - Time Dependent (1990-2020) Income/Rurality, 1969-2020 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, released April 2023, based on the November 2022 submission.

24. Surveillance E, and End Results (SEER) Program (www.seer.cancer.gov).

SEER*Stat Database: Mortality - All COD, Aggregated With State, Total U.S. (1969-2020) <Katrina/Rita Population Adjustment>, National Cancer Institute, DCCPS, Surveillance Research Program, released September 2022. Underlying mortality data provided by NCHS (www.cdc.gov/nchs).

25. Waks AG, Winer EP. Breast Cancer Treatment: A Review. Jama. 2019;321(3):288-300.

26. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.

27. Bertucci F, Finetti P, Cervera N, Esterni B, Hermitte F, Viens P, et al. How basal are triple-negative breast cancers? Int J Cancer. 2008;123(1):236-40.

28. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. N Engl J Med. 2010;363(20):1938-48.

29. Klein CA. Cancer. The metastasis cascade. Science. 2008;321(5897):1785-7.

30. Chaffer CL, Weinberg RA. A perspective on cancer cell metastasis. Science. 2011;331(6024):1559-64.

31. Weigelt B, Peterse JL, van 't Veer LJ. Breast cancer metastasis: markers and models. Nat Rev Cancer. 2005;5(8):591-602.

32. Gogate A, Wheeler SB, Reeder-Hayes KE, Ekwueme DU, Fairley TL, Drier S, et al. Projecting the Prevalence and Costs of Metastatic Breast Cancer From 2015 through 2030. JNCI Cancer Spectr. 2021;5(4).

33. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, et al. Metastatic behavior of breast cancer subtypes. J Clin Oncol. 2010;28(20):3271-7.

34. Fidler IJ. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nat Rev Cancer. 2003;3(6):453-8.

35. Simpson CD, Anyiwe K, Schimmer AD. Anoikis resistance and tumor metastasis. Cancer Lett. 2008;272(2):177-85.

36. Mehlen P, Puisieux A. Metastasis: a question of life or death. Nat Rev Cancer. 2006;6(6):449-58.

37. Fendt SM. Metabolic vulnerabilities of metastasizing cancer cells. BMC Biol. 2019;17(1):54.

38. Grassian AR, Coloff JL, Brugge JS. Extracellular matrix regulation of metabolism

and implications for tumorigenesis. Cold Spring Harb Symp Quant Biol. 2011;76:313-24.

39. Grassian AR, Metallo CM, Coloff JL, Stephanopoulos G, Brugge JS. Erk regulation of pyruvate dehydrogenase flux through PDK4 modulates cell proliferation. Genes Dev. 2011;25(16):1716-33.

40. Endo H, Owada S, Inagaki Y, Shida Y, Tatemichi M. Metabolic reprogramming sustains cancer cell survival following extracellular matrix detachment. Redox Biol. 2020;36:101643.

41. Elia I, Rossi M, Stegen S, Broekaert D, Doglioni G, van Gorsel M, et al. Breast cancer cells rely on environmental pyruvate to shape the metastatic niche. Nature. 2019;568(7750):117-21.

42. Asleh K, Negri GL, Spencer Miko SE, Colborne S, Hughes CS, Wang XQ, et al. Proteomic analysis of archival breast cancer clinical specimens identifies biological subtypes with distinct survival outcomes. Nat Commun. 2022;13(1):896.

43. Katayama H, Tsou P, Kobayashi M, Capello M, Wang H, Esteva F, et al. A plasma protein derived TGFβ signature is a prognostic indicator in triple negative breast cancer. NPJ Precis Oncol. 2019;3:10.

44. Zhong X, Rescorla FJ. Cell surface adhesion molecules and adhesion-initiated signaling: understanding of anoikis resistance mechanisms and therapeutic opportunities. Cell Signal. 2012;24(2):393-401.

45. Morrison BJ, Hastie ML, Grewal YS, Bruce ZC, Schmidt C, Reynolds BA, et al. Proteomic comparison of mcf-7 tumoursphere and monolayer cultures. PLoS One. 2012;7(12):e52692.

46. Lamb R, Harrison H, Hulit J, Smith DL, Lisanti MP, Sotgia F. Mitochondria as new therapeutic targets for eradicating cancer stem cells: Quantitative proteomics and functional validation via MCT1/2 inhibition. Oncotarget. 2014;5(22):11029-37.

47. Tasdemir N, Ding K, Savariau L, Levine KM, Du T, Elangovan A, et al. Proteomic and transcriptomic profiling identifies mediators of anchorage-independent growth and roles of inhibitor of differentiation proteins in invasive lobular carcinoma. Sci Rep. 2020;10(1):11487.

48. Johansson HJ, Socciarelli F, Vacanti NM, Haugen MH, Zhu Y, Siavelis I, et al. Breast cancer quantitative proteome and proteogenomic landscape. Nat Commun. 2019;10(1):1600.

Chapter 2

Systemic Vitamin Intake Impacting Tissue Proteomes*

* Jeong H, Vacanti NM. Systemic vitamin intake impacting tissue proteomes.

Nutrition & Metabolism. 2020 Dec;17(1):1-6.

Abstract

The kinetics and localization of the reactions of metabolism are coordinated by the enzymes that catalyze them. These enzymes are controlled via a myriad of mechanisms including inhibition/activation by metabolites, compartmentalization, thermodynamics, and nutrient sensing-based transcriptional or post-translational regulation; all of which are influenced as a network by the activities of metabolic enzymes and have downstream potential to exert direct or indirect control over protein abundances. Considering many of these enzymes are active only when one or more vitamin cofactors are present; the availability of vitamin cofactors likely yields a systems-influence over tissue proteomes. Furthermore, vitamins may influence protein abundances as nuclear receptor agonists, antioxidants, substrates for post-translational modifications, molecular signal transducers, and regulators of electrolyte homeostasis. Herein, studies of vitamin intake are explored for their contribution to unraveling vitamin influence over protein expression. As a body of work, these studies establish vitamin intake as a regulator of protein abundance; with the most powerful demonstrations reporting regulation of proteins directly related to the vitamin of interest. However, as a whole, the field has not kept pace with advances in proteomic platforms and analytical methodologies, and has not moved to validate mechanisms of regulation or potential for clinical application.

Introduction

Regulatory Mechanisms

Cellular metabolism is a system of chemical reactions in which cells harness the

energy stored in the chemical bonds of substrate molecules to perform their biological functions, maintain homeostasis, or to synthesize building blocks for structural maintenance or cellular division. The kinetics of these reactions are dependent on the activity of the proteins which catalyze them; thus proteins are key modulators of metabolism.

Metabolic activity also exerts network control over itself by a diverse array of mechanisms which finely tune protein expression responses via nutrient sensing machineries (1). Products or intermediates of a metabolic pathway can inhibit or activate metabolic enzymes; e.g. malate inhibits the succinate dehydrogenase complex (2) and fructose-2,6-bisphosphate activates phosphofructokinase (3). The oxidative status of a cell can drive the directionality of redox reactions and impact abundances of redox reaction-catalyzing proteins; e.g. the KEAP1/NRF2 network responds to oxidative stress by upregulating expression of antioxidant-functioning proteins (4). Splice-variant or isozyme expression can impact relative pathway utilization at metabolic network nodes; e.g. splice variants and isozymes of pyruvate and lactate dehydrogenase respectively impact the bridge between glycolysis and the TCA cycle (5, 6). Additionally, local metabolite concentrations and thermodynamics can dictate the directionality of reactions catalyzed by compartment-specific isozymes; e.g. reductive activity of isocitrate dehydrogenase can be confined to the cytosol-specific isozyme (7). The impacts of the above-mentioned regulations are closely monitored by nutrient sensing proteins which initiate molecular events altering protein activation and expression; e.g. serine/threonine kinase 11, AMP-activated protein kinase, mammalian target of rapamycin 1, and sterol regulatory element-binding protein 1 are

part of overlapping protein networks that orchestrate protein-expression and post-translational modification responses to nutrient availability (8, 9). Considering that many metabolic enzymes do not function in isolation and, as detailed in the sections that follow, require vitamin cofactors to stabilize intermediates, donate/accept electrons, shuttle substrates, and hold reactants in close proximity; vitamin status is a critical consideration when examining protein-mediated regulation of metabolism and the impacts of metabolism on protein expression.

In addition to their potential regulatory roles as cofactors, vitamins orchestrate other direct or indirect mechanisms influencing protein abundance. Retinoic acid (vitamin A) interacts with nuclear receptors impacting gene transcription (10), ascorbic acid (vitamin C) impacts oxidative status and associated protein networks (11) and is reported to exhibit epigenetic regulation over protein expression (12), vitamin D regulates calcium signaling machinery, activates nuclear receptors, and exerts hormonal regulation over protein expression (13, 14), and niacin (vitamin B_3) and biotin (vitamin B_7) can be incorporated as post-translational modifications impacting protein function (15, 16).

Herein, studies on systemic intake (dietary, injection, oral gavage) of vitamins and their impacts on tissue proteomes are examined, and their contributions to unraveling vitamin-based regulation of protein expression and tissue function are explored. The current work is intended to provide background information to understand each vitamin's (Figures 2.1 and 2.2) molecular functions and highlight its role as a cofactor or substrate in the reactions of central metabolism (Figure 2.3, Appendix Tables 2.1-2.13). Finally, this work is intended as a resource for identifying regulation of proteins

related to vitamin metabolism in published works. The public domain of proteomic data sets is ever expanding, but is rarely searched for effects related to vitamin metabolism. To that end, all proteins are specified by their HUGO Gene Nomenclature Committee (HGNC) gene symbol, or the HGNC gene symbol of the human ortholog when identified in another species, and proteins requiring a vitamin as a cofactor or substrate are tabulated (Appendix Tables 2.1-2.13).

Proteomics Platforms

Proteomics platforms of the discussed studies are provided to place them on a technological timeline. Platforms are described with the terms: orbitrap, QTOF (quadrupole time-of-flight), triple-TOF (triple – time of flight), QQQ (triple-quadrupole), 2DGE-MS (two-dimensional gel electrophoresis – mass spectrometry), and 2DGE. In brief, orbitrap platforms are the workhorses of modern proteomics because their high achievable mass resolutions combined with high sensitivity are best suited for maximizing the number of proteins identified in a complex sample (17, 18); though QTOF and triple-TOF instruments, capable of maintaining mass resolution at higher scan speeds (19), hold a substantial influence in this arena. Within the categories of orbitrap, QTOF, and triple-TOF, there are major technological advances not discussed here. QQQ platforms are best suited for quantifying a pre-determined list of proteins. Lower scan speeds and mass resolution render them less capable than orbitrap, QTOF, or triple-TOF systems for non-targeted applications (17). Advances in nano-flow liquid chromatography coupled directly to mass spectrometry have improved proteomic depth by orders of magnitude over that achievable by 2DGE-MS,

where the upstream selection of protein spots predates the modern definition of non-targeted proteomics. Similarly, identifying differentially intense protein spots using 2DGE alone is considered an important milestone in the development of proteomics; but is rarely discussed outside the topic of the field's history.

Vitamin A

Vitamin A exists in alcohol, aldehyde, acid, and ester forms known as retinol, retinal, retinoic acid, and retinyl esters respectively (Figure 2.1) (20). Several carotenoids are precursors to vitamin A including α- and β-carotene (21). β-carotene is converted to two molecules of retinal by beta carotene oxygenases (*BCO1* or *BCO2*) (22). Retinal is an important component of rhodopsin (*RHO*), a protein in rod cells responsible for detecting low levels of light (23). Thus night blindness is telltale characteristic of vitamin A deficiency (24). Retinoic acid serves as a signaling molecule, acting through nuclear retinoic acid (*RARA*, *RARB*, *RARG*) and retinoid X (*RXRA*, *RXRB*, *RXRG*) receptors which regulate growth and differentiation (25, 26). Cellular and organismal trafficking of vitamin A is dependent on retinol/retinoic acid binding proteins (RBP family, *CRABP1*, *CRABP2*) and retinol esterification via lecithin retinol acyltransferase (*LRAT*) (27). Retinal is oxidized to retinol via aldehyde dehydrogenases (ALDH family) and retinol is oxidized to retinoic acid by retinol dehydrogenases (*RDH* and *DHRS* families) (28). In addition to inducing night blindness, vitamin A deficiency adversely impacts cellular growth, bone development, and antibody-based immune responses (29).

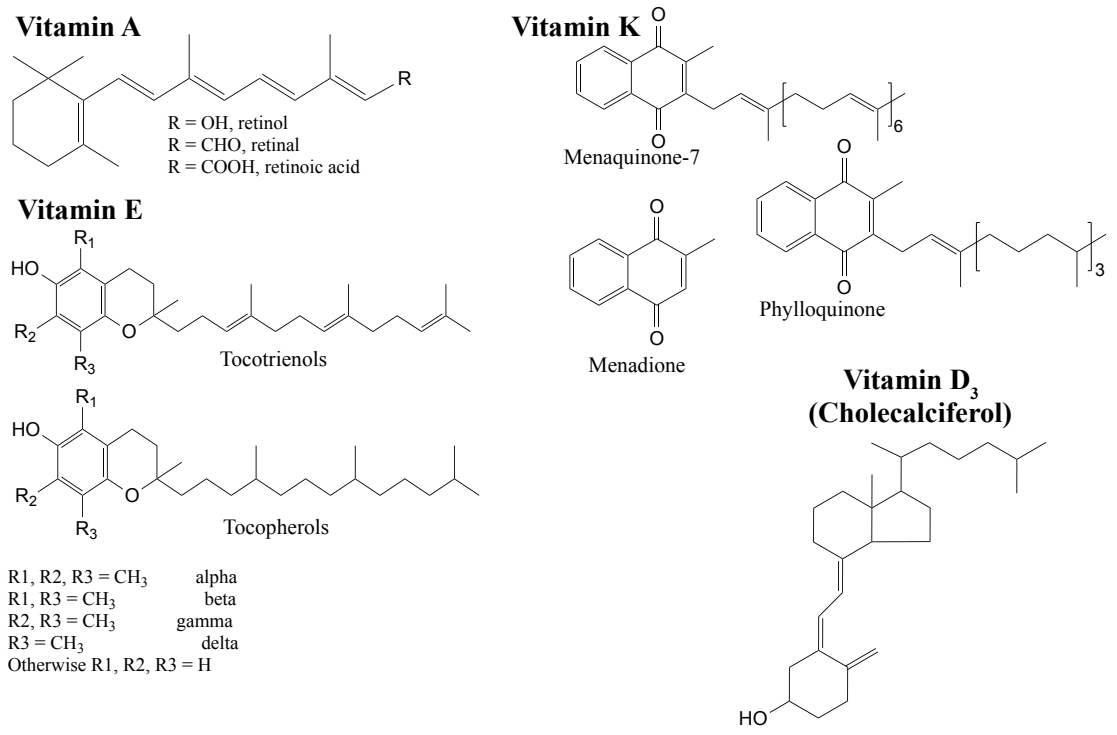


Figure 2.1. Fat-Soluble Vitamin Structures

In an orbitrap-based study of mouse embryo heads, toxic levels of prenatal retinoic acid exposure intended to model an established risk factor for craniofacial birth defects are reported to induce abundance alterations in proteins associated with craniofacial development and neural crest processes (30). In a parallel triple-TOF-based study of gerbil plasma and 2DGE-MS-based study of gerbil liver and white adipose tissue, a few dozen protein abundances linked to a handful of biological processes are reported to respond to dietary retinol, β-carotene, lutein, or lycopene; though process or pathway enrichment analyses are not reported. As the authors discuss, plasma was not depleted of common highly abundant proteins upstream of analysis by mass spectrometry which are known to adversely impact data quality (31). In an orbitrap-based study of plasma from Nepalese children, dozens of proteins are associated with circulating carotenoid

abundances; potentiating development of low-cost antibody-based tests for carotenoid deficiencies (32). A pair of 2DGE-MS-based studies link tissue function to protein abundance responses to vitamin A status in mice brains (33) and bovine muscle (34).

Vitamin B_1

Thiamine (vitamin B_1) is composed of linked pyrimidine and thiazole rings decorated with methyl, amine, and alkyl-hydroxyl functional groups (Figure 2.2) (35). Thiamine is transported through the plasma membrane via thiamine transporters (*SLC19A2* and *SLC19A3*) (36) and then twice phosphorylated on the alkyl-hydroxyl functional group by thiamine pyrophosphokinase (*TPK1*), rendering it active as thiamine diphosphate (TDP) (35). TDP is a cofactor for enzymes catalyzing the oxidative decarboxylation of ketoacids including the pyruvate dehydrogenase complex (*PDHA*, *PDHB*, *PDHX*, *DLAT*, *DLD*), the oxoglutarate dehydrogenase complex (*OGDH*, *DLST*, *DLD*), and the branched chain keto acid dehydrogenase complex (*BCKDHA*, *DBT*, *DLD*) (37). It is also a cofactor for transketolase (*TKT*) in the non-oxidative branch of the pentose phosphate pathway (38). Independent from its role as a cofactor, thiamine is believed to regulate ion transport activity in the nervous system (39).

Vitamin B_1 deficiency is marked by a broad range of neurological, respiratory, and cardiovascular pathophysiologies and is termed beriberi. Symptoms of beriberi are difficult to directly link to the molecular functions of vitamin B_1 (40).

In a 2DGE-MS-based study of type 2 diabetic and healthy control subjects, authors report treatment with thiamine reduces albumin (*ALB*) abundance in urine; indicating

the vitamin serves a protective role of kidney function (41). In a QTOF-based study of rat thalami under thiamine deficiency, glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) is the most up-regulated protein (50 fold) while regulated proteins are most enriched in the synaptic vesicle cycle pathway (according to the KEGG database). Proteomic changes are accompanied by diminished performances on cognitive tests (42).

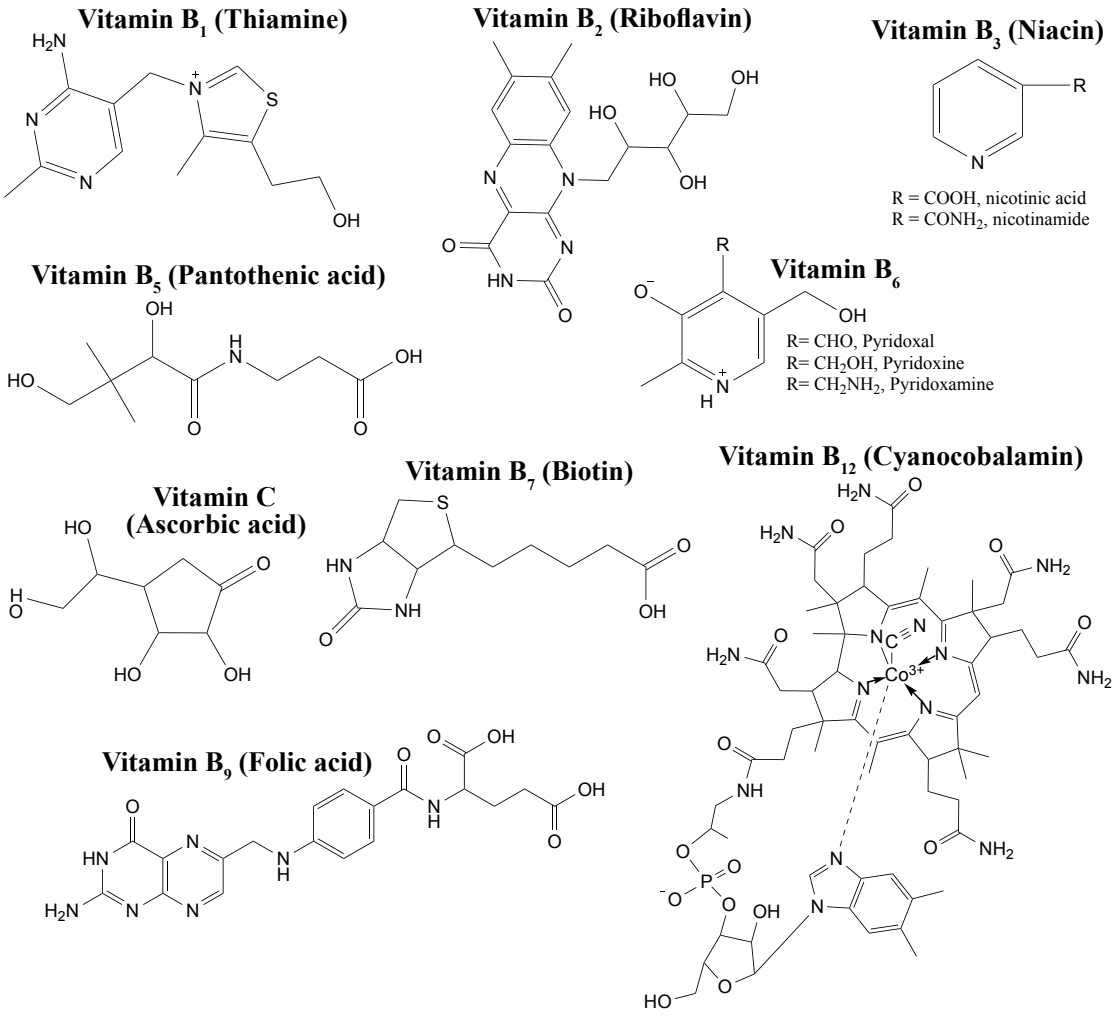


Figure 2.2. Water-Soluble Vitamin Structures

Vitamin B_2

Riboflavin (vitamin B_2) is composed of an isoalloxazine ring and a bound ribitol (Figure 2.2) (43). It is activated by riboflavin kinase (*RFK*), forming flavin mononucleotide (FMN); and by flavin adenine dinucleotide synthase 1 (*FLAD1*), forming flavin adenine dinucleotide (FAD) (44). Bound FMN or FAD serves as an electron carrier for redox-reaction-catalyzing proteins (flavoproteins) including the succinate dehydrogenase complex (*SDHA*, *SDHB*, *SDHC*, *SDHD*), the pyruvate dehydrogenase complex (*PDHA*, *PDHB*, *PDHX*, *DLAT*, *DLD*), acyl-CoA dehydrogenases (*ACAD*s), and methylene tetrahydrofolate reductase (*MTHFR*) (45).

Riboflavin deficiency in humans predominantly occurs in combination with that of other nutrients. However animal studies link it to impaired fetal and intestinal development (46, 47), iron absorption (48), and lipid metabolism (49, 50).

In a QTOF-based study of duck livers, riboflavin deficiency is accompanied by a reduced abundance of small-chain-specific acyl-coenzyme A dehydrogenases (*ACADs*), for which riboflavin serves as a cofactor, and concordant elevation of hepatic small chain fatty-acid lipid content. Dramatic decreases in protein abundance are reported for *INPP1* (involved in inositol signaling), *THRSP* (purported regulator of lipid metabolism), *BDH2* (a regulator of lipid metabolism), *FXN* (involved in mitochondrial iron-sulfur complex assembly), and *NDUFS1* (a subunit of electron transport chain complex I) (51). In a QTOF-based study of maternal riboflavin deficiency, reductions in fetal duck hepatic TCA cycle, beta-oxidation, and electron transport chain proteins are reported, with *IDH3A* being the lone member of these pathways whose abundance increases (52).

Vitamin B_3

Niacin (vitamin B_3) is inclusive of nicotinic acid and nicotinamide (Figure 2.2) which are converted to their mononucleotide forms by nicotinate phosphoribosyltransferase (*NAPRT*) and nicotinamide phosphoribosyltransferase (*NAMPT*) respectively (53). Both forms of the mononucleotide are subsequently converted to their adenosine dinucleotide forms by nicotinamide/nicotinic acid mononucleotide adenylyltransferases (*NMNAT1*, *NMNAT2*, *NMNAT3*). Nicotinamide adenine dinucleotide (NAD) is a cofactor form of the vitamin whereas nicotinic acid dinucleotide is subsequently converted to NAD by NAD synthase (*NADSYN1*) (54). NAD is reduced to NADH by oxidative reactions of glycolysis, the TCA cycle, and β-oxidation; and subsequently serves as a redox equivalent carrier to the electron transport chain (55) and to regenerate reduced ascorbic acid (vitamin C) (56), glutathione (57), and thioredoxin (58). NAD can also be phosphorylated by NAD kinases (*NADK*, *NADK2*) to form a distinct redox shuttling cofactor, NADP (59). NADP is reduced by reactions in the oxidative pentose phosphate pathway (*G6PD*, *PGD*) and other enzymes (e.g. *ME1*, *ME3*, *IDH1*, *IDH2*) to NADPH. NADPH provides reducing equivalents for biosynthetic reactions in fatty acid, cholesterol, and deoxyribonucleotide synthesis (60). Outside its role as a reducing equivalent shuttle, NAD provides adenine dinucleotide phosphate (ADP) ribose for synthesis of the second messenger, cyclic adenosine monophosphate (cAMP), via the activity of adenylate cyclases (*ADCY* family) (61). NAD also provides ADP-ribose and poly-ADP-ribose for post translational modifications of proteins via activity of ADP-ribosyl transferases (*ART* family) and ADP ribose polymerases (*PARP* family) (62, 63). cAMP and protein (poly)ADP-

ribosylation are important mediators of cell signaling and protein expression (64). Niacin is synthesized from tryptophan, but in small quantities relative to a healthy dietary intake (65). Deficiency, known as pellagra, is marked by dermatitis and severe gastrointestinal/neurological pathophysiologies which are fatal if untreated (66). No proteomic studies on systemic intake of vitamin B_3 were found at the time of writing this review.

Vitamin B_5

Pantothenic acid (vitamin B_5) is composed of a molecule of pantoic acid bound to β-alanine (Figure 2.2) (67). Its primary metabolic function is as an acyl-carrier (68). Pantothenic acid is a substrate in the first reaction of coenzyme A (CoA) biosynthesis catalyzed by pantothenate kinases (*PANK1*, *PANK2*, *PANK3*, *PANK4*) (69). CoA is a substrate for enzymes catalyzing the oxidative decarboxylation of ketoacids including the pyruvate dehydrogenase complex (*PDHA*, *PDHB*, *PDHX*, *DLAT*, *DLD*), the oxoglutarate dehydrogenase complex (*OGDH*, *DLST*, *DLD*), and the branched chain keto acid dehydrogenase complex (*BCKDHA*, *DBT*, *DLD*) (70-72). Acyl species are activated by conjugation with CoA and are substrates in or products of glycolysis, the TCA cycle, fatty-acid synthesis/β-oxidation, cholesterol synthesis, ketogenesis, branched-chain amino acid catabolism, and protein acetylation/O-GlcNAcylation (73). Finally, 4'-phosphopanthetheine (product of *PANK* proteins' activities) is a cofactor of the acyl carrier protein domain of fatty acid synthase (*FASN*) (74). Vitamin B_5 deficiency is rare and usually accompanied by that of other nutrients (75). Burning of the feet and numbness in the toes is a characteristic manifestation along with variety of

other symptoms (76). No proteomic studies on systemic intake of vitamin B_5 were found at the time of writing this review.

Vitamin B_6

Vitamin B_6 has aldehyde, alcohol, and amine forms (Figure 2.2); of which the phosphorylated aldehyde form (pyridoxal phosphate) acts as a cofactor to over 100 enzymes (77). All three forms of vitamin B_6 are phosphorylated by pyridoxal kinase (*PDXK*) (78). Both the phosphorylated alcohol and amine forms (pyridoxine phosphate and pyridoxamine phosphate) are converted to pyridoxal phosphate by pyridoxine phosphate oxidase (*PNPO*) (79). Pyridoxal phosphate is a cofactor for enzymes catalyzing decarboxylase reactions in gamma-aminobutyric acid (*GAD1*, *GAD2*) (80) and serotonin/dopamine biosynthesis (*DDC*) (81); as well as for enzymes catalyzing transamination reactions (e.g. *GOT1*, *GOT2*, *GPT*, *GPT2*) (82), cysteine synthesis (*CTH*) (83), heme synthesis (*ALAS1*, *ALAS2*) (84), carnitine synthesis (3-hydroxy-6-N-trimethyllysine aldolase, gene unidentified) (85), niacin synthesis (*KYNU*) (86), and sphingolipid synthesis (*SPTLC1*, *SPTLC2*) (87). Pyridoxal phosphate is also an important cofactor for enzymes of one-carbon metabolism (*SHMT1* and *SHMT2*) (88) and glycogen catabolism (*PYGL* and *PYGM*) (89). Vitamin B_6 deficiency is rare because of its availability in many foods, and pathophysiologies can be diverse (90).

In a triple-TOF-based study of streptozotocin-induced diabetic rat hippocampi, pyridoxamine treatment prevented long-term recognition memory impairment and regulated protein abundances in a number of diverse pathways; notably upregulating half of the proteins involved in ubiquinol biosynthesis (91). In a 2DGE-MS-based study

of mice hippocampi, the abundances of phosphoglycerate mutase (*PGAM1*) and cannabinoid receptor-interacting protein 1 (*CNRIP1*) are reported to be elevated/reduced, respectively, upon administration of pyridoxine. Proteomic changes are accompanied by improved novel object recognition (92).

Vitamin B_7

Biotin (vitamin B_7) is composed of a fused-ring structure bound to a valeric acid side chain (Figure 2.2) (93). It is transported across the plasma membrane by the sodium-dependent solute carriers *SLC5A6* and *SLC19A3* (94, 95). As a cofactor/post-translational modification, biotin covalently binds lysine residues (96). It is a cofactor for pyruvate carboxylase (*PC*), acetyl-CoA carboxylase (*ACACA*), propionyl-CoA carboxylase (*PCCA*), and the methylcrotonyl-CoA carboxylase complex (*MCCC1*, *MCCC2*) (97). Histones are also biotinylated, regulating gene expression (98). The post-translational modification occurs via the activity of holocarboxylase synthetase (*HLCS*) (99).

Biotin deficiency is rare and has wide ranging pathophysiologies. Eating raw egg whites can prevent its absorption (leading to deficiency) because of its affinity for avidin, a chemical in egg whites that is denatured upon cooking. This observation led to the vitamin's eventual discovery (100). No proteomic studies on systemic intake of vitamin B_7 were found at the time of writing this review.

Vitamin B_9

The term folate (vitamin B_9) is inclusive of a group of compounds composed of a

pteridine ring linked to para-aminobenzoic acid with a mono- or polyglutamate tail (Figure 2.2) (101). In its reduced form (tetrahydrofolate), a one-carbon unit cross-links (as CH or CH_2) amine groups on the ring structure and aminobenzoic acid, or binds the secondary amine (as a formyl group) on the aminobenzoic acid group (102, 103). This one-carbon unit is utilized in the synthesis of purines and thymidine, conversion of homocysteine to methionine, interconversion of serine and glycine, and catabolism of histidine; reactions collectively termed one-carbon metabolism (104, 105). At the cellular level, one-carbon metabolism is tightly regulated by compartmentalization (104, 106, 107) while whole-body folate homeostasis is predominantly maintained by the liver through the enterohepatic cycle (108)

Folate deficiency induces megaloblastic macrocytic anemia and fetal neural tube defects, purportedly via its adverse impact on nucleotide synthesis (109, 110). Low intake of folate is also linked to cardiovascular disease (111, 112), neurodegenerative disease (113, 114), Alzheimer's disease (115, 116) and cancer (117-119).

In an orbitrap-based study of follicle fluid of women undergoing in vitro fertilization, the folate supplemented group is reported to have elevated abundances of apolipoproteins from high density lipoproteins and reduced reactive protein c (*CRP*). The study is performed on women who did not become pregnant (120). In a QTOF-based study of a folate-deficiency-induced intestinal neoplasia mouse model, the combinatorial impacts of folate deficiency and methylene tetrahydrofolate reductase heterozygous deletion ($mthfr^{+/-}$) are reported to impact protein abundances spanning diverse cellular functions. However 40% of samples are discarded as outliers and the simultaneous examination of $mthfr^{+/-}$ and dietary folate deficiency does not allow

proteomic adaptations to be attributed to either in isolation (121). In a 2DGE-MS-based study of adult rats, aortic calmodulin (*CALM1*, calcium signaling) protein abundances are positively correlated with folate dose while abundances of triose phosphate isomerase (*TPI1*, glycolysis), transgelin (*TAGLN*, cytoskeleton), and glutathione s-transferase alpha 3 (*GSTA3*, reductive detoxification) respond inversely (122). In an 2DGE-MS-based study of rat livers, *PRDX6* and *GPX1* are reported to be elevated while cofilin (*CFL1*) is reported to be depleted under folate deficiency (123). Other studies report protein abundance differences due to folate intake in rat urinary exosomes (QQQ-based) (124), human plasma (2DGE-MS) (125), fetal brain tissue from pregnant mice fed ethanol (2DGE-MS) (126), pregnant rat livers (2DGE-MS) (127), fetal rat livers (2DGE-MS) (128), adult rat livers and brains (2DGE-MS) (129), and livers of piglets born to folate deficient mothers (2DGE-MS) (130).

Vitamin B_{12}

Cobalamin (vitamin B_{12}) encompasses a group of molecules with four linked pyrrole ring derivatives (forming a corrin ring) and a cobalt atom bound at the center of the corrin ring. The cobalt atom also binds a 5,6-dimethylbenzimidazole nucleotide and a functional group (Figure 2.2) (131). The identity of the functional group distinguishes the vitamin B_{12} compounds as cyanocobalamin, hydroxycobalamin, hydrocobalamin, nitrocobalamin, 5'-deoxyadenosylcobalamin (also called adenosylcobalamin), and methyl cobalamin (132, 133). Methylcobalamin serves as a coenzyme in the conversion of homocysteine to methionine by methionine synthase (*MTR*) in the cytosol (134) and adenosylcobalamin is required for conversion of L-methylmalonyl-CoA to succinyl-

CoA by methylmalonyl-CoA mutase (*MUT*) in mitochondria (135).

Vitamin B_{12} deficiency is closely related to folate deficiency and can lead to megaloblastic anemia by impairment in the activity of methionine synthase (*MTR*) (109): 5-methyl tetrahydrofolate cannot be converted to one-carbon donors required for purine and thymidine synthesis without vitamin B_{12} as a cofactor, thus interfering with DNA synthesis and erythrocyte production (136). Vitamin B_{12} deficiency is also linked to neurological disorders independent of anemia (137).

Ruoppolo and colleagues performed a 2DGE-MS-based study of lymphocytes isolated from methylmalonic acidemia with homocystinuria, cobalamin deficiency type C (MMACHC) patients (an inborn error in metabolism marked by inactivity of the *MMACHC* gene product) receiving a standard treatment of hydroxycobalamin, betaine, folate, and carnitine. Protein products of *ME2*, *GLUD1*, and *GPD2*, genes involved in anaplerosis and redox equivalent shuttling, are up-regulated while variant 2 of protein pyruvate kinase muscle isozyme (*PKM*) and lactate dehydrogenase B (*LDHB*) are down-regulated relative to lymphocytes isolated from healthy control donors (138). In a 2DGE-based study of adult rat cerebral spinal fluid, protein abundance shifts are reported to peak after several months on a cobalamin deficient diet (modest shifts) or after a total gastrectomy (more severe shifts), and return to near control values at later time points (139). In a 2DGE-MS-based study, glutathione s-transferase P (*GSTP1*) abundances are diminished and glutathione peroxidase 1 (*GPX1*) abundances are elevated in rat pup kidneys under maternal vitamin B_{12} deficient and maternal folate deficient conditions (140); suggesting maternal dietary intake of these vitamins impacts offspring kidney redox homeostasis mechanisms. In a similar 2DGE-MS-based study

of maternal vitamin B_{12} deficiency, the same group reports that several dozen rat kidney pup proteins revert to control levels upon administration of vitamin B_{12} at birth. Additionally, diminished abundance of beta-oxidation proteins in kidneys of pups born to vitamin B_{12} deficient mothers is accompanied by elevated *PPARA* (141), a positive regulator of fatty acid oxidation, suggesting attempted compensation at the cellular level.

Vitamin C

Vitamin C (ascorbic acid) is absorbed at the brush-border and distributed to cells throughout the body by the sodium-dependent plasma membrane solute carriers *SLC23A1* and *SLC23A2* (142). The oxidized form of vitamin C (dehydroascorbate) is also transported via plasma membrane glucose transporters *SLC2A1*, *SLC2A3*, and *SLC2A4* (also known as *GLUT1*, *GLUT3*, and *GLUT4*) (143) and reduced intracellularly to ascorbic acid by glutathione (144) and the activity of thioredoxin reductases (*TXNRD1*, *TXNRD2*, or *TXNRD3*) (145).

Vitamin C is a cofactor in the function of prolyl and lysyl hydroxylases, which consume oxygen and alpha-ketoglutarate to form the hydroxylated amino acid residue and succinate (146). The Fe^{2+} of these enzymes is restored from Fe^{3+} by oxidation of vitamin C (147). In the presence of oxygen, prolyl hydroxylases (*EGLN1*, *EGLN2*, *EGLN3*; also known as *PHD2*, *PHD1*, *PHD3* respectively) hydroxylate the *HIF1A* protein; providing a necessary signal for its degradation and preventing a hypoxic response at the cellular level (148). Prolyl and lysyl hydroxylase activities are also necessary for post-translational modifications to form functional collagen (149). Lysyl

hydroxylases include *PLOD1*, *PLOD2*, and *PLOD3* (150). Vitamin C serves a nearly identical function in reducing Fe^{3+} as a cofactor for trimethyllysine dioxygenase (*TMLH*), which catalyzes the first reaction in carnitine biosynthesis (151). Carnitine is essential for fatty acid catabolism in the mitochondria as only fatty acyl carnitines formed via the activity of carnitine palmitoyl transferases *CPT1A*, *CPT1B*, and *CPT1C* cross the inner mitochondrial membrane through the solute carrier *SLC25A20* (152). Vitamin C similarly serves as a cofactor for tyrosine hydroxylase (*TH*), which catalyzes the first reaction in catecholamine (e.g. dopamine, epinephrine, and norepinephrine) synthesis (153). Additionally, vitamin C serves and as a general antioxidant (154). Vitamin C deficiency leads to the condition known as scurvy with symptoms largely attributed to malformed connective tissue due to improperly folded collagen (155).

In a orbitrap-based study on a pig model of hemorrhagic shock, vitamin C administration is reported to impact plasma protein abundances in the complement pathway and those in poly-trauma related processes; including the stabilization of *ADAMTS13* abundance, an important regulator of clot formation (156). An orbitrap-based study of endoplasmic reticulum enriched fractions of livers in Werner syndrome mouse models identifies around a dozen proteins whose abundances are impacted by administration of vitamin C (157). A QTOF-based study of zebrafish reports upregulation of glutamate dehydrogenase (*GLUD1*) and downregulation of pyruvate kinase muscle isozyme (*PKM*) upon administration of vitamin C in a vitamin E deficient background (158). In a QQQ-based study of human plasma, ascorbic acid concentration is reported to be inversely related to vitamin D binding protein (*GC*) abundance (159). 2DGE-MS-based studies identify protein abundance regulations in mouse models of

sarcoma metastases in the liver (160) and tumor nodules of adenocarcinoma due to administration of vitamin C (161). Another 2DGE-MS-based study reports polypeptide abundance shifts in hemodialysis patient plasma upon vitamin C supplementation (162).

Vitamin D

Vitamins D_2 and D_3 are respectively distinguished by their ergosterol and cholesterol backbones (163). Though only vitamin D_3 is synthesized in animals, both can be converted to active forms. Exposure of 7-dehydrocholesterol (an intermediate in cholesterol synthesis) to ultra-violet radiation in the skin and subsequent isomerization produces cholecalciferol (vitamin D_3, Figure 2.1) (164). Whether 7-dehydrocholesterol is derived from cholesterol via activity of 7-dehydrocholesterol reductase (*DHCR7*) or synthesized de novo in the skin is disputed (165). 7-dehydrocholesterol is successively hydroxylated by activity of cytochrome p450 enzymes (e.g. *CYP2R1* and *CYP27B1*) in the liver and kidney to its active 1,25-$(OH)_2$ cholecalciferol [$1,25(OH)_2D_3$] form (166). Transport of vitamin D and its metabolites occurs bound to vitamin D binding protein (*GC*) (167). Ergocalciferol is the vitamin D_2 equivalent of cholecalciferol and is activated analogously (168).

$1,25(OH)_2D_3$ influences cellular function via nuclear receptor-dependent and nuclear receptor-independent mechanisms. The former involves $1,25(OH)_2D_3$-bound vitamin D receptor (*VDR*) forming a heterodimer complex with a retinoid X receptor (*RXRA*, *RXRB*, *RXRG*) and subsequently binding vitamin D response elements regulating transcription of genes largely involved modulating calcium and phosphorous transport (169) and maintaining homeostasis by regulating their absorption in the

kidneys, intestines, and bones (170, 171). The rapid-onset extracellular impacts (nuclear receptor-independent) of $1,25(OH)_2D_3$ are mediated by a membrane-associated rapid response steroid binding protein, identified as *PDIA3* (172), and diversely impact cell growth, survival, and immune response (173).

Deficiency in vitamin D impairs bone mineralization causing rickets in infants/children and osteomalacia in adults (174). Vitamin D deficiency is also linked to cardiovascular diseases (175, 176), cancer (177, 178), neurological impairments (179, 180) and autoimmune diseases (181, 182); though underlying mechanisms are not completely understood.

In an orbitrap-based study of mouse fetal and postnatal lung tissue, maternal vitamin D deficiency is reflected in total proteome adaptations which are unexpectedly strongest at postnatal day 7 opposed to fetal time points. Impacted proteins include several associated with lung development (183). An orbitrap-based study of a mouse brain tissue model of remyelination in multiple sclerosis reports calcium binding protein abundances to be upregulated upon treatment with $1,25(OH)_2D_3$, consistent with the vitamin's regulatory role over calcium absorption (184). In an orbitrap-based study of serum from overweight adults, vitamin D deficiency is reported to differentially affect abundances of proteins related to blood coagulation in males and females. However, abundances of these proteins are likely impacted by the production of serum from whole blood. The authors report quantifying 1,841 proteins (Table 2.1); an impressive analytical depth for serum (185). In a 2DGE-based study, vitamin D deficient children are reported to have diminished serum abundances of adiponectin (*ADIPOQ*) (186). In a separate 2DGE-based study, the same group reports fetuin-b (*FETUB*) to be elevated

in the plasma of obese vitamin D deficient children compared with their vitamin D sufficient counterparts (187). However the authors do not directly identify *FETUB* and rely on comparison of their findings to those of another study (188). Two 2DGE-MS-based studies, of rat left ventricular and aortic tissue, identify proteins whose abundances respond upon inducing arterial calcification or atherosclerosis by co-administration of vitamin D_3 with nicotine or a high cholesterol diet respectively (189, 190). Two studies (2DGE-MS and 2DGE-based respectively) examine the impacts of vitamin D deficiency on the rat brain proteome. The former reports the progeny of vitamin D deficient mothers to have diminished abundances of ATP synthase β (*ATPB*) and enolase 2 (*ENO2*) in both the cortex and hippocampus, and diminished calmodulin (*CALM1*) in the hippocampus amongst a variety of other regulated proteins (191). The latter finds low vitamin D diets to be accompanied by diminished cortical abundances of three glycolytic enzymes: triose phosphate isomerase (*TPI1*); phosphofructokinase, platelet (*PFKP*); and pyruvate kinase, muscle (*PKM*) (192).

Table 2.1. Summary of Key Findings

Vitamin	Key Findings
Vitamin A	Creation of a model of craniofacial disorders induced by prenatal retinoic acid exposure is reported to impact protein abundances whose functions are associated with neural crest processes (30). Plasma carotenoids abundances are reported to be associated with plasma proteins of diverse functions in Nepalese children, potentiating development of inexpensive assays to predict carotenoids deficiency (32).
Vitamin B_1	Treatment with thiamine is presented as a potential strategy to improve kidney function in type 2 diabetic patients (41). Thiamine deficiency is reported to impact cognition in rats (42).
Vitamin B_2	Dietary and maternal dietary riboflavin to impact the machineries of lipid metabolism and fetal lipid metabolism in ducks (51, 52).
Vitamin B_6	A rat model of diabetes reports pyridoxamine treatment to impact proteins involved in synaptic plasticity in hippocampi and have protective effects on long-term memory (91).
Vitamin B_9	Folate supplementation in women undergoing in vitro fertilization is reported to increase abundances of apolipoproteins of high-density lipoproteins in monofollicular fluid (120).
Vitamin B_{12}	Maternal cobalamin deficiency is reported to impact abundances of proteins related to lipid metabolism in the offspring kidneys of rats (141).
Vitamin C	Treatment with ascorbic acid is reported to impact plasma abundances of proteins involved in complement and coagulation pathway in a pig model of hemorrhagic shock (156).
Vitamin D	Maternal vitamin D deficiency is reported to impact abundances of proteins involved in mouse neonatal lung development during alveolar development stages without affecting gross lung structure (183). Treatment with $1,25(OH)_2D_3$ is reported to increase abundances of proteins involved in calcium homeostasis in a mouse brain model of remyelination (184).
Vitamin E	A quantitative model based on plasma protein abundances is reported to predict plasma α-tocopherol status, potentiating the development of an inexpensive assay to detect α-tocopherol deficiency (193). A mouse model of Alzheimer's disease reports treatment with the tocotrienol-rich fraction of palm oil reduces the abundance of amyloid beta A4 protein, the primary component of amyloid plaques, in hippocampi (194).
Vitamin K	A quantitative model based on five plasma protein abundances is reported to predict vitamin K deficiency with moderate accuracy (195).

Vitamin E

Members of the vitamin E class of molecules all contain fused phenyl and chromanol rings linked to a 16-carbon side-chain (196). Methyl group placement on the phenyl ring dictates α, β, γ, and δ designation while side-chain saturation state distinguishes tocopherols from tocotrienols (Figure 2.1). Furthermore, all forms of vitamin E have three chiral centers resulting in 8 stereoisomers (197). RRR α-tocopherol is the most biologically active form, likely due to specificity of α-tocopherol transfer protein (*TTPA*) whose binding is necessary for packaging and transport to tissues from the liver (198). α-tocopherol primarily localizes to membranes (i.e. plasma, endoplasmic reticulum, and mitochondrial) and functions as an antioxidant for unsaturated, lipid-bound fatty acids (196). A-tocopherol also has non-antioxidant signal-transduction functions impacting a broad range of cellular activities (199).

Vitamin E deficiency is rare due to the availability of the vitamin in the diet (200), though it may be caused by a genetic defect in α-tocopherol transfer protein (*TTPA*) and diseases associated with fat malabsorption (201, 202). Severe vitamin E deficiency can result in hemolytic anemia, neurological disorders, and ataxia (201, 203-205).

An orbitrap-based study of plasma from undernourished Nepalese children reports plasma α-tocopherol concentration to be positively correlated with abundances of a number of apolipoproteins (*APO*s) and negatively correlated with the muscle isozyme of the protein pyruvate kinase (*PKM*). Authors establish a linear model based on a handful of protein quantities that accounts for most variance in α-tocopherol plasma concentration and suggest an inexpensive, portable, antibody-based methodology can be used to assay plasma α-tocopherol abundance in low-income countries (193). An

orbitrap-based study of hippocampi, medial prefrontal cortices, and striata tissue in a mouse model of Alzheimer's disease reports administration of a tocotrienol-rich fraction of palm oil down-regulates hippocampi expression of the amyloid beta A4 protein (*APP*). Amyloid beta A4 is the principle component of amyloid plaques characteristic of Alzheimer's disease (194). In a QTOF-based study of rabbit aortae, vitamin E supplementation is reported to impact protein abundances including the apolipoprotein, *APOA1*, and several related to oxidation/reduction processes (206). 2DGE-MS-based studies also reports vitamin E supplementation to impact apolipoprotein abundances in human plasma (207, 208). In a 2DGE-based study of high-density reared rainbow trout livers, vitamin E supplementation is reported to regulate the abundances of a handful of heat shock and metabolic proteins (209). Finally, a 2DGE-MS-based study reports vitamin E supplementation to regulate a number of plasma protein abundance in patients harboring prostate tumors (210).

Vitamin K

Vitamin K compounds all share a common fused benzyl and methyl-naphthoquinone ring moiety (Figure 2.1). Naturally occurring vitamin K compounds include phylloquinone and menaquinones (211). Vitamin K is a necessary cofactor of gamma-glutamyl carboxylase (*GGCX*), an enzyme which catalyzes the carboxylation of glutamate protein residues to carboxyglutamate residues (212). This post-translational modification is necessary for the function of proteins of the coagulation cascade (*F2*, *F7*, *F9*, *F10*), proteins inhibiting coagulation (*PROC*, *PROS1*, *PROZ*), and those associated with connective tissue matrix formation (*BGLAP*, *MGP*) (213).

Newborn infants are among the most at-risk for vitamin K deficiency because they do not have adequate stores and milk is not a sufficient source. Thus a phylloquinone injection shortly after birth is recommended (214). Elevated risk of hemorrhage is associate with vitamin K deficiency (215).

In an orbitrap-based study of plasma from Nepalese children, authors create a model based on five protein abundances which can predict vitamin K deficiency with moderate accuracy. Vitamin K status is based on a surrogate measurement of an abundance of an abnormal form of prothrombin (195).

Conclusions

Proteomic studies have established dietary vitamin status as a regulator of tissue protein abundances. The regulatory feedback between vitamin status and protein expression is highlighted by findings where the abundances of proteins directly related to the vitamin are impacted by systemic intake of that vitamin, including: abundances of proteins related to craniofacial development and neural crest processes are impacted in an established maternal retinoic acid toxicity-driven model of craniofacial birth defects (30), deficiency in their riboflavin cofactor is accompanied by reduced abundances of acyl-coenzyme A dehydrogenases and accumulation of the enzymes' substrates (51), treatment with the active form of vitamin D is accompanied by increased expression of calcium binding proteins (184), and vitamin E supplementation impacts proteins related to redox processes (206) (Table 2.1). However, the literature in this field is sparse and, in all likelihood, the vast majority of vitamin-status to protein abundance relationships are undescribed; especially considering the void in the literature for several

vitamins. Moreover, the field has not advanced to explore the mechanisms of these regulations, their biological impacts, or their potential to shape clinical interventions. Modern proteomic platforms are ever increasing achievable depth, where analysis of whole mammalian tissue (216-220) or plasma (221, 222) routinely results in 10,000 or 1,500 unique proteins quantified per sample respectively. Furthermore, advances continue in capacity to detect post-translational modifications (223, 224), determine compartmental localization (225), and apply findings in clinical settings (226). In a golden age of proteomic technological advances, studies of vitamin intake have not kept pace. Of those using orbitrap, QTOF, or triple-TOF systems, many fall short of the cutting edge of analytical depth (Table 2.2); whereas most studies have relied on antiquated 2DGE-MS platforms. Outdated platforms rendering fewer quantified proteins are likely contributors to clustering (227) or enrichment analysis techniques (228-234) not being widely employed. These high-throughput methods of data analysis provide systems-level stratification of proteome-wide adaptations and can guide targeted inquiries. As the study of precision nutrition advances in an era of big data, fundamental questions of nutrient-protein interactions will be at the forefront of understanding molecular mechanisms of nutrient and substrate processing. Where the sparsity of the literature leaves fundamental questions unanswered, opportunity for rapid advancement lies with application of cutting-edge technologies in well-designed and executed studies.

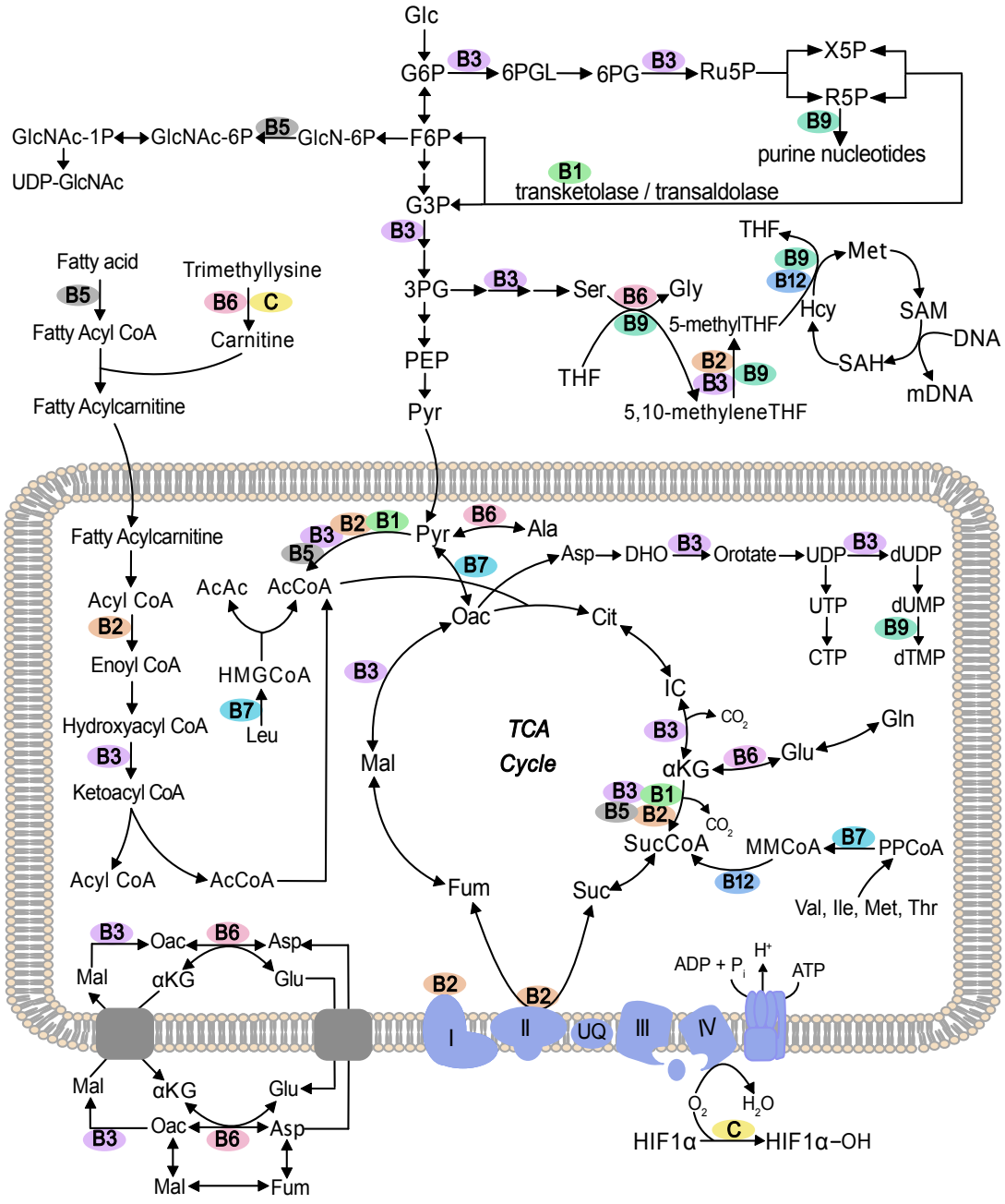


Figure 2.3. Schematic of Vitamin Involvement in Reactions of Central Carbon Metabolism. Abbreviations defined in the list of abbreviations section. I, II, III, and V refer to electron transport chain complex numbers.

Table 2.2. Summary of Technical Depth of Orbitrap-, QTOF-, and Triple-TOF-based Studies.

Study	Year	Platform	Tissue	# Proteins Identified	# Samples
(30)	2018	Orbitrap	mouse embryo heads	group 1: 313 total group 2: 372 total	2 groups
(31)	2018	Triple-TOF	gerbil plasma	109 total	30
(32, 193, 195)	2015	Orbitrap	human plasma	4,705 total 589/set	72 iTRAQ sets
(42)	2018	QTOF	rat thalami	1,440 total	6 x 3 tech. reps.
(51)	2017	QTOF	duck livers	1,749 total	3 iTRAQ sets
(52)	2019	QTOF	fetal duck livers	3,801 total	1 iTRAQ set
(91)	2019	Triple-TOF	rat hippocampi	4,807 total	2 iTRAQ sets
(120)	2015	Orbitrap	human follicular fluid	227 total	1 TMT set
(121)	2014	QTOF	mouse intestine	2,039 total	10
(156)	2019	Orbitrap	pig plasma	534 total	45
(157)	2018	Orbitrap	mouse liver fraction	4,058 total	9
(158)	2014	QTOF	zebrafish	2,956 total	19
(183)	2016	Orbitrap	mouse lung	1,160 total, 240 common to all	34
(184)	2018	Orbitrap	mouse brain	5,062 total	1 TMT set
(185)	2016	Orbitrap	human serum	1,841 total	1 iTRAQ set
(194)	2019	Orbitrap	mouse brain tissues	group 1: 5,847 total group 2: 6,047 total	2 groups of 6
(206)	2013	QTOF	rabbit aortae	100 total	24 x 3 tech. reps.

A protein only needs to be identified in one sample to contribute to the total number of unique proteins (indicated as "total" below), thus the total is typically larger in studies with greater sample numbers. An iTRAQ or TMT set is a pool of samples that are run on the LC-MS/MS concurrently. Because the samples in a set are all analyzed

simultaneously, a set's contribution to the total number of unique proteins is similar to that of a single sample.

Acknowledgements

The authors wish to acknowledge invaluable resources for fundamental information presented in this work (235-237). This chapter is a reprint of “Systemic vitamin intake impacting tissue proteomes”, *Nutrition & Metabolism*, vol. 17, 2020. Heesoo Jeong is the first author of the publication.

REFERENCES

1. Metallo CM, Vander Heiden MG. Understanding metabolic regulation and its influence on cell physiology. Mol Cell. 2013;49(3):388-98.

2. Dervartanian DV, Veeger C. Studies on succinate dehydrogenase. II. On the nature of the reaction of competitive inhibitors and substrates with succinate dehydrogenase. Biochim Biophys Acta. 1965;105(3):424-36.

3. Underwood AH, Newsholme EA. Control of glycolysis and gluconeogenesis in rat kidney cortex slices. Biochem J. 1967;104(1):300-5.

4. Nguyen T, Nioi P, Pickett CB. The Nrf2-antioxidant response element signaling pathway and its activation by oxidative stress. J Biol Chem. 2009;284(20):13291-5.

5. Kopperschläger G, Kirchberger J. Methods for the separation of lactate dehydrogenases and clinical significance of the enzyme. J Chromatogr B Biomed Appl. 1996;684(1-2):25-49.

6. Patel MS, Nemeria NS, Furey W, Jordan F. The pyruvate dehydrogenase complexes: structure-based function and regulation. J Biol Chem. 2014;289(24):16615-23.

7. Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, et al. Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature. 2011;481(7381):380-4.

8. Efeyan A, Comb WC, Sabatini DM. Nutrient-sensing mechanisms and pathways. Nature. 2015;517(7534):302-10.

9. Kandel ES, Hay N. The regulation and activities of the multifunctional serine/threonine kinase Akt/PKB. Exp Cell Res. 1999;253(1):210-29.

10. Lefebvre P, Martin PJ, Flajollet S, Dedieu S, Billaut X, Lefebvre B. Transcriptional activities of retinoic acid receptors. Vitam Horm. 2005;70:199-264.

11. Kojo S. Vitamin C: basic metabolism and its function as an index of oxidative stress. Curr Med Chem. 2004;11(8):1041-64.

12. Minor EA, Court BL, Young JI, Wang G. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. J Biol Chem. 2013;288(19):13669-74.

13. Carlberg C, Seuter S. A genomic perspective on vitamin D signaling. Anticancer Res. 2009;29(9):3485-93.

14. Pike JW, Christakos S. Biology and Mechanisms of Action of the Vitamin D Hormone. Endocrinol Metab Clin North Am. 2017;46(4):815-43.

15. Kamanna VS, Kashyap ML. Mechanism of action of niacin on lipoprotein metabolism. Curr Atheroscler Rep. 2000;2(1):36-46.

16. Chapman-Smith A, Cronan JE, Jr. The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. Trends Biochem Sci. 1999;24(9):359-63.

17. Rochat B. Quantitative and Qualitative LC-High-Resolution MS: The Technological and Biological Reasons for a Shift of Paradigm. Recent Advances in Analytical Chemistry: IntechOpen; 2018.

18. Bromirski M. First choice in high resolution mass spectrometry with Orbitrap mass analyzer technology for screening, confirmative and quantitative analyses. Thermo Scientific. 2018(White Paper 65146).

19. Alelyunas YW, Wrona MD, Cook K, McDonald S, Rainville PD. Effect of MS Scan Speed on UPLC Peak Separation and Metabolite Identification: Time-of-Flight HRMS vs. Orbitrap. Waters Corporation. 2013.

20. Sporn MB, Dunlop NM, Newton DL, Henderson WR. Relationships between structure and activity of retinoids. Nature. 1976;263(5573):110-3.

21. von Lintig J. Provitamin A metabolism and functions in mammalian biology. Am J Clin Nutr. 2012;96(5):1234s-44s.

22. Amengual J, Widjaja-Adhi MA, Rodriguez-Santiago S, Hessel S, Golczak M, Palczewski K, et al. Two carotenoid oxygenases contribute to mammalian provitamin A metabolism. J Biol Chem. 2013;288(47):34081-96.

23. Bownds D. Site of attachment of retinal in rhodopsin. Nature. 1967;216(5121):1178-81.

24. Jacobson SG, Cideciyan AV, Regunath G, Rodriguez FJ, Vandenburgh K, Sheffield VC, et al. Night blindness in Sorsby's fundus dystrophy reversed by vitamin A. Nat Genet. 1995;11(1):27-32.

25. Petkovich M. Regulation of gene expression by vitamin A: the role of nuclear retinoic acid receptors. Annu Rev Nutr. 1992;12:443-71.

26. Szanto A, Narkar V, Shen Q, Uray IP, Davies PJ, Nagy L. Retinoid X receptors: X-ploring their (patho)physiological functions. Cell Death Differ. 2004;11 Suppl 2:S126-43.

27. Herr FM, Ong DE. Differential interaction of lecithin-retinol acyltransferase with cellular retinol binding proteins. Biochemistry. 1992;31(29):6748-55.

28. Kim CI, Leo MA, Lieber CS. Retinol forms retinoic acid via retinal. Arch Biochem Biophys. 1992;294(2):388-93.

29. Wiseman EM, Bar-El Dadon S, Reifen R. The vicious cycle of vitamin a deficiency: A review. Crit Rev Food Sci Nutr. 2017;57(17):3703-14.

30. Berenguer M, Darnaudery M, Claverol S, Bonneu M, Lacombe D, Rooryck C. Prenatal retinoic acid exposure reveals candidate genes for craniofacial disorders. Sci Rep. 2018;8(1):17492.

31. Bohn T, Planchon S, Leclercq CC, Renaut J, Mihaly J, Beke G, et al. Proteomic responses of carotenoid and retinol administration to Mongolian gerbils. Food Funct. 2018;9(7):3835-44.

32. Eroglu A, Schulze KJ, Yager J, Cole RN, Christian P, Nonyane BAS, et al. Plasma proteins associated with circulating carotenoids in Nepalese school-aged children. Arch Biochem Biophys. 2018;646:153-60.

33. Zhang M, Huang K, Zhang Z, Ji B, Zhu H, Zhou K, et al. Proteome alterations of cortex and hippocampus tissues in mice subjected to vitamin A depletion. J Nutr Biochem. 2011;22(11):1003-8.

34. Campos CF, Costa TC, Rodrigues RTS, Guimaraes SEF, Moura FH, Silva W, et al. Proteomic analysis reveals changes in energy metabolism of skeletal muscle in beef cattle supplemented with vitamin A. J Sci Food Agric. 2020.

35. Brown G. Defects of thiamine transport and metabolism. J Inherit Metab Dis. 2014;37(4):577-85.

36. Ganapathy V, Smith SB, Prasad PD. SLC19: the folate/thiamine transporter family. Pflugers Arch. 2004;447(5):641-6.

37. Jordan F. Current mechanistic understanding of thiamin diphosphate-dependent enzymatic reactions. Nat Prod Rep. 2003;20(2):184-201.

38. Kochetov GA, Solovjeva ON. Structure and functioning mechanism of transketolase. Biochim Biophys Acta. 2014;1844(9):1608-18.

39. Bettendorff L, Wins P, Schoffeniels E. Regulation of ion uptake in membrane vesicles from rat brain by thiamine compounds. Biochem Biophys Res Commun. 1990;171(3):1137-44.

40. Lonsdale D. Thiamin. Adv Food Nutr Res. 2018;83:1-56.

41. Riaz S, Skinner V, Srai SK. Effect of high dose thiamine on the levels of urinary protein biomarkers in diabetes mellitus type 2. J Pharm Biomed Anal. 2011;54(4):817-25.

42. Nunes PT, Gomez-Mendoza DP, Rezende CP, Figueiredo HCP, Ribeiro AM. Thalamic Proteome Changes and Behavioral Impairments in Thiamine-deficient Rats. Neuroscience. 2018;385:181-97.

43. Massey V. The chemical and biological versatility of riboflavin. Biochem Soc Trans. 2000;28(4):283-96.

44. Barile M, Giancaspero TA, Leone P, Galluccio M, Indiveri C. Riboflavin transport and metabolism in humans. J Inherit Metab Dis. 2016;39(4):545-57.

45. Lienhart WD, Gudipati V, Macheroux P. The human flavoproteome. Arch Biochem Biophys. 2013;535(2):150-62.

46. Chan J, Deng L, Mikael LG, Yan J, Pickell L, Wu Q, et al. Low dietary choline and low dietary riboflavin during pregnancy influence reproductive outcomes and heart development in mice. Am J Clin Nutr. 2010;91(4):1035-43.

47. Yates CA, Evans GS, Powers HJ. Riboflavin deficiency: early effects on post-weaning development of the duodenum in rats. Br J Nutr. 2001;86(5):593-9.

48. Powers HJ, Weaver LT, Austin S, Wright AJ, Fairweather-Tait SJ. Riboflavin deficiency in the rat: effects on iron utilization and loss. Br J Nutr. 1991;65(3):487-96.

49. Bian X, Gao W, Wang Y, Yao Z, Xu Q, Guo C, et al. Riboflavin deficiency affects lipid metabolism partly by reducing apolipoprotein B100 synthesis in rats. J Nutr Biochem. 2019;70:75-81.

50. Olpin SE, Bates CJ. Lipid metabolism in riboflavin-deficient rats. 1. Effect of dietary lipids on riboflavin status and fatty acid profiles. Br J Nutr. 1982;47(3):577-96.

51. Tang J, Hegeman MA, Hu J, Xie M, Shi W, Jiang Y, et al. Severe riboflavin deficiency induces alterations in the hepatic proteome of starter Pekin ducks. Br J Nutr. 2017;118(9):641-50.

52. Tang J, Hu J, Xue M, Guo Z, Xie M, Zhang B, et al. Maternal diet deficient in riboflavin induces embryonic death associated with alterations in the hepatic proteome of duck embryos. Nutr Metab (Lond). 2019;16:19.

53. Cantó C, Menzies KJ, Auwerx J. NAD(+) Metabolism and the Control of Energy Homeostasis: A Balancing Act between Mitochondria and the Nucleus. Cell Metab. 2015;22(1):31-53.

54. Magni G, Amici A, Emanuelli M, Orsomando G, Raffaelli N, Ruggieri S. Enzymology of NAD+ homeostasis in man. Cell Mol Life Sci. 2004;61(1):19-34.

55. Meyer-Ficca M, Kirkland JB. Niacin. Adv Nutr. 2016;7(3):556-8.

56. Njus D, Kelley PM, Harnadek GJ, Pacquing YV. Mechanism of ascorbic acid regeneration mediated by cytochrome b561. Ann N Y Acad Sci. 1987;493:108-19.

57. Vogel R, Wiesinger H, Hamprecht B, Dringen R. The regeneration of reduced glutathione in rat forebrain mitochondria identifies metabolic pathways providing the NADPH required. Neurosci Lett. 1999;275(2):97-100.

58. Arnér ES, Holmgren A. Physiological functions of thioredoxin and thioredoxin reductase. Eur J Biochem. 2000;267(20):6102-9.

59. Lerner F, Niere M, Ludwig A, Ziegler M. Structural and functional characterization of human NAD kinase. Biochem Biophys Res Commun. 2001;288(1):69-74.

60. Pollak N, Dölle C, Ziegler M. The power to reduce: pyridine nucleotides--small molecules with a multitude of functions. Biochem J. 2007;402(2):205-18.

61. Gill DM, Meren R. ADP-ribosylation of membrane proteins catalyzed by cholera toxin: basis of the activation of adenylate cyclase. Proc Natl Acad Sci U S A. 1978;75(7):3050-4.

62. Schreiber V, Dantzer F, Ame JC, de Murcia G. Poly(ADP-ribose): novel functions for an old molecule. Nat Rev Mol Cell Biol. 2006;7(7):517-28.

63. Corda D, Di Girolamo M. Functional aspects of protein mono-ADP-ribosylation. Embo j. 2003;22(9):1953-8.

64. Ziegler M. New functions of a long-known molecule. Emerging roles of NAD in cellular signaling. Eur J Biochem. 2000;267(6):1550-64.

65. Ikeda M, Tsuji H, Nakamura S, Ichiyama A, Nishizuka Y, Hayaishi O. STUDIES ON THE BIOSYNTHESIS OF NICOTINAMIDE ADENINE DINUCLEOTIDE. II. A ROLE OF PICOLINIC CARBOXYLASE IN THE BIOSYNTHESIS OF NICOTINAMIDE ADENINE DINUCLEOTIDE FROM TRYPTOPHAN IN MAMMALS. J Biol Chem. 1965;240:1395-401.

66. Hegyi J, Schwartz RA, Hegyi V. Pellagra: dermatitis, dementia, and diarrhea. Int J

Dermatol. 2004;43(1):1-5.

67. Williams RJ, Major RT. THE STRUCTURE OF PANTOTHENIC ACID. Science. 1940;91(2358):246.

68. Tahiliani AG, Beinlich CJ. Pantothenic acid in health and disease. Vitam Horm. 1991;46:165-228.

69. Leonardi R, Jackowski S. Biosynthesis of Pantothenic Acid and Coenzyme A. EcoSal Plus. 2007;2(2).

70. Pettit FH, Pelley JW, Reed LJ. Regulation of pyruvate dehydrogenase kinase and phosphatase by acetyl-CoA/CoA and NADH/NAD ratios. Biochem Biophys Res Commun. 1975;65(2):575-82.

71. McMinn CL, Ottaway JH. Studies on the mechanism and kinetics of the 2-oxoglutarate dehydrogenase system from pig heart. Biochem J. 1977;161(3):569-81.

72. Coon MJ, Robinson WG. Amino acid metabolism. Annu Rev Biochem. 1958;27(3):561-612.

73. Novelli GD. Metabolic functions of pantothenic acid. Physiol Rev. 1953;33(4):525-43.

74. Larrabee AR, McDaniel EG, Bakerman HA, Vagelos PR. Acyl carrier protein. V. Identification of 4'-phosphopantetheine bound to a mammalian fatty acid synthetase preparation. Proc Natl Acad Sci U S A. 1965;54(1):267-73.

75. Hodges RE, Ohlson MA, Bean WB. Pantothenic acid deficiency in man. J Clin Invest. 1958;37(11):1642-57.

76. Bean WB, Hodges RE, Daum K. Pantothenic acid deficiency induced in human subjects. J Clin Invest. 1955;34(7, Part 1):1073-84.

77. Snell EE. Chemical structure in relation to biological activities of vitamin B6. Vitam Horm. 1958;16:77-125.

78. Hanna MC, Turner AJ, Kirkness EF. Human pyridoxal kinase. cDNA cloning, expression, and modulation by ligands of the benzodiazepine receptor. J Biol Chem. 1997;272(16):10756-60.

79. McCormick DB, Chen H. Update on interconversions of vitamin B-6 with its coenzyme. J Nutr. 1999;129(2):325-7.

80. Denner LA, Wu JY. Two forms of rat brain glutamic acid decarboxylase differ in their dependence on free pyridoxal phosphate. J Neurochem. 1985;44(3):957-65.

81. Bertoldi M. Mammalian Dopa decarboxylase: structure, catalytic activity and inhibition. Arch Biochem Biophys. 2014;546:1-7.

82. Karpeisky MY, Ivanov VI. A molecular mechanism for enzymatic transamination. Nature. 1966;210(5035):493-6.

83. Kabil O, Banerjee R. Redox biochemistry of hydrogen sulfide. J Biol Chem. 2010;285(29):21903-7.

84. Brown BL, Kardon JR, Sauer RT, Baker TA. Structure of the Mitochondrial Aminolevulinic Acid Synthase, a Key Heme Biosynthetic Enzyme. Structure. 2018;26(4):580-9.e4.

85. Strijbis K, Vaz FM, Distel B. Enzymology of the carnitine biosynthesis pathway. IUBMB Life. 2010;62(5):357-62.

86. Phillips RS. Structure and mechanism of kynureninase. Arch Biochem Biophys. 2014;544:69-74.

87. Bourquin F, Capitani G, Grütter MG. PLP-dependent enzymes as entry and exit gates of sphingolipid metabolism. Protein Sci. 2011;20(9):1492-508.

88. Schirch V, Szebenyi DM. Serine hydroxymethyltransferase revisited. Curr Opin Chem Biol. 2005;9(5):482-7.

89. Helmreich EJ. How pyridoxal 5'-phosphate could function in glycogen phosphorylase catalysis. Biofactors. 1992;3(3):159-72.

90. Wiss O, Weber F. BIOCHEMICAL PATHOLOGY OF VITAMIN B6 DEFICIENCY. Vitam Horm. 1964;22:495-501.

91. Kassab S, Begley P, Church SJ, Rotariu SM, Chevalier-Riffard C, Dowsey AW, et al. Cognitive dysfunction in diabetic rats is prevented by pyridoxamine treatment. A multidisciplinary investigation. Mol Metab. 2019;28:107-19.

92. Jung HY, Kim DW, Nam SM, Kim JW, Chung JY, Won MH, et al. Pyridoxine improves hippocampal cognitive function via increases of serotonin turnover and tyrosine hydroxylase, and its association with CB1 cannabinoid receptor-interacting protein and the CB1 cannabinoid receptor pathway. Biochim Biophys Acta Gen Subj. 2017;1861(12):3142-53.

93. DeTitta GT, Edmonds JW, Stallings W, Donohue J. Molecular structure of biotin.

Results of two independent crystal structure investigations. J Am Chem Soc. 1976;98(7):1920-6.

94. Balamurugan K, Ortiz A, Said HM. Biotin uptake by human intestinal and liver epithelial cells: role of the SMVT system. Am J Physiol Gastrointest Liver Physiol. 2003;285(1):G73-7.

95. Zeng WQ, Al-Yamani E, Acierno JS, Jr., Slaugenhaupt S, Gillis T, MacDonald ME, et al. Biotin-responsive basal ganglia disease maps to 2q36.3 and is due to mutations in SLC19A3. Am J Hum Genet. 2005;77(1):16-26.

96. Gitlin G, Bayer EA, Wilchek M. Studies on the biotin-binding site of avidin. Lysine residues involved in the active site. Biochem J. 1987;242(3):923-6.

97. Tong L. Structure and function of biotin-dependent carboxylases. Cell Mol Life Sci. 2013;70(5):863-91.

98. Wiedmann S, Rodriguez-Melendez R, Ortega-Cuellar D, Zempleni J. Clusters of biotin-responsive genes in human peripheral blood mononuclear cells. J Nutr Biochem. 2004;15(7):433-9.

99. Ingaramo M, Beckett D. Selectivity in post-translational biotin addition to five human carboxylases. J Biol Chem. 2012;287(3):1813-22.

100. Eakin RE, McKinley WA, Williams RJ. EGG-WHITE INJURY IN CHICKS AND ITS RELATIONSHIP TO A DEFICIENCY OF VITAMIN H (BIOTIN). Science. 1940;92(2384):224-5.

101. Stokstad EL, Koch J. Folic acid metabolism. Physiol Rev. 1967;47(1):83-116.

102. Fox JT, Stover PJ. Folate-mediated one-carbon metabolism. Vitam Horm. 2008;79:1-44.

103. Tibbetts AS, Appling DR. Compartmentalization of Mammalian folate-mediated one-carbon metabolism. Annu Rev Nutr. 2010;30:57-81.

104. Ducker GS, Rabinowitz JD. One-Carbon Metabolism in Health and Disease. Cell Metab. 2017;25(1):27-42.

105. Lucock M. Folic acid: nutritional biochemistry, molecular biology, and role in disease processes. Mol Genet Metab. 2000;71(1-2):121-38.

106. Stover PJ, Durga J, Field MS. Folate nutrition and blood-brain barrier dysfunction. Curr Opin Biotechnol. 2017;44:146-52.

107. Lewis CA, Parker SJ, Fiske BP, McCloskey D, Gui DY, Green CR, et al. Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. Mol Cell. 2014;55(2):253-63.

108. Steinberg SE. Mechanisms of folate homeostasis. Am J Physiol. 1984;246(4 Pt 1):G319-24.

109. Green R, Datta Mitra A. Megaloblastic Anemias: Nutritional and Other Causes. Med Clin North Am. 2017;101(2):297-317.

110. Pitkin RM. Folate and neural tube defects. Am J Clin Nutr. 2007;85(1):285s-8s.

111. Blom HJ, Smulders Y. Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. J Inherit Metab Dis. 2011;34(1):75-81.

112. Ueland PM, Refsum H, Beresford SA, Vollset SE. The controversy over homocysteine and cardiovascular risk. Am J Clin Nutr. 2000;72(2):324-32.

113. Mattson MP, Shea TB. Folate and homocysteine metabolism in neural plasticity and neurodegenerative disorders. Trends Neurosci. 2003;26(3):137-46.

114. Sharma M, Tiwari M, Tiwari RK. Hyperhomocysteinemia: Impact on Neurodegenerative Diseases. Basic Clin Pharmacol Toxicol. 2015;117(5):287-96.

115. Shen L, Ji HF. Associations between Homocysteine, Folic Acid, Vitamin B12 and Alzheimer's Disease: Insights from Meta-Analyses. J Alzheimers Dis. 2015;46(3):777-90.

116. Troesch B, Weber P, Mohajeri MH. Potential Links between Impaired One-Carbon Metabolism Due to Polymorphisms, Inadequate B-Vitamin Status, and the Development of Alzheimer's Disease. Nutrients. 2016;8(12).

117. Duthie SJ. Folate and cancer: how DNA damage, repair and methylation impact on colon carcinogenesis. J Inherit Metab Dis. 2011;34(1):101-9.

118. Mason JB. Unraveling the complex relationship between folate and cancer risk. Biofactors. 2011;37(4):253-60.

119. Zhao Y, Guo C, Hu H, Zheng L, Ma J, Jiang L, et al. Folate intake, serum folate levels and esophageal cancer risk: an overall and dose-response meta-analysis. Oncotarget. 2017;8(6):10458-69.

120. Twigt JM, Bezstarosti K, Demmers J, Lindemans J, Laven JS, Steegers-Theunissen RP. Preconception folic acid use influences the follicle fluid proteome. Eur J Clin Invest. 2015;45(8):833-41.

121. Leclerc D, Dejgaard K, Mazur A, Deng L, Wu Q, Nilsson T, et al. Quantitative proteomics reveals differentially expressed proteins in murine preneoplastic intestine in a model of intestinal tumorigenesis induced by low dietary folate and MTHFR deficiency. Proteomics. 2014;14(21-22):2558-65.

122. Gerard N, Chanson-Rolle A, Rock E, Brachet P. Proteomic analysis identifies cytoskeleton-interacting proteins as major downstream targets of altered folate status in the aorta of adult rat. Mol Nutr Food Res. 2014;58(12):2307-19.

123. Chanson A, Sayd T, Rock E, Chambon C, Sante-Lhoutellier V, Potier de Courcy G, et al. Proteomic analysis reveals changes in the liver protein pattern of rats exposed to dietary folate deficiency. J Nutr. 2005;135(11):2524-9.

124. Rattanasinganchan P, Sopitthummakhun K, Doi K, Hu X, Payne DM, Pisitkun T, et al. A folic acid-induced rat model of renal injury to identify biomarkers of tubulointerstitial fibrosis from urinary exosomes. Asian Biomedicine. 2017;10(5):491-502.

125. Duthie SJ, Horgan G, de Roos B, Rucklidge G, Reid M, Duncan G, et al. Blood folate status and expression of proteins involved in immune function, inflammation, and coagulation: biochemical and proteomic changes in the plasma of humans in response to long-term synthetic folic acid supplementation. J Proteome Res. 2010;9(4):1941-50.

126. Xu Y, Tang Y, Li Y. Effect of folic acid on prenatal alcohol-induced modification of brain proteome in mice. Br J Nutr. 2008;99(3):455-61.

127. McNeil CJ, Hay SM, Rucklidge GJ, Reid M, Duncan G, Maloney CA, et al. Disruption of lipid metabolism in the liver of the pregnant rat fed folate-deficient and methyl donor-deficient diets. Br J Nutr. 2008;99(2):262-71.

128. McNeil CJ, Hay SM, Rucklidge GJ, Reid MD, Duncan GJ, Rees WD. Maternal diets deficient in folic acid and related methyl donors modify mechanisms associated with lipid metabolism in the fetal liver of the rat. Br J Nutr. 2009;102(10):1445-52.

129. Lan W, Guhaniyogi J, Horn MJ, Xia JQ, Graham B. A density-based proteomics sample fractionation technology: folate deficiency induced oxidative stress response in liver and brain. J Biomol Tech. 2007;18(4):213-25.

130. Liu J, Yao Y, Yu B, Mao X, Huang Z, Chen D. Effect of maternal folic acid supplementation on hepatic proteome in newborn piglets. Nutrition. 2013;29(1):230-4.

131. Hodgkin DC, Kamper J, Mackay M, Pickworth J, Trueblood KN, White JG. Structure of vitamin B12. Nature. 1956;178(4524):64-6.

132. Seetharam B, Alpers DH. Absorption and transport of cobalamin (vitamin B12). Annu Rev Nutr. 1982;2:343-69.

133. Bauer JA. Synthesis, characterization and nitric oxide release profile of nitrosylcobalamin: a potential chemotherapeutic agent. Anticancer Drugs. 1998;9(3):239-44.

134. Banerjee RV, Matthews RG. Cobalamin-dependent methionine synthase. Faseb j. 1990;4(5):1450-9.

135. Thomä NH, Leadlay PF. Mechanistic and structural studies on methylmalonyl-CoA mutase. Biochem Soc Trans. 1998;26(3):293-8.

136. Nixon PF, Bertino JR. Interrelationships of vitamin B12 and folate in man. Am J Med. 1970;48(5):555-61.

137. Lindenbaum J, Healton EB, Savage DG, Brust JC, Garrett TJ, Podell ER, et al. Neuropsychiatric disorders caused by cobalamin deficiency in the absence of anemia or macrocytosis. N Engl J Med. 1988;318(26):1720-8.

138. Caterino M, Pastore A, Strozziero MG, Di Giovamberardino G, Imperlini E, Scolamiero E, et al. The proteome of cblC defect: in vivo elucidation of altered cellular pathways in humans. J Inherit Metab Dis. 2015;38(5):969-79.

139. Gianazza E, Veber D, Eberini I, Buccellato FR, Mutti E, Sironi L, et al. Cobalamin (vitamin B12)-deficiency-induced changes in the proteome of rat cerebrospinal fluid. Biochem J. 2003;374(Pt 1):239-46.

140. Ahmad S, Basak T, Anand Kumar K, Bhardwaj G, Lalitha A, Yadav DK, et al. Maternal micronutrient deficiency leads to alteration in the kidney proteome in rat pups. J Proteomics. 2015;127(Pt A):178-84.

141. Ahmad S, Kumar KA, Basak T, Bhardwaj G, Yadav DK, Lalitha A, et al. PPAR signaling pathway is a key modulator of liver proteome in pups born to vitamin B(12) deficient rats. J Proteomics. 2013;91:297-308.

142. Bürzle M, Suzuki Y, Ackermann D, Miyazaki H, Maeda N, Clémençon B, et al. The sodium-dependent ascorbic acid transporter family SLC23. Mol Aspects Med. 2013;34(2-3):436-54.

143. Wilson JX. Regulation of vitamin C transport. Annu Rev Nutr. 2005;25:105-25.

144. Winkler BS, Orselli SM, Rex TS. The redox couple between glutathione and ascorbic acid: a chemical and physiological perspective. Free Radic Biol Med. 1994;17(4):333-49.

145. May JM, Mendiratta S, Hill KE, Burk RF. Reduction of dehydroascorbate to ascorbate by the selenoenzyme thioredoxin reductase. J Biol Chem. 1997;272(36):22607-10.

146. Myllylä R, Majamaa K, Günzler V, Hanauske-Abel HM, Kivirikko KI. Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. J Biol Chem. 1984;259(9):5403-5.

147. de Jong L, Albracht SP, Kemp A. Prolyl 4-hydroxylase activity in relation to the oxidation state of enzyme-bound iron. The role of ascorbate in peptidyl proline hydroxylation. Biochim Biophys Acta. 1982;704(2):326-32.

148. Fong GH, Takeda K. Role and regulation of prolyl hydroxylase domain proteins. Cell Death Differ. 2008;15(4):635-41.

149. Gjaltema RA, Bank RA. Molecular insights into prolyl and lysyl hydroxylation of fibrillar collagens in health and disease. Crit Rev Biochem Mol Biol. 2017;52(1):74-95.

150. Passoja K, Rautavuoma K, Ala-Kokko L, Kosonen T, Kivirikko KI. Cloning and characterization of a third human lysyl hydroxylase isoform. Proc Natl Acad Sci U S A. 1998;95(18):10482-6.

151. Henderson LM, Nelson PJ, Henderson L. Mammalian enzymes of trimethyllysine conversion to trimethylaminobutyrate. Fed Proc. 1982;41(12):2843-7.

152. Longo N, Frigeni M, Pasquali M. Carnitine transport and fatty acid oxidation. Biochim Biophys Acta. 2016;1863(10):2422-35.

153. May JM, Qu ZC, Meredith ME. Mechanisms of ascorbic acid stimulation of norepinephrine synthesis in neuronal cells. Biochem Biophys Res Commun. 2012;426(1):148-52.

154. Meister A. On the antioxidant effects of ascorbic acid and glutathione. Biochem Pharmacol. 1992;44(10):1905-15.

155. Hemilä H. Vitamin C and Infections. Nutrients. 2017;9(4).

156. Cudjoe EK, Jr., Hassan ZH, Kang L, Reynolds PS, Fisher BJ, McCarter J, et al. Temporal map of the pig polytrauma plasma proteome with fluid resuscitation and intravenous vitamin C treatment. J Thromb Haemost. 2019;17(11):1827-37.

157. Aumailley L, Roux-Dalvai F, Kelly I, Droit A, Lebel M. Vitamin C alters the amount of specific endoplasmic reticulum associated proteins involved in lipid metabolism in the liver of mice synthesizing a nonfunctional Werner syndrome (Wrn) mutant protein. PLoS One. 2018;13(3):e0193170.

158. Motorykin I, Traber MG, Tanguay RL, Maier CS. Proteome-driven elucidation of adaptive responses to combined vitamin E and C deficiency in zebrafish. J Proteome Res. 2014;13(3):1647-56.

159. Da Costa LA, Garcia-Bailo B, Borchers CH, Badawi A, El-Sohemy A. Association between the plasma proteome and serum ascorbic acid concentrations in humans. J Nutr Biochem. 2013;24(5):842-7.

160. Park S, Ahn ES, Lee S, Jung M, Park JH, Yi SY, et al. Proteomic analysis reveals upregulation of RKIP in S-180 implanted BALB/C mouse after treatment with ascorbic acid. J Cell Biochem. 2009;106(6):1136-45.

161. Lee J, Lee G, Park JH, Lee S, Yeom CH, Na B, et al. Proteomic analysis of tumor tissue in CT-26 implanted BALB/C mouse after treatment with ascorbic acid. Cell Mol Biol Lett. 2012;17(1):62-76.

162. Weissinger EM, Nguyen-Khoa T, Fumeron C, Saltiel C, Walden M, Kaiser T, et al. Effects of oral vitamin C supplementation in hemodialysis patients: a proteomic assessment. Proteomics. 2006;6(3):993-1000.

163. Henry HL, Norman AW. Vitamin D: metabolism and biological actions. Annu Rev Nutr. 1984;4:493-520.

164. Bikle DD. Vitamin D metabolism, mechanism of action, and clinical applications. Chem Biol. 2014;21(3):319-29.

165. Glossmann HH. Origin of 7-dehydrocholesterol (provitamin D) in the skin. J Invest Dermatol. 2010;130(8):2139-41.

166. Jones G, Prosser DE, Kaufmann M. Cytochrome P450-mediated metabolism of vitamin D. J Lipid Res. 2014;55(1):13-31.

167. White P, Cooke N. The multifunctional properties and characteristics of vitamin D-binding protein. Trends Endocrinol Metab. 2000;11(8):320-7.

168. Holick MF, DeLuca HF. Chemistry and biological activity of vitamin D, its metabolites and analogs. Adv Steroid Biochem Pharmacol. 1974;4(0):111-55.

169. Haussler MR, Jurutka PW, Mizwicki M, Norman AW. Vitamin D receptor (VDR)-

mediated actions of 1alpha,25(OH)(2)vitamin D(3): genomic and non-genomic mechanisms. Best Pract Res Clin Endocrinol Metab. 2011;25(4):543-59.

170. Lips P. Vitamin D physiology. Prog Biophys Mol Biol. 2006;92(1):4-8.

171. Goltzman D. Functions of vitamin D in bone. Histochem Cell Biol. 2018;149(4):305-12.

172. Nemere I, Farach-Carson MC, Rohe B, Sterling TM, Norman AW, Boyan BD, et al. Ribozyme knockdown functionally links a 1,25(OH)2D3 membrane binding protein (1,25D3-MARRS) and phosphate uptake in intestinal cells. Proc Natl Acad Sci U S A. 2004;101(19):7392-7.

173. Hii CS, Ferrante A. The Non-Genomic Actions of Vitamin D. Nutrients. 2016;8(3):135.

174. Holick MF. Vitamin D deficiency. N Engl J Med. 2007;357(3):266-81.

175. Faridi KF, Lupton JR, Martin SS, Banach M, Quispe R, Kulkarni K, et al. Vitamin D deficiency and non-lipid biomarkers of cardiovascular risk. Arch Med Sci. 2017;13(4):732-7.

176. Kheiri B, Abdalla A, Osman M, Ahmed S, Hassan M, Bachuwa G. Vitamin D deficiency and risk of cardiovascular diseases: a narrative review. Clin Hypertens. 2018;24:9.

177. Dimitrakopoulou VI, Tsilidis KK, Haycock PC, Dimou NL, Al-Dabhani K, Martin RM, et al. Circulating vitamin D concentration and risk of seven cancers: Mendelian randomisation study. Bmj. 2017;359:j4761.

178. Hossain S, Beydoun MA, Beydoun HA, Chen X, Zonderman AB, Wood RJ. Vitamin D and breast cancer: A systematic review and meta-analysis of observational studies. Clin Nutr ESPEN. 2019;30:170-84.

179. Littlejohns TJ, Henley WE, Lang IA, Annweiler C, Beauchet O, Chaves PH, et al. Vitamin D and the risk of dementia and Alzheimer disease. Neurology. 2014;83(10):920-8.

180. Moretti R, Morelli ME, Caruso P. Vitamin D in Neurological Diseases: A Rationale for a Pathogenic Impact. Int J Mol Sci. 2018;19(8).

181. Ishikawa LLW, Colavite PM, Fraga-Silva TFC, Mimura LAN, Franca TGD, Zorzella-Pezavento SFG, et al. Vitamin D Deficiency and Rheumatoid Arthritis. Clin Rev Allergy Immunol. 2017;52(3):373-88.

182. Pierrot-Deseilligny C, Souberbielle JC. Vitamin D and multiple sclerosis: An update. Mult Scler Relat Disord. 2017;14:35-45.

183. Chen L, Wilson R, Bennett E, Zosky GR. Identification of vitamin D sensitive pathways during lung development. Respir Res. 2016;17:47.

184. Oveland E, Nystad A, Berven F, Myhr KM, Torkildsen O, Wergeland S. 1,25-Dihydroxyvitamin-D3 induces brain proteomic changes in cuprizone mice during remyelination involving calcium proteins. Neurochem Int. 2018;112:267-77.

185. Al-Daghri NM, Alokail MS, Manousopoulou A, Heinson A, Al-Attas O, Al-Saleh Y, et al. Sex-specific vitamin D effects on blood coagulation among overweight adults. Eur J Clin Invest. 2016;46(12):1031-40.

186. Walker GE, Ricotti R, Roccio M, Moia S, Bellone S, Prodam F, et al. Pediatric obesity and vitamin D deficiency: a proteomic approach identifies multimeric adiponectin as a key link between these conditions. PLoS One. 2014;9(1):e83685.

187. Walker GE, Follenzi A, Bruscaggin V, Manfredi M, Bellone S, Marengo E, et al. Fetuin B links vitamin D deficiency and pediatric obesity: Direct negative regulation by vitamin D. J Steroid Biochem Mol Biol. 2018;182:37-49.

188. Jung SH, Won KJ, Lee KP, Kim HJ, Seo EH, Lee HM, et al. The serum protein fetuin-B is involved in the development of acute myocardial infarction. Clin Sci (Lond). 2015;129(1):27-38.

189. Wang QQ, Zhao X, Pu XP. Proteome analysis of the left ventricle in the vitamin D(3) and nicotine-induced rat vascular calcification model. J Proteomics. 2011;74(4):480-9.

190. Almofti MR, Huang Z, Yang P, Rui Y, Yang P. Proteomic analysis of rat aorta during atherosclerosis induced by high cholesterol diet and injection of vitamin D3. Clin Exp Pharmacol Physiol. 2006;33(4):305-9.

191. Almeras L, Eyles D, Benech P, Laffite D, Villard C, Patatian A, et al. Developmental vitamin D deficiency alters brain protein expression in the adult rat: implications for neuropsychiatric disorders. Proteomics. 2007;7(5):769-80.

192. Keeney JTR, Forster S, Sultana R, Brewer LD, Latimer CS, Cai J, et al. Dietary vitamin D deficiency in rats from middle to old age leads to elevated tyrosine nitration and proteomics changes in levels of key proteins in brain: implications for low vitamin D-dependent age-related cognitive decline. Free Radic Biol Med. 2013;65:324-34.

193. West KP, Jr., Cole RN, Shrestha S, Schulze KJ, Lee SE, Betz J, et al. A Plasma

alpha-Tocopherome Can Be Identified from Proteins Associated with Vitamin E Status in School-Aged Children of Nepal. J Nutr. 2015;145(12):2646-56.

194. Hamezah HS, Durani LW, Yanagisawa D, Ibrahim NF, Aizat WM, Makpol S, et al. Modulation of Proteome Profile in AbetaPP/PS1 Mice Hippocampus, Medial Prefrontal Cortex, and Striatum by Palm Oil Derived Tocotrienol-Rich Fraction. J Alzheimers Dis. 2019;72(1):229-46.

195. Lee SE, Schulze KJ, Cole RN, Wu LS, Yager JD, Groopman J, et al. Biological Systems of Vitamin K: A Plasma Nutriproteomics Study of Subclinical Vitamin K Deficiency in 500 Nepalese Children. OMICS. 2016;20(4):214-23.

196. Herrera E, Barbas C. Vitamin E: action, metabolism and perspectives. J Physiol Biochem. 2001;57(1):43-56.

197. Kamal-Eldin A, Appelqvist LA. The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids. 1996;31(7):671-701.

198. Brigelius-Flohe R. Bioactivity of vitamin E. Nutr Res Rev. 2006;19(2):174-86.

199. Zingg JM. Vitamin E: A Role in Signal Transduction. Annu Rev Nutr. 2015;35:135-73.

200. Schneider C. Chemistry and biology of vitamin E. Mol Nutr Food Res. 2005;49(1):7-30.

201. Di Donato I, Bianchi S, Federico A. Ataxia with vitamin E deficiency: update of molecular diagnosis. Neurol Sci. 2010;31(4):511-5.

202. Sitrin MD, Lieberman F, Jensen WE, Noronha A, Milburn C, Addington W. Vitamin E deficiency and neurologic disease in adults with cystic fibrosis. Ann Intern Med. 1987;107(1):51-4.

203. Gomez-Pomar E, Hatfield E, Garlitz K, Westgate PM, Bada HS. Vitamin E in the Preterm Infant: A Forgotten Cause of Hemolytic Anemia. Am J Perinatol. 2018;35(3):305-10.

204. Muller DP. Vitamin E and neurological function. Mol Nutr Food Res. 2010;54(5):710-8.

205. Ulatowski LM, Manor D. Vitamin E and neurodegeneration. Neurobiol Dis. 2015;84:78-83.

206. Kaga E, Karademir B, Baykal AT, Ozer NK. Identification of differentially expressed proteins in atherosclerotic aorta and effect of vitamin E. J Proteomics.

2013;92:260-73.

207. Panachan J, Chokchaichamnankit D, Weeraphan C, Srisomsap C, Masaratana P, Hatairaktham S, et al. Differentially expressed plasma proteins of beta-thalassemia/hemoglobin E patients in response to curcuminoids/vitamin E antioxidant cocktails. Hematology. 2019;24(1):300-7.

208. Aldred S, Sozzi T, Mudway I, Grant MM, Neubert H, Kelly FJ, et al. Alpha tocopherol supplementation elevates plasma apolipoprotein A1 isoforms in normal healthy subjects. Proteomics. 2006;6(5):1695-703.

209. Naderi M, Keyvanshokooh S, Salati AP, Ghaedi A. Proteomic analysis of liver tissue from rainbow trout (Oncorhynchus mykiss) under high rearing density after administration of dietary vitamin E and selenium nanoparticles. Comp Biochem Physiol Part D Genomics Proteomics. 2017;22:10-9.

210. Kim J, Sun P, Lam YW, Troncoso P, Sabichi AL, Babaian RJ, et al. Changes in serum proteomic patterns by presurgical alpha-tocopherol and L-selenomethionine supplementation in prostate cancer. Cancer Epidemiol Biomarkers Prev. 2005;14(7):1697-702.

211. Shearer MJ. Vitamin K. Lancet. 1995;345(8944):229-34.

212. Stafford DW. The vitamin K cycle. J Thromb Haemost. 2005;3(8):1873-8.

213. Berkner KL, Runge KW. The physiology of vitamin K nutriture and vitamin K-dependent protein function in atherosclerosis. J Thromb Haemost. 2004;2(12):2118-32.

214. Shearer MJ. Vitamin K deficiency bleeding (VKDB) in early infancy. Blood Rev. 2009;23(2):49-59.

215. Greer FR. Vitamin K deficiency and hemorrhage in infancy. Clin Perinatol. 1995;22(3):759-77.

216. Johansson HJ, Socciarelli F, Vacanti NM, Haugen MH, Zhu Y, Siavelis I, et al. Breast cancer quantitative proteome and proteogenomic landscape. Nat Commun. 2019;10(1):1600.

217. Zhu Y, Orre LM, Johansson HJ, Huss M, Boekel J, Vesterlund M, et al. Discovery of coding regions in the human genome by integrated proteogenomics analysis workflow. Nat Commun. 2018;9(1):903.

218. Müller JB, Geyer PE, Colaço AR, Treit PV, Strauss MT, Oroshi M, et al. The proteome landscape of the kingdoms of life. Nature. 2020;582(7813):592-6.

219. Meier F, Geyer PE, Virreira Winter S, Cox J, Mann M. BoxCar acquisition method enables single-shot proteomics at a depth of 10,000 proteins in 100 minutes. Nat Methods. 2018;15(6):440-8.

220. Lundby A, Lage K, Weinert BT, Bekker-Jensen DB, Secher A, Skovgaard T, et al. Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. Cell Rep. 2012;2(2):419-31.

221. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. Mol Cell Proteomics. 2007;6(12):2212-29.

222. Pernemalm M, Sandberg A, Zhu Y, Boekel J, Tamburro D, Schwenk JM, et al. In-depth human plasma proteome analysis captures tissue proteins and transfer of protein variants across the placenta. Elife. 2019;8.

223. Hogrebe A, von Stechow L, Bekker-Jensen DB, Weinert BT, Kelstrup CD, Olsen JV. Benchmarking common quantification strategies for large-scale phosphoproteomics. Nat Commun. 2018;9(1):1045.

224. Panizza E, Branca RMM, Oliviusson P, Orre LM, Lehtio J. Isoelectric point-based fractionation by HiRIEF coupled to LC-MS allows for in-depth quantitative analysis of the phosphoproteome. Sci Rep. 2017;7(1):4513.

225. Orre LM, Vesterlund M, Pan Y, Arslan T, Zhu Y, Fernandez Woodbridge A, et al. SubCellBarCode: Proteome-wide Mapping of Protein Localization and Relocalization. Mol Cell. 2019;73(1):166-82 e7.

226. Tamborero D, Dienstmann R, Rachid MH, Boekel J, Baird R, Brana I, et al. Support systems to guide clinical decision-making in precision oncology: The Cancer Core Europe Molecular Tumor Board Portal. Nat Med. 2020;26(7):992-4.

227. Vacanti NM. The Fundamentals of Constructing and Interpreting Heat Maps. Methods Mol Biol. 2019;1862:279-91.

228. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128.

229. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1):W90-7.

230. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of

large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57.

231. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13.

232. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.

233. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. PLoS Comput Biol. 2007;3(3):e39.

234. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.

235. Gropper SS, Smith JL, Groff JL. Advanced Nutrition and Human Metabolism. 6th ed: Cengage Learning; 2012.

236. Michal G, Schomburg D. Biochemical Pathways. 2nd ed: Wiley; 2012.

237. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30.

Chapter 3

Correcting for Naturally Occurring Mass Isotopologue Abundances in Stable-Isotope Tracing Experiments with PolyMID*

* Jeong H, Yu Y, Johansson HJ, Schroeder FC, Lehtiö J, Vacanti NM. Correcting for naturally occurring mass isotopologue abundances in stable-isotope tracing experiments with polymid. *Metabolites*. 2021 May 12;11(5):310.

Abstract

Stable-isotope tracing is a method to measure intracellular metabolic pathway utilization by feeding a cellular system a stable-isotope-labeled tracer nutrient. The power of the method to resolve differential pathway utilization is derived from the enrichment of metabolites in heavy isotopes that are synthesized from the tracer nutrient. However, the readout is complicated by the presence of naturally occurring heavy isotopes that are not derived from the tracer nutrient. Herein we present an algorithm, and a tool that applies it (PolyMID-Correct, part of the PolyMID software package), to computationally remove the influence of naturally occurring heavy isotopes. The algorithm is applicable to stable-isotope tracing data collected on low- and high- mass resolution mass spectrometers. PolyMID-Correct, part of the PolyMID software package, is open source and available under an MIT license.

Introduction

Utilization of intracellular metabolic pathways can be assayed by tracing with stable-isotope-labeled nutrients. Cells cultured in medium containing a stable-isotope-labeled nutrient, i.e. a tracer, internalize and metabolize it; allowing the heavy isotope(s) to be incorporated in metabolic pathway reactants, products, and intermediates. Heavy isotope presence in metabolites can then be assayed by mass spectrometry and used to infer pathway utilization by quantitative modelling using metabolic flux analysis (MFA) or manual inspection of mass isotopologue distributions (MIDs, Figures 3.1A-B).

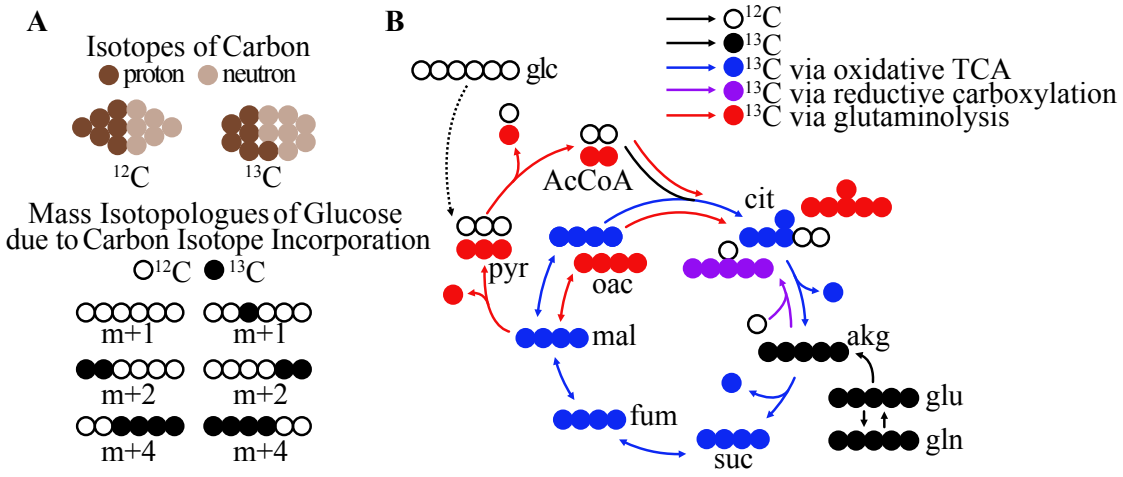


Figure 3.1. Principles of Stable-Isotope Tracing. A) Isotopes of an element are atoms of that element with different masses due to the number of neutrons they contain. Each element has a characteristic isotope mass distribution (IMD) describing the proportion of the atoms having the nominal mass (m) and those greater than nominal mass by n integer atomic mass units (m+n). Mass isotopologues of a compound are molecules of that compound with different masses due to the number of heavy isotopes they contain. Each compound has a mass isotopologue distribution (MID) describing the proportion of molecules having the monoisotopic mass (m) and those greater than the monoisotopic mass by n integer atomic mass units (m+n). Molecules of mass m have no extra neutrons due to heavy isotope incorporation while those of mass m+n have n extra neutrons due to heavy isotope incorporation. B) Schematic of possible TCA cycle metabolite labeling patterns due to incorporation of ^{13}C from a [U-$^{13}C_5$]glutamine tracer.

If data is incorporated into an MFA model, it is done as either time-course or static (steady-state) measurements where the set of metabolic fluxes that minimizes the sum of squared differences between model-computed and measured values is returned. However, MFA models are generally constructed taking great liberties in making simplifying assumptions of biochemical networks; limiting their representation of cellular metabolism. Thus, manual interpretation is generally performed in lieu of the burdensome step of fitting an MFA model to measured MIDs.

As an illustration of flux inference by manual inspection, consider a glutamine tracer with all five carbons labeled as ^{13}C, denoted [U-$^{13}C_5$]glutamine, that is metabolized by enzymes of the tricarboxylic acid (TCA) cycle and associated anaplerotic pathways.

Examination of label incorporation on citrate can then distinguish relative utilization of the oxidative tricarboxylic acid (TCA) cycle, reductive carboxylation, and glutaminolysis. If via the oxidative TCA cycle, [U-$^{13}C_5$]glutamine will label citrate with four heavy carbon atoms; if via the reductive carboxylation, [U-$^{13}C_5$]glutamine will label citrate with five heavy carbon atoms; and if via glutaminolysis, [U-$^{13}C_5$]glutamine will label citrate with six heavy carbon atoms (Figure 3.1B).

However, inference of pathway utilization by manual inspection of metabolite MIDs is confounded by incorporation of naturally occurring heavy isotopes. Posed as a question related to the above described example using a [U-$^{13}C_5$]glutamine tracer: How does the researcher know if citrate with five heavy isotopes derives them all from the tracer or derives four from the tracer and one due to incorporation of naturally present ^{13}C, ^{2}H, or ^{17}O? Whether synthesis of the citrate molecule in question is attributed to oxidative TCA cycle flux or reductive carboxylation is dependent on distinguishing these two scenarios (Figure 3.1B). Note that [U-$^{13}C_5$]glutamine labeling citrate with five heavy carbon atoms via simultaneous flux through glutaminolysis and pyruvate carboxylase is not considered in this example for simplicity.

This and analogous questions can be answered by post-processing of measured mass isotopologue distributions in a manner that removes the influence of naturally occurring heavy isotopes. Herein, we present the PolyMID-Correct tool, part of the PolyMID software package, that relies on known distributions of naturally occurring atomic isotopes (1) and applies the principles of polynomial expansion to correct stable-isotope tracing data for the presence of naturally occurring heavy isotopes.

PolyMID-Correct is open source and written in Python 3. Information about the

metabolite, including its measured MID, is passed to this tool within a user-created text file or via the command line. The user can specify whether mass differences due to naturally occurring isotopes of specified atoms are resolved from mass differences due to the tracer atom, i.e. if the data is high mass resolution; or if MIDs are inclusive of all heavy isotopes, i.e. if the data is low mass resolution (henceforth referred to as high- and low- resolution). The identity of the tracer atom is also specified by the user along with the purity of the tracer molecule and the atom enrichment of the heavy-isotope label on the tracer.

Other programs that are freely available to correct for naturally occurring isotopes in high- and low- resolution data include AccuCor (1), IsoCor (2), IsoCorrectoR (3), and PyNAC (4). However, PolyMID-Correct is the only of these that allows users to specify input parameters programmatically; allowing for smooth integration into data processing pipelines. PolyMID-Correct is also the only option that allows users to specify the elemental identity of atoms whose heavy isotope incorporation into metabolites is resolved from incorporation of heavy isotopes of the tracer element. AccuCor and IsoCor make this distinction based on the user-input instrument resolution whereas IsoCorrectoR and PyNAC either allow all or no other elemental heavy isotopes to be distinguished from the tracer atoms. PolyMID-Correct also includes the option for data to be input from a single, simply-formatted, text file; providing a shallower learning curve for command-line novices. PolyMID-Correct, IsoCor, and PyNAC are implemented in Python whereas AccuCor and IsoCorrectoR are implemented in R.

Results

General Correction Algorithm

PolyMID-Correct transforms a measured MID of a metabolite into a set of component MIDs. Each component MID is a theoretical MID of the metabolite where a known number of heavy tracer atoms are incorporated. Consider a [U-$^{13}C_6$]glucose tracing experiment where ^{13}C incorporation on the amino acid serine is being examined. Theoretically, serine can be labeled with zero, one, two, or three heavy carbon atoms from [U-$^{13}C_6$]glucose. The portion of serine molecules containing zero ^{13}C atoms derived from [U-$^{13}C_6$]glucose will have an MID consistent with naturally occurring isotope incorporation. The portion of serine containing one ^{13}C atom derived from [U-$^{13}C_6$]glucose will also have an MID consistent with naturally occurring isotope incorporation with the exception that one of the carbon atoms is guaranteed to be labeled as ^{13}C. Analogously, the portions of serine containing two or three ^{13}C atom derived from [U-$^{13}C_6$]glucose will have MIDs consistent with naturally occurring isotope incorporation with the exceptions that, respectively, two or three of the carbon atoms are guaranteed to be labeled as ^{13}C. The MIDs accounting for natural isotope incorporation of serine guaranteed to contain zero, one, two, or three ^{13}C atoms derived from [U-$^{13}C_6$]glucose can be computed theoretically, as described in a subsequent section. Thus, the measured MID of serine can be formulated as Equation 1 where the weights of each of the component vectors are the components of the corrected MID.

$$\begin{bmatrix} S_{m,M0} \\ S_{m,M1} \\ S_{m,M2} \\ S_{m,M3} \\ S_{m,M4} \\ S_{m,M5} \end{bmatrix} = S_{c,M0} * \begin{bmatrix} S_{0L,M0} \\ S_{0L,M1} \\ S_{0L,M2} \\ S_{0L,M3} \\ S_{0L,M4} \\ S_{0L,M5} \end{bmatrix} + S_{c,M1} * \begin{bmatrix} S_{1L,M0} \\ S_{1L,M1} \\ S_{1L,M2} \\ S_{1L,M3} \\ S_{1L,M4} \\ S_{1L,M5} \end{bmatrix} + S_{c,M2} * \begin{bmatrix} S_{2L,M0} \\ S_{2L,M1} \\ S_{2L,M2} \\ S_{2L,M3} \\ S_{2L,M4} \\ S_{2L,M5} \end{bmatrix} + S_{c,M3} * \begin{bmatrix} S_{3L,M0} \\ S_{3L,M1} \\ S_{3L,M2} \\ S_{3L,M3} \\ S_{3L,M4} \\ S_{3L,M5} \end{bmatrix}$$

(Eq. 1)

where S indicates an MID component of serine, the subscript m indicates a measured value, the subscript c indicates a value corrected for incorporation of naturally occurring isotopes, the subscript L (for *labeled*) preceded by a number indicates a molecule of serine guaranteed to have that number of labeled carbon atoms, and the subscript M followed by a number indicates the MID component corresponds to that number of atomic mass units above the monoisotopic mass of serine. As examples; $S_{m,M0}$ is the measured MID component of serine that is zero atomic mass units above the monoisotopic mass, $S_{1L,M2}$ is the theoretical MID component two atomic mass units above the monoisotopic mass of a serine molecule guaranteed to have one carbon derived from [U-$^{13}C_6$]glucose, and $S_{c,M1}$ is the corrected MID component of serine that is one atomic mass unit above the monoisotopic mass.

Equation 1 can be rewritten as Equation 2 because the multiplication of a matrix by a column vector is equivalent to weighting the columns of the matrix by the components of the vector (5).

$$\begin{bmatrix} S_{m,M0} \\ S_{m,M1} \\ S_{m,M2} \\ S_{m,M3} \\ S_{m,M4} \\ S_{m,M5} \end{bmatrix} = \begin{bmatrix} S_{0L,M0} & S_{1L,M0} & S_{2L,M0} & S_{3L,M0} \\ S_{0L,M1} & S_{1L,M1} & S_{2L,M1} & S_{3L,M1} \\ S_{0L,M2} & S_{1L,M2} & S_{2L,M2} & S_{3L,M2} \\ S_{0L,M3} & S_{1L,M3} & S_{2L,M3} & S_{3L,M3} \\ S_{0L,M4} & S_{1L,M4} & S_{2L,M4} & S_{3L,M4} \\ S_{0L,M5} & S_{1L,M5} & S_{2L,M5} & S_{3L,M5} \end{bmatrix} * \begin{bmatrix} S_{c,M0} \\ S_{c,M1} \\ S_{c,M2} \\ S_{c,M3} \end{bmatrix} \quad \text{(Eq. 2)}$$

This elegant formulation, proposed by Brunengraber and colleagues (6), collects the corrected MID components as a vector and maps it to the measured MID by multiplication of what is termed a correction matrix. Equation 2 is written in succinct matrix form as Equation 3.

$$\overline{S_m} = \overline{\overline{CM}} * \overline{S_c} \quad \text{(Eq. 3)}$$

where $\overline{S_m}$ is the measured MID of serine, $\overline{\overline{CM}}$ is the correction matrix for serine in a ^{13}C tracing experiment, and $\overline{S_c}$ is the MID of serine corrected for abundances of naturally occurring isotopes. Thus, the MID corrected for abundances of naturally occurring isotopes is given by Equation 4.

$$\overline{S_c} = \overline{\overline{CM}}^{+} * \overline{S_m} \quad \text{(Eq. 4)}$$

where $\overline{\overline{CM}}^{+}$ is the Moore-Penrose inverse of the correction matrix. The above example is worked through for a ^{13}C tracer and serine. However, it can be generalized to any tracer and metabolite by formation of analogous component vectors in Equation 1. PolyMID-Correct does not allow the corrected MID to have more components than physically possible, e.g. $\overline{S_c}$ can only have four components in a carbon-tracing experiment because serine only has three carbon atoms. However, $\overline{S_m}$ may have more than four components due to naturally occurring heavy isotopes of other atoms or those on derivatization reagents. This leads to situations where the correction matrix

is not square and Equation 4 is overdetermined. Thus, Equation 4 yields a best fit for $\overline{S_c}$ with associated residuals.

The Correction Matrix

As determined in the previous section, the component vectors of Equation 1 are the columns of the correction matrix. Herein, the algorithm developed for and applied by PolyMID-Correct to compute the columns of the correction matrix is presented. In contrast to earlier efforts (7, 8), it considers the impacts on the computation of columns in the correction matrix of substituting tracer atoms for naturally occurring atoms in the chemical formulas of the molecules represented by those columns.

Each component vector in Equation 1 is an MID of serine, where serine is guaranteed to contain a specified number of ^{13}C atoms from [U-$^{13}C_6$]glucose. The first component vector corresponds to serine with zero ^{13}C atoms derived from the tracer, the second component vector corresponds to serine with one ^{13}C atom derived from [U-$^{13}C_6$]glucose, and so on. If the MID of an intact and unmodified serine has been measured, its chemical formula is: $C_3H_7NO_3$. Each carbon on serine is capable of being labeled as ^{13}C from the tracer. The first component vector has no ^{13}C atoms derived from the tracer, so it is simply the MID of serine found in nature. The components of this vector can be computed from the isotope mass distributions (IMDs) of each atom (1), specified as Equation 5. Note the distinction between an IMD and an MID as illustrated in Figure 3.1A.

$$IMD_C = [C_{M0} \quad C_{M1} \quad C_{M2}]$$

$$IMD_H = [H_{M0} \quad H_{M1} \quad H_{M2}] \qquad \text{(Eq. 5)}$$

$$IMD_N = [N_{M0} \quad N_{M1} \quad N_{M2}]$$

$$IMD_O = [O_{M0} \quad O_{M1} \quad O_{M2}]$$

where C, H, N, and O represent the elements for which they are chemical symbols. The subscript M followed by a number indicates the relative abundance of an isotope that is that number of atomic mass units heavier than the nominal mass. These values are known constants. To compute the MID of serine ($C_3H_7NO_3$), the IMDs of the atoms are multiplied together as if they were polynomials (9) as in Equation 6.

$$MID_{C3H7NO3} = IMD_C{}^3 * IMD_H{}^7 * IMD_N * IMD_O{}^3 \quad \text{(Eq.6)}$$

$$MID_{C3H7NO3} =$$

$$[C_{M0} \quad C_{M1} \quad C_{M2}]^3[H_{M0} \quad H_{M1} \quad H_{M2}]^7[N_{M0} \quad N_{M1} \quad N_{M2}][O_{M0} \quad O_{M1} \quad O_{M2}]^3$$

Terms that have the same collective mass are grouped as illustrated for a simple example molecule, C_2, in Equation 7.

$$MID_{C2} = IMD_C{}^2$$

$$MID_{C2} = [C_{M0} \quad C_{M1} \quad C_{M2}] * [C_{M0} \quad C_{M1} \quad C_{M2}] \quad \text{(Eq. 7)}$$

$$MID_{C2} = [C_{M0}C_{M0} \quad 2C_{M0}C_{M1} \quad (2C_{M0}C_{M2} + C_{M1}C_{M1}) \quad 2C_{M1}C_{M2} \quad C_{M2}C_{M2}]$$

The second column of the correction matrix in Equation 2 is the MID of a serine molecule as it is found in nature, but with one carbon replaced by a labeled carbon atom derived from the tracer, C_L. Thus, the second column of the correction matrix is the MID of the molecule $C_LC_2H_7NO_3$, where the IMD of C_L is considered to be 100% enriched for isotopes one atomic mass unit above the nominal mass, as per the definition of a labeled carbon atom.

$$IMD_{C_L} = [C_{L,M0} \quad C_{L,M1} \quad C_{L,M2}] \quad \text{(Eq.8)}$$

$$IMD_{C_L} = [0 \quad 1 \quad 0]$$

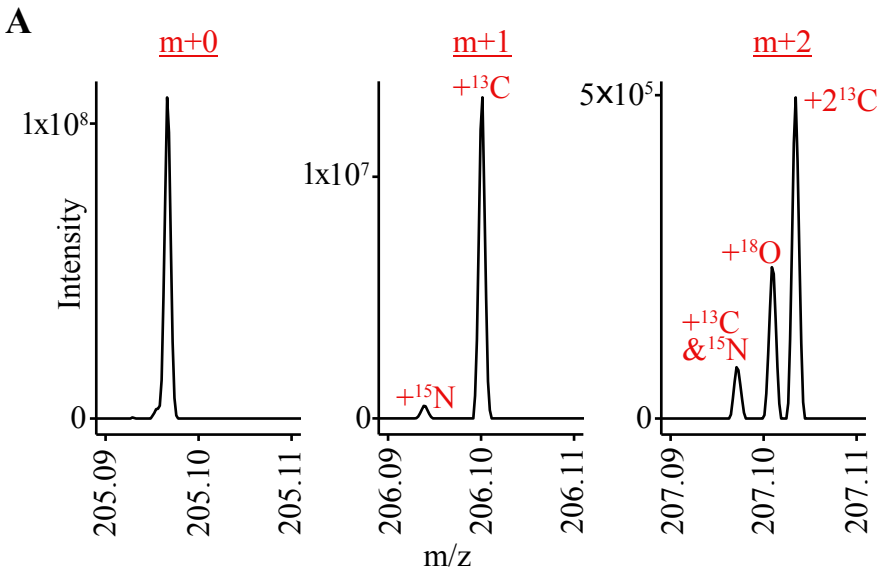
where IMD_{C_L} is the isotope mass distribution of a labeled carbon atom and $C_{L,M0}$ is the component of IMD_{C_L} corresponding to 0 atomic mass units above the nominal mass of carbon (as indicated by the subscript $M0$).

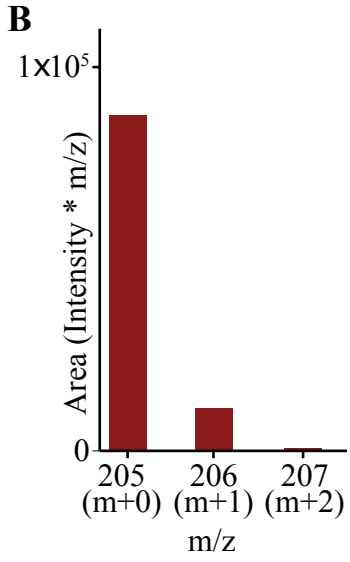
Values within the remaining columns of the correction matrix are computed analogously, thus providing all information required to apply Equation 4 to correct for incorporation of naturally occurring heavy isotopes in a serine MID as measured in a ^{13}C tracing experiment. This algorithm can be generalized to compute the correction matrix corresponding to any metabolite-tracer combination. For larger metabolites with tens of atoms that can acquire label from the tracer, such as lipids in a carbon tracing experiment, many expansions of long polynomials must be performed to compute the columns of the correction matrix; thus PolyMID-Correct may require a few minutes to compute their corrected MIDs.

High Resolution Data

The distinction between high- and low- resolution mass spectrometer data in stable-isotope tracing is whether the mass differences due to incorporation of heavy isotopes of tracer element can be distinguished from incorporation of those of other elements. This is illustrated in the mass spectrographs displayed in Figures 3.2A-B for a molecule of tryptophan. The spectrograph in Figure 3.2A is collected from a high-resolution Orbitrap instrument capable of resolving mass differences due to incorporation of heavy isotopes of carbon, nitrogen, oxygen, and hydrogen. The data in Figure 3.2B is a representation of the same data as it would be measured on a low-

resolution instrument, i.e. one that is not capable of resolving mass differences due to incorporation of a heavy isotope of carbon from that of any other heavy isotope. PolyMID-Correct considers high resolution measurements by setting the IMD for elements whose isotope mass differences are resolved from those of the tracer element to be 100% nominal mass. The measured and corrected high- and low- resolution MIDs of tryptophan are displayed in Figure 3.2C. The m+0 term (i.e. the term corresponding to zero atomic mass units above monoisotopic mass) is near unity when measured MID of tryptophan is corrected for abundances of naturally occurring heavy isotopes because it was extracted from a sample not exposed to a stable-isotope tracer. High-resolution data is entered as the relative areas of peaks due to incorporation of heavy isotopes of atoms of the same element as the tracer. Thus, for tryptophan in a carbon tracing experiment, the m+0 term would be the relative peak area corresponding to monoisotopic tryptophan (with no heavy isotopes of any atoms), the m+1 term would be the relative peak area corresponding to tryptophan with one ^{13}C atom, the m+2 term would be the relative peak area corresponding to tryptophan with two ^{13}C atoms, and so on. Peak areas would be relative to the sum of the m+0 peak and all peaks of tryptophan corresponding to incorporation of one or more ^{13}C atoms.





C

	High Resolution		Low Resolution	
mass	meas	corr	meas	corr
m+0	0.889	1.000	0.882	1.007
m+1	0.107	0.001	0.111	-0.003
m+2	0.004	-0.002	0.007	-0.003
m+3	-	0.000	-	0.000
m+4	-	0.000	-	0.000
m+5	-	0.000	-	0.000
m+6	-	0.000	-	0.000
m+7	-	0.000	-	0.000
m+8	-	0.000	-	0.000
m+9	-	0.000	-	0.000
m+10	-	0.000	-	0.000
m+11	-	0.000	-	0.000

Figure 3.2. Correction of High- and Low-Resolution Data. A) High-resolution mass spectrograph of protonated tryptophan. B) Low-resolution representation of the same mass spectrograph of protonated tryptophan. C) PolyMID correction of tryptophan MIDs for naturally-occurring isotope abundances in a 13C tracer experiment. Abbreviations: meas: measured, corr: corrected.

Tracer Purity and Tracer Atom Enrichment

Stable-isotope tracing experiments may be performed with tracers that are less than 100% pure and/or having heavy isotope enrichments of less than 100% at the

labeled atom positions (Figures 3.3). Consider an experiment performed with [U-$^{13}C_6$]glucose where it is 50% pure with an atom enrichment of 95%. After correcting the MID of serine for natural isotopic abundances, the labeling on serine can be interpreted as carbon atoms derived from labeled atoms of the tracer, i.e. 50% [U-$^{13}C_6$]glucose with 95% atom enrichment. However, after subsequent correction for tracer purity and enrichment, the labeling on serine can be interpreted as carbon atoms derived from glucose; a much more physiologically relevant readout. PolyMID-Correct removes the influences of tracer impurity and less than unity atom enrichment by adjusting the IMD of the labeled element as in Equation 9.

$$IMD_{E_L} = P * E * [0 \quad 1 \quad 0 \ldots] + (1 - P * E) * IMD_E \qquad \text{(Eq. 9}$$

where P is the purity of the tracer, E is the label-atom enrichment of the tracer, IMD_{E_L} is the isotope mass distribution of the labeled element, and IMD_E is the isotope mass distribution of the element as it is found in nature. In other words, the IMD of the labeled element has a component that is 100% labeled and a component with that element's naturally occurring IMD.

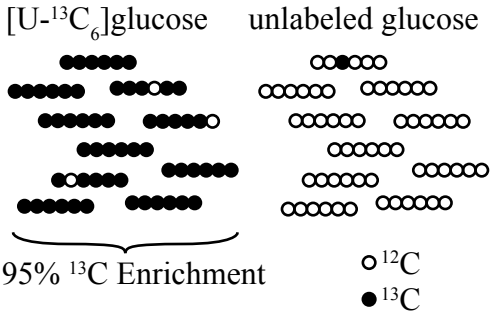


Figure 3.3. Illustrations of Tracer Enrichment and Label Enrichment. [U-$^{13}C_6$]glucose tracer enrichment is 50% and ^{13}C label enrichment is 95%.

Methods

PolyMID Availability and Installation

PolyMID is free open source software available under the MIT license. Source code is available at http://VacantiLab.github.com/PolyMID. PolyMID runs on Python 3 and is operating system independent. Python 3 is required to install and run PolyMID. To determine if Python 3 is installed, type the following command in the Terminal (MacOS and Linux) or Cmd (Windows) window:

python3 --version

If a version of Python 3 is listed, then Python 3 is installed. If not, then Python 3 must be installed prior to continuing. With Python 3 installed, use the following command in the Terminal or Cmd window to install PolyMID:

pip install PolyMID

Running PolyMID-Correct

Inputs from a Text File

The following provides instructions to perform the correction of a measured MID of tryptophan (Figure 3.2) for naturally occurring heavy isotope abundances. In a plain text editor, create a .txt file having the following format:

FragmentName: TryptophanProtonated

FragmentFormula: C11H13N2O2

CanAcquireLabel: C11H13N2O2

MIDm: 0.88885 0.106829 0.004322

LabeledElement: C

TracerEnrichment: 1

LabelEnrichment: 1

HighRes: N O H

FragmentName and *FragmentFormula* specify the name and formula of the metabolite fragment. For analyses on LCMS systems using electrospray ionization, the fragment will generally be the whole metabolite. For analyses on GCMS systems using electron ionization, the fragment will generally be a portion of the derivatized form of the whole metabolite. The value, *CanAcquireLabel*, specifies which atoms of the fragment can possibly acquire label from the tracer. Only the atoms that are of the same element as the tracer are considered. If the fragment formula is identical to the metabolite formula, then this input could be the same as the input *FragmentFormula*. However, in cases where the fragment is a derivatized version of the metabolite, atoms of the derivatizing agent cannot acquire label from the tracer and should be excluded from this input. The value, *MIDm*, is the mass isotopologue distribution of the metabolite fragment as it is measured. The value, *LabeledElement*, is the chemical symbol of the heavy isotope-labeled element on the tracer molecule. The value, *TracerEnrichment*, is the percent of the chemical species of the tracer molecule that is labeled. The value, *LabelEnrichment*, is the fraction of atoms in labeled positions on the tracer that are labeled as heavy isotopes. Finally, *HighRes* is a series of element chemical symbols, separated by spaces, whose mass shifts due to incorporation of heavy isotopes are distinguished from those of the labeled element. *HighRes* can also take on values of *all* or *none* to indicate mass shifts due to incorporation of heavy isotopes of all or no chemical species, respectively, can be distinguished from those of the labeled element.

Open the Terminal (MacOS or Linux) or Cmd (Windows) window and type the command, *Python3*, to start the Python3 interpreter. The Python3 interpreter should now be running in the Terminal or Cmd window. Type the command, *Import PolyMID*, to load the PolyMID software. PolyMID-Correct is part of the PolyMID software package. To run Poly-MID-Correct, type the command, *Output = PolyMID.Correct()*. A window accessing the operating system's directories will open. Navigate and select the text file specifying the input values as formatted according to instructions in the previous section. The program will run and print the corrected MID and associated sum of squared residuals when finished. The corrected MID and sum of squared residuals can also be accessed with the commands *Output.MIDc* and *Output.SSE*.

Inputs from the Command Line

The ability to define all inputs and call PolyMID-Correct directly from the command line allows it to be integrated into data processing workflows. The following provides instructions to perform the correction of a measured MID of tryptophan (Figure 3.2) for naturally occurring heavy isotope abundances. Open the Terminal (MacOS or Linux) or Cmd (Windows) window and type the command, *Python3*, to start the Python3 interpreter. The Python3 interpreter should now be running in the Terminal or Cmd window. Type the command, *Import PolyMID*, to load the PolyMID software. To implement the same correction for naturally occurring heavy isotopes as performed above, first define the following variables:

```
>>> FragmentName = 'TryptophanProtonated'
>>> FragmentFormula = 'C11H13N2O2'
```

```
>>> CanAcquireLabel = 'C11H13N2O2'
>>> MIDm = [0.88885, 0.106829, 0.004322]
>>> LabeledElement = 'C'
>>> TracerEnrichment = 1
>>> LabelEnrichment = 1
>>> HighRes = ['N', 'O', 'H']
```

Next, curate the variables into a single "Fragment" object, *Input*:

```
>>> Input = PolyMID.Fragment(FragmentName, FragmentFormula,
                             CanAcquireLabel, MIDm, LabeledElement,
                             TracerEnrichment, LabelEnrichment, HighRes)
```

Finally, run PolyMID-Correct passing it the *Input* variable:

```
>>> Output = PolyMID.Correct(Input)
```

The program will run and print the corrected MID and associated sum of squared residuals when finished. The corrected MID and sum of squared residuals can also be accessed with the commands *Output.MIDc* and *Output.SSE*.

Discussion

Interpretation of labeling patterns from stable-isotope tracing experiments is widely applied and dependent on accounting for the influence of naturally-occurring heavy isotopes. Investigating flux through metabolic pathways is used to probe complex mechanisms in eukaryotic systems, including: compartmental regulation of cellular metabolism (10, 11), stress-induced adaptations of tumor cells (12), nutrient selection for biosynthesis (13, 14), and the maintenance of stem cell pluripotency (15, 16). It is

also used in the process of optimizing micro-organisms to produce desirable industrial products (17) and to investigate the function of the human microbiome (18).

Furthermore, a push towards understanding biology on a genome-wide level lends towards combining systems analyses of metabolic function, i.e. stable-isotope tracing measurements, with high-throughput molecular measurements. We focus on measurements of the proteome because that is where our expertise lies and proteins are a major purveyor of gene function. The sharpening of quantitative resolution of proteomics can now link unbiased and unsupervised groupings similar tissue samples based on proteome profiles to distinct metabolic phenotypes (19) and the ever-increasing sensitivity of detecting protein phosphorylation provides troves of data linking cell signaling cascades to treatments or stresses impacting metabolism (20-23). Thus clean interpretations, less the influence of natural isotopic abundances, of label incorporation from stable-isotope tracing-based measurements will be indispensable when systems analyses of genome-wide molecular measurements (24) are applied to advance understanding of the interplay between gene function and nutrient metabolism (25, 26). PolyMID-Correct, with options for single text-file inputs or a full command-line interface, is well-suited for manual operation of multiple software components, or full integration into computational pipelines analyzing multiple levels of systems measurements.

Acknowledgements

This chapter is a reprint of “Correcting for Naturally Occurring Mass Isotopologue Abundances in Stable-Isotope Tracing Experiments with PolyMID”, *Metabolites*, vol.

11, 2021. Heesoo Jeong is the first author of the publication.

REFERENCES

1. Rosman KJR, Taylor PDP. Isotopic Compositions of the Elements 1997. Pure and Applied Chemistry. 1998;70(1):217-35.

2. Millard P, Delepine B, Guionnet M, Heuillet M, Bellvert F, Letisse F. IsoCor: isotope correction for high-resolution MS labeling experiments. Bioinformatics. 2019;35(21):4484-7.

3. Heinrich P, Kohler C, Ellmann L, Kuerner P, Spang R, Oefner PJ, et al. Correcting for natural isotope abundance and tracer impurity in MS-, MS/MS- and high-resolution-multiple-tracer-data from stable isotope labeling experiments with IsoCorrectoR. Sci Rep. 2018;8(1):17910.

4. Carreer WJ, Flight RM, Moseley HN. A Computational Framework for High-Throughput Isotopic Natural Abundance Correction of Omics-Level Ultra-High Resolution FT-MS Datasets. Metabolites. 2013;3(4).

5. Strang G. Introduction to Linear Algebra. 5th ed: Wellesley-Cambridge Press; 2016.
6. Fernandez CA, Des Rosiers C, Previs SF, David F, Brunengraber H. Correction of 13C mass isotopomer distributions for natural stable isotope abundance. J Mass Spectrom. 1996;31(3):255-62.

7. Lee WN, Byerley LO, Bergner EA, Edmond J. Mass isotopomer analysis: theoretical and practical considerations. Biol Mass Spectrom. 1991;20(8):451-8.

8. van Winden WA, Wittmann C, Heinzle E, Heijnen JJ. Correcting mass isotopomer distributions for naturally occurring isotopes. Biotechnol Bioeng. 2002;80(4):477-9.

9. Yergey J. A General Approach to Calculating Isotopic Distributions for Mass Spectrometry. International Journal of Mass Spectrometry and Ion Physics. 1983;52(2-3):337-49.

10. Vacanti NM, Divakaruni AS, Green CR, Parker SJ, Henry RR, Ciaraldi TP, et al. Regulation of substrate utilization by the mitochondrial pyruvate carrier. Mol Cell. 2014;56(3):425-35.

11. Lewis CA, Parker SJ, Fiske BP, McCloskey D, Gui DY, Green CR, et al. Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. Mol Cell. 2014;55(2):253-63.

12. Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, et al. Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature. 2011;481(7381):380-4.

13. Green CR, Wallace M, Divakaruni AS, Phillips SA, Murphy AN, Ciaraldi TP, et al. Branched-chain amino acid catabolism fuels adipocyte differentiation and lipogenesis. Nat Chem Biol. 2016;12(1):15-21.

14. Wallace M, Green CR, Roberts LS, Lee YM, McCarville JL, Sanchez-Gurmaches J, et al. Enzyme promiscuity drives branched-chain fatty acid synthesis in adipose tissues. Nat Chem Biol. 2018;14(11):1021-31.

15. Zhang H, Badur MG, Divakaruni AS, Parker SJ, Jager C, Hiller K, et al. Distinct Metabolic States Can Support Self-Renewal and Lipogenesis in Human Pluripotent Stem Cells under Different Culture Conditions. Cell Rep. 2016;16(6):1536-47.

16. Vacanti NM, Metallo CM. Exploring metabolic pathways that contribute to the stem cell phenotype. Biochim Biophys Acta. 2013;1830(2):2361-9.

17. Ghosh A, Ando D, Gin J, Runguphan W, Denby C, Wang G, et al. (13)C Metabolic Flux Analysis for Systematic Metabolic Engineering of S. cerevisiae for Overproduction of Fatty Acids. Front Bioeng Biotechnol. 2016;4:76.

18. Berry D, Loy A. Stable-Isotope Probing of Human and Animal Microbiome Function. Trends Microbiol. 2018;26(12):999-1007.

19. Johansson HJ, Socciarelli F, Vacanti NM, Haugen MH, Zhu Y, Siavelis I, et al. Breast cancer quantitative proteome and proteogenomic landscape. Nat Commun. 2019;10(1):1600.

20. Sacco F, Seelig A, Humphrey SJ, Krahmer N, Volta F, Reggio A, et al. Phosphoproteomics Reveals the GSK3-PDX1 Axis as a Key Pathogenic Signaling Node in Diabetic Islets. Cell Metab. 2019;29(6):1422-32 e3.

21. Humphrey SJ, James DE, Mann M. Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. Trends Endocrinol Metab. 2015;26(12):676-87.

22. Panizza E, Branca RMM, Oliviusson P, Orre LM, Lehtio J. Isoelectric point-based fractionation by HiRIEF coupled to LC-MS allows for in-depth quantitative analysis of the phosphoproteome. Sci Rep. 2017;7(1):4513.

23. Panizza E, Zhang L, Fontana JM, Hamada K, Svensson D, Akkuratov EE, et al. Ouabain-regulated phosphoproteome reveals molecular mechanisms for Na(+), K(+)-ATPase control of cell adhesion, proliferation, and survival. FASEB J. 2019;33(9):10193-206.

24. Vacanti NM. The Fundamentals of Constructing and Interpreting Heat Maps. Methods Mol Biol. 2019;1862:279-91.

25. Jeong H, Vacanti NM. Systemic vitamin intake impacting tissue proteomes. Nutr Metab (Lond). 2020;17:73.

26. Metallo CM, Vander Heiden MG. Understanding metabolic regulation and its influence on cell physiology. Mol Cell. 2013;49(3):388-98.

Chapter 4

Proteomic Regulation of Metabolic Plasticity in Breast Cancer Cells*

* Jeong H, Johansson HJ, Lehtiö J, Vacanti NM. Proteomic regulation of metabolic plasticity in breast cancer cells. Manuscript in consideration for *Nature Metabolism* (to be submitted on June 15, 2023).

Abstract

Altered cellular metabolism is a well-established trait of cancer. The current work establishes alterations in metabolic adaptations as defining traits of heterogenous metastatic potential. When challenged with inhibition of oxidative phosphorylation (OXPHOS) or loss of anchorage, breast cancer cells have individual metabolic, proliferative, and proteomic response profiles. Parallel application of stable-isotope tracing and high-throughput proteomics reveals functional relationships between metabolic flux alterations and protein expression. While some relationships are expected, such as abundances of electron transport chain proteins correlating with the flux response through pyruvate dehydrogenase; others are not so obvious, such as the link between glucose anaplerosis and cell cycle regulation. Protein network and metabolic flux responses are found to be additive in inhibited OXPHOS and loss of anchorage, and contribute to driving cells to a metastatic phenotype. Furthermore, *in situ* and *in silico* measurements of metabolic fluxes by application of stable-isotope tracing and metabolic flux analysis were necessary to observe the described regulatory mechanisms, as metabolic enzyme abundances are found not to correlate with their activities. Finally, the study of glucose anaplerosis required shortcomings in established methods to be addressed. Thus, the current work commences with the development of a robust metabolic flux analysis model to measure pyruvate carboxylase activity *in situ*.

Introduction

The tricarboxylic acid (TCA) cycle provides building blocks for synthesis of lipids, amino acids, and nucleotides, and generates reducing equivalents for oxidative

phosphorylation (OXPHOS) and ATP synthesis (1). When TCA cycle intermediates are utilized as precursors for macromolecule synthesis, they are replenished by a process termed anaplerosis; which is necessary to keep the TCA cycle active and maintain its biosynthetic and energetic functions (2). Glucose provides substrate for a predominant anaplerotic pathway involving the activity of pyruvate carboxylase (PC); yielding the TCA cycle intermediate, oxaloacetate (2). PC-mediated anaplerosis has been extensively studied as a metabolic regulator in diabetes (3, 4), liver metabolism (5), tumorigeneses (6, 7), and metastases (8-10). Elevated PC expression or its activity is implicated as a mechanism promoting breast tumor cell proliferation (11) and metastasis (11-13), distinguishes breast tumor subtypes (11, 14, 15), and is reported as a marker for poorer patient prognosis (12). Collectively, the above-mentioned findings from clinical, *in vivo*, and *in vitro* studies highlight the importance of PC-mediated anaplerosis in the metabolism of breast tumor cells. Thus, the current work examines the intracellular activity of PC in breast tumor cell lines.

Herein, inaccuracies in established methods to measure intracellular PC flux are exposed and a novel method based on metabolic flux analysis (16) (MFA) of [U-$^{13}C_6$]glucose tracing data is presented. Performance of the MFA-based method is evaluated in breast cancer cell lines that are exposed to small molecules targeting metabolism. Across conditions and cell lines, the MFA-based approach yields results nearest to an accepted gold-standard methodology, using [3-^{13}C]glucose (17, 18), or even exposes the gold standard's limitations. The MFA-based method does not require a separate experiment with a more costly and generally less informative metabolic tracer, [3-^{13}C]glucose, and in some cases, outperforms its application. Thus, the MFA-

based approach may be widely applicable to data collected from [U-$^{13}C_6$]glucose tracing experiments in any cell culture system.

The finding that breast cancer cell lines have different PC-flux responses to inhibited OXPHOS and glutaminase activity indicates a heterogeneity in breast tumor metabolic regulation. Exploration of this heterogeneity at the proteome level reveals metabolic flux responses are coordinated with protein expression adaptation programs linked to cellular functions, either directly, or via seemingly more distant connections. For instance, in response to inhibited OXPHOS, OXPHOS protein and the upstream pyruvate dehydrogenase (PDH) complex's flux alterations correlate. Conversely, elevated reductive carboxylation is correlated with upregulated expression of the distally related RNA polymerase complex. The direct and indirect functional relationships between metabolic activity and protein expression underscore a need to study relationships between tumor metabolism and cellular regulatory mechanisms as interacting networks. Furthermore, widely applied inferences of metabolic activities by enzyme expression (19) are limited as *in situ* measurements of enzyme activity are not found to be correlated with their abundances; highlighting the necessity of metabolic flux measurements by stable isotope tracing and MFA.

Considering the metastatic process involves inhibited OXPHOS, through limited access to oxygen, and lost cellular anchorage (20), herein, the metabolic and proteomic network adaptations to metastasis are considered as the sum of responses to the individual processes. Indeed, findings indicate that heterogenous responses in PC activity to individual application of inhibited OXPHOS and lost anchorage sum to a homogenous response in the metastatic process, and subsequently, simultaneous

application of the stresses can drive a cell line toward a metastatic phenotype.

Results

Implementation of an MFA model corrects inaccuracies in current methods to assay PC flux

Culture with [U-$^{13}C_6$]glucose is widely applied to assay intracellular utilization of glycolysis, the TCA cycle, amino acid synthesis, fatty acid synthesis, cholesterol synthesis, one-carbon metabolism, glycosylation, nucleotide synthesis, and other metabolic pathways (21-24). Culture with [U-$^{13}C_6$]glucose and readout of the relative abundance of M3 malate (M3-Mal method) is commonly used to assay flux through pyruvate carboxylase (13, 25-29), as all labeled pyruvate carbons are retained on malate while flux through PDH would surrender a labeled carbon to CO_2 (Figure 4.1A). The M3-Mal method is applied as a proxy for the gold standard of culture with [3-^{13}C]glucose (3Glc method). In application of the 3Glc method, the labeled carbon is preserved in PC-catalyzed anaplerosis, but not in the reaction catalyzed by pyruvate dehydrogenase (Figure 4.1B). Assuming 50% of pyruvate produced from glycolysis carries a labeled carbon, relative PC flux can be calculated as the abundance of M1 malate derived from [3-^{13}C]glucose multiplied by two. Application of the M3-Mal method indicates that flux through PC is elevated upon inhibition of glutaminase, with CB839, in BT20 and MCF7 cells (indicated in red in Figures 4.1C-D, Appendix Figure 4.1A-B); and that while inhibition of complex I, with rotenone, is a strong activator of PC flux in BT20 cells, it is a strong inhibitor of the same flux in MCF7 cells (indicated in green in Figures 4.1C-D, Appendix Figure 4.1A-B). However, the gold-standard

(3Glc) method reveals upregulation of PC flux is overestimated by the M3-Mal method in BT20 and MCF7 cells, both untreated or treated with CB839 (indicated in blue in Figures 4.1C-D). Multiple rounds of the TCA cycle also produce malate with three labeled carbon atoms from [U-$^{13}C_6$]glucose (Appendix Figure 4.1C); potentially leading to overestimation of PC flux by the M3-Mal method. Furthermore, the 3Glc method reveals treatment with rotenone modestly activates PC flux in MCF7 cells (indicated in purple in Figure 4.1D), in contrast to the dramatic deactivation predicted by the M3-Mal method (indicated in green in Figure 4.1D, Appendix Figure 4.1B).

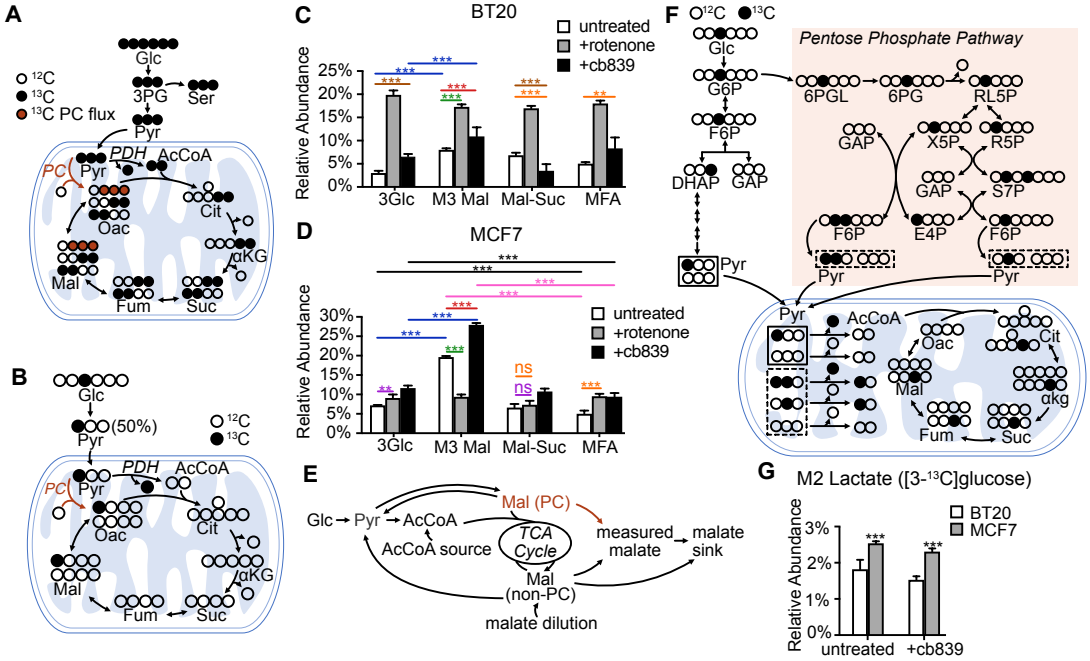


Figure 4.1. PC Flux Quantification. (A) Schematic of [U-$^{13}C_6$]glucose labeling of carbon atoms in de novo serine synthesis and the TCA cycle. (B) Schematic of [3-^{13}C]glucose labeling of carbon atoms in the TCA cycle. (C) PC flux measured with the 3Glc, M3 Mal, Mal-Suc, and MFA methods in the BT20 breast cancer cells. (D) PC flux measured with the 3Glc, M3 Mal, Mal-Suc, and MFA methods in the MCF7 breast cancer cells. (E) MFA model network schematic. (F) Schematic of [3-^{13}C]glucose labeling of carbon atoms in glycolysis, the pentose phosphate pathway, and the TCA cycle. (G) Relative abundances of M2 lactate. Error bars represent standard deviation (C, D, and G). *p<0.05, **p<0.01, and ***p<0.001 by ANOVA with Tukey's post hoc test (C and D), or a two-tailed, equal variance, Student's t test (G). AcCoA: acetyl CoA, αKG: alpha-ketoglutarate, Cit: citrate, E4P: erythrose 4-phosphate, Fum: fumarate, F6P:

fructose 6-phosphate, GAP: glyceraldehyde 3-phosphate, Glc: glucose, Mal: malate, Oac: oxaloacetate, Pyr: pyruvate, PC: pyruvate carboxylase, PDH: pyruvate dehydrogenase, RL5P: ribulose 5-phosphate, R5P: ribose 5-phosphate, Ser: serine, Suc: succinate, S7P: sedoheptulose 7-phosphate, X5P: xylulose 5-phosphate, 3PG: 3-phosphoglycerate, 6PG: 6-phosphogluconate, 6PGL: 6-phosphogluconolactone.

A proposed correction of overestimation of PC flux by the M3-Mal method involves subtraction of M3 succinate from M3 malate (Mal-Suc method). Following the assumption that succinate dehydrogenase (SDH) is irreversible, all M3 succinate would be derived from multiple rounds of the TCA cycle, and its subtraction from M3 malate would yield the contribution of M3 malate due only to flux through PC (18). However, flux through SDH is not always irreversible (30, 31). In fact, the Mal-Suc method predicts a two-fold decrease in flux through PC when BT20 cells are treated with CB839 where the 3Glc method predicts more than a two-fold increase (indicated in brown in Figure 4.1C). Additionally, application of the Mal-Suc method does not resolve increased flux through PC detected by the 3Glc method when MCF7 cells are treated with rotenone (indicated in purple in Figure 4.1D). Mass isotopologue distributions (MIDs) of succinate are available in the Supplement (Appendix Figures 4.1D-E).

Though PC flux can be measured more accurately by the 3Glc method, it requires culture with [3-^{13}C]glucose; which provides essentially no information regarding glucose metabolism downstream of PDH and is considerably more costly than [U-$^{13}C_6$]glucose. Herein, a metabolic network with nodes representing glycolytic and TCA cycle products/intermediates is built into an MFA model designed to resolve relative PC flux when fit to experimental MIDs resulting from cell culture with [U-$^{13}C_6$]glucose (MFA method). The node representing both TCA cycle intermediates and the product

of PC is compartmentalized to resolve contributions from PC and the TCA cycle (Figure 4.1E). Details of the reactions in the modeled network are available in the Methods section. When benchmarked against the 3Glc method, the MFA method corrects overestimation of the M3-Mal method in MCF7 cells, both untreated and treated with CB839 (indicated in pink in Figure 4.1D). The MFA method corrects the underestimation of the Mal-Suc method in BT20 cells treated with CB839 (indicated in orange in Figure 4.1C) and in MCF7 cells treated with rotenone (indicated in orange in Figure 4.1D). Interestingly, the MFA method predicts lower fluxes through PC in untreated MCF7 cells and in those treated with CB839 when compared to the 3Glc method (indicated in black in Figure 1D). [3-^{13}C]glucose can label malate via pyruvate dehydrogenase flux when the pentose phosphate pathway (PPP) is active (Figure 4.1F), leading to an overestimation of PC flux by the 3Glc method. Relative flux through the PPP can be assayed with the [3-^{13}C]glucose tracer by examination of M2 pyruvate (Figure 4.1F). Indeed, MCF7 cells, untreated and treated with CB839, exhibit higher relative fluxes through PPP than their BT20-cell counterparts (Figure 4.1G). Thus, the 3Glc method may be overestimating PC flux in MCF7 cells in a manner that is corrected by the MFA method.

Metabolic fluxes are differentially regulated across breast cancer cell lines

In consideration of breast cancer heterogeneity, the cellular proliferation and intracellular metabolic pathway utilization responses to rotenone were measured. The rate of proliferation is reported as the coefficient relating the derivative of cell quantity to cell quantity in an exponential model of cell growth (Methods). All cells proliferated

when left untreated (Appendix Figure 4.2A) and responses to rotenone range from decreased proliferation (ratio between zero and one) in MDAMB157 cells, arrested proliferation (ratio near zero) in T47D, HCC1419, and MCF7 cells, to cell death (negative ratio) in BT20 cells (Figure 4.2A).

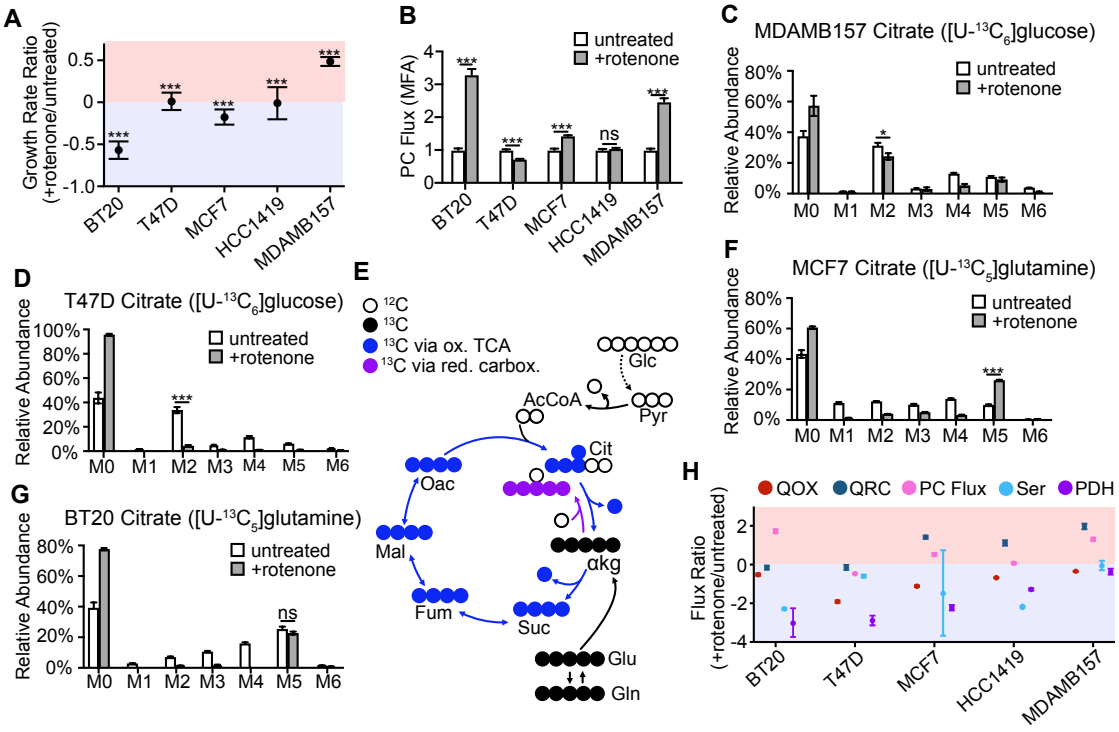


Figure 4.2. Metabolic adaptations upon OXPHOS inhibition. (A) Growth rate ratios of rotenone-treated to untreated cells. (B) PC flux tabulated with the MFA method. Fluxes are normalized to the untreated group mean for each cell-type. (C) Citrate MID from MDAMB157 cells cultured with [U-$^{13}C_6$]glucose. (D) Citrate MID from T47D cells cultured with [U-$^{13}C_6$]glucose. (E) Schematic of [U-$^{13}C_5$]glutamine labeling of carbon atoms in the TCA cycle. (F) Citrate MID from MCF7 cells cultured with [U-$^{13}C_5$]glutamine. (G) Citrate MID of BT20 cells cultured with [U-$^{13}C_5$]glutamine. (H) Log_2-transformed ratios of metabolic fluxes, rotenone-treated to untreated conditions. Error bars represent standard deviation (A-D and F-G). *p<0.05, **p<0.01, and ***p<0.001 by a two-tailed, equal variance, Student's t test (A-D and F-G). Gln: glutamine, Glu: glutamate, PDH: pyruvate dehydrogenase flux, PC Flux: Pyruvate carboxylase flux, QOX: glutamine oxidation flux, QRC: glutamine reductive carboxylation flux, Ser: *de novo* serine synthesis flux.

Metabolic pathway utilization varies similarly, but no single pathway response

correlates with cellular proliferation. In response to rotenone treatment, pyruvate carboxylase (PC) flux increases in MDAMB157 and BT20 cells, but is far less impacted in the other cell lines (Figure 4.2B). Pyruvate dehydrogenase (PDH) flux into the TCA cycle, as assayed by the relative abundance of M2 citrate in a [U-$^{13}C_6$]glucose tracing experiment (Figure 4.1A), modestly decreases in MDAMB157 cells upon treatment with rotenone (Figure 4.2C), but drops dramatically in T47D and other cell lines (Figure 4.2D and Appendix Figure 4.2B). Serine synthesis is assayed by the relative abundance of M3 serine in a [U-$^{13}C_6$]glucose tracing experiment (Figure 4.1A) and also displays a characteristic response profile to rotenone across the five breast cancer cell lines; dropping dramatically in the only cell line (HCC1419) synthesizing a substantial portion of its serine pool (Appendix Figure 4.2C).

Stable-isotope tracing with [U-$^{13}C_5$]glutamine can inform of reductive carboxylation, by labeling citrate with five heavy carbon atoms, and of glutamine oxidation, by labeling malate with four heavy carbon atoms (Figure 4.2E). In response to rotenone, MDAMB157, HCC1419, and MCF7 cells upregulate reductive carboxylation (Figures 4.2F and Appendix Figure 4.2D), while the same flux is unaffected in T47D and BT20 cell lines (Figure 4.2G and Appendix Figure 4.2D). Finally, glutamine oxidation is either decreased or unaffected by rotenone in the five breast cancer cell lines (Appendix Figure 4.2E). Findings regarding the five metabolic pathway responses to rotenone discussed above are summarized in Figure 4.2H.

Proteomic and metabolite network responses

To assess whether the metabolic network response is driven by a proteomic network

response, proteome profiles of the five breast cancer cell lines were measured when the cells were exposed to rotenone. Mass spectrometry-based proteomics methods quantified 9,869 proteins based on 161,489 uniquely identified peptides. Dimensional reduction by t-distributed stochastic neighbor embedding (t-SNE) reveals cell-type to be the strongest driver of network protein expression when cells are challenged with the major metabolic disruption of inhibited oxidative phosphorylation (Figure 4.3A, left panel). Strikingly, when protein network adaptations (ΔAbundances) are considered, rotenone treatment is the clear driver of proteomic separation (Figure 4.3A, right panel). These findings indicate that cell proteomes do not adopt expression patterns characteristic of inhibited oxidative phosphorylation, rather respond with defined shifts from cell-type driven protein expression programs.

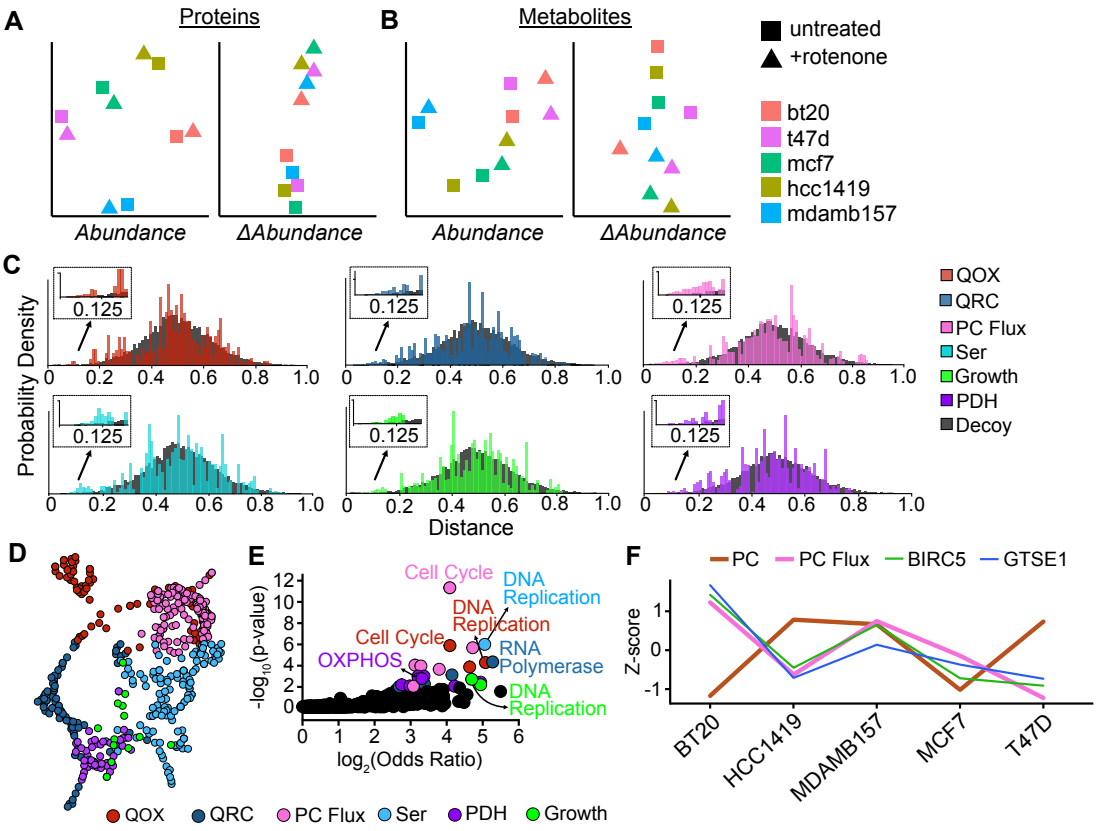


Figure 4.3. Proteomic regulation of metabolic function upon OXPHOS inhibition. (A-B) t-SNE dimensional reduction plots for proteins (A) and metabolites (B). (C) CREC-CDDA plots. (D) Correlation network of top correlating protein abundance shifts with those of growth and metabolic fluxes. Proteins with distances less than 0.125 in Figure 4.3C define the nodes. Nodes are color-coded based on the metabolic flux the represented proteins are correlated with. (E) Volcano plot showing enriched KEGG pathways from Enrichr pathway analysis performed on the top correlating proteins with growth and metabolic fluxes. Distance<0.125 in figure 4.3C was used to define the top correlating proteins. Enriched pathways with adjusted p-values less than 0.01 are color-coded based on metabolic flux the pathways are correlated with. (F) Changes of protein abundances and PC flux upon OXPHOS inhibition across the five breast cancer cell lines. Z-scores of $\log_2$-transformed ratios displayed.

To assess whether metabolism is also regulated by deviating from characteristic cell-type metabolic network function, independent measurements of metabolite abundances were examined. Indeed, t-SNE analyses of metabolite abundances do not stratify based on OXPHOS status as do analyses of metabolite abundance adaptations (Figure 4.3B).

Taken together, these data indicate both proteomic and metabolic network regulation to inhibited OXPHOS occur as deviations from cell-type driven programs and suggest the existence of metabolome-proteome feedback mechanisms regulating network responses.

Proteomic regulation of metabolic fluxes

For growth and each indicator of intracellular metabolic flux, relative responses to rotenone were correlated with protein quantity responses. Purported protein regulators of each metabolic flux and false discovery rate are determined by cross-referenced enrichment-corrected correlation distance distribution analysis (CREC-CDDA, Methods). The enrichment of computed protein-flux correlation distances at the tail near zero, relative to those of a decoy set, determines the false discovery rate (Figure 4.3C). Quantified proteins whose responses are found to correlate with a metabolic flux response to rotenone (CREC-CDDA distance < 0.125) are visualized as a network. Each node represents a protein. Nodes are joined by an edge (not shown to reduce clutter) if corresponding protein responses to rotenone are correlated beyond the indicated threshold, and arranged so that highly connected nodes are in proximity according to the Yifian Hu algorithm (32). Correlating protein abundance responses to metabolic flux responses are indicated by node color, revealing protein regulatory networks of metabolism (Figure 4.3D). Protein members of the metabolic regulatory networks are analyzed for enrichment in pathways curated by the Kyoto Encyclopedia of Genes and Genomes (33) (Figure 4.3E, Appendix Table 4.1).

Notably, protein regulatory networks whose abundance responses to rotenone

correlate to that of cell proliferation are enriched in proteins that participate in DNA replication (Figure 4.3E, Appendix Table 4.1). The functional relevance of DNA replication to cell proliferation lends confidence to the analogous findings on functional drivers of metabolic flux that are not so well established. For instance, protein regulatory networks whose abundance responses correlate to that of PC flux are enriched in regulators of the cell cycle. It is not obvious why the cell cycle and PC activity would be co-regulated in the face of inhibited OXPHOS. However, proteins whose responses to PC flux also include BIRC5 (a.k.a. Survivin) and GTSE1, potent inhibitors of apoptosis (Figure 4.3F). Thus, upregulating PC flux may permit escape of apoptosis and continued execution of the cell cycle. Protein abundance responses correlating to PDH flux are enriched in OXPHOS proteins (Figure 4.3E, Appendix Table 4.1). PDH and OXPHOS are functionally linked, as flux through PDH is the canonical pathway providing glycolytic substrate for production of reduced cofactors in the TCA cycle. Thus, mechanisms to co-regulate PDH activity and proteins of OXPHOS could be functionally advantageous. Interestingly, protein abundance responses correlating with reductive carboxylation of glutamine are enriched in members of the RNA polymerase (Figure 4.3E, Appendix Table 4.1). Reductive carboxylation was originally described as a tumorigenic means to proliferate under conditions of hypoxia or dysfunctional mitochondria (34-36). Considering hypoxia and dysfunctional mitochondria are forms of inhibited OXPHOS, it is not surprising that pharmacological inhibition of complex I also induces reductive carboxylation in breast cancer cells (Figure 4.2F-H, Appendix Figure 4.2C). Of note, the current work suggests that the reductive carboxylation response is not ubiquitously hard-wired, but rather cell-type specific. The high

correlation of the reductive carboxylation response with RNA polymerase proteins suggests reductive carboxylation also has a functional role in the transcription response when OXPHOS is inhibited. Finally, protein abundance responses correlating with oxidative glutamine anaplerosis include regulators of the cell cycle and DNA replication (Figure 4.3E, Appendix Table 4.1). Notably, these proteins participate in accomplishing similar functions as those correlating with cell proliferation and PC flux, though their identities are unique. Thus, different metabolic responses may orchestrate unique aspects of cellular function.

Finally, metabolic activity responses to inhibited OXPHOS are, in most cases, found to be divergent from the abundance responses of the enzymes catalyzing the activities. The PC flux (Figure 4.3F), serine synthesis (Appendix Figure 4.3A), and reductive carboxylation (Appendix Figure 4.3B) flux responses do not correlate with enzyme abundance responses in the respective pathways. Interestingly, the glutamine oxidation flux response correlates with the protein abundance response of GLS, but not GLS2 (Appendix Figure 4.3C), suggesting GLS as the dominant isoform catalyzing glutamine anaplerosis in breast cancer cells. The flux response of pyruvate dehydrogenase deviates sharply from the abundances of protein subunits catalyzing the reaction in MDAMB157 cells, suggesting the cell line may have some sort of compensatory mechanism to maintain PDH activity when the enzyme complex's abundance is low (Appendix Figure 4.3D).

Proteome networks respond to anchorage independence

Because the metastatic process involves both inhibited OXPHOS, through hypoxia,

and matrix detachment (37) the proteomes and metabolomes of two cell lines with diverging behaviors in tissue culture were profiled in anchored and forced-suspension conditions. Under culture conditions allowing for anchorage, BT20 cells spontaneously detach at a rate more than an order of magnitude greater than that of T47D cells (Figure 4.4A). When cells are cultured on agarose-coated surfaces, forcing suspension, BT20 cells aggregate in greater numbers of smaller clusters, whereas T47D cells form fewer, larger clusters (Figures 4.4B and 4.4C). Dimensional reduction of protein abundances of anchored and forced-suspension cells, by tSNE, groups based on cell type; however, the separation is driven by anchorage status when abundance shifts are analyzed (Figures 4.4D). Analogously, dimensional reduction of metabolomics data only groups by anchorage status when metabolite abundance shifts are considered (Figure 4.4E). Thus, as in the inhibition of OXPHOS, cells do not acquire suspension-defined proteome and metabolite profiles and both BT20 and T47D cells execute anchorage-loss-defined proteomic and metabolomic program shifts from basal states, and the program shifts results in diverging behaviors in suspension.

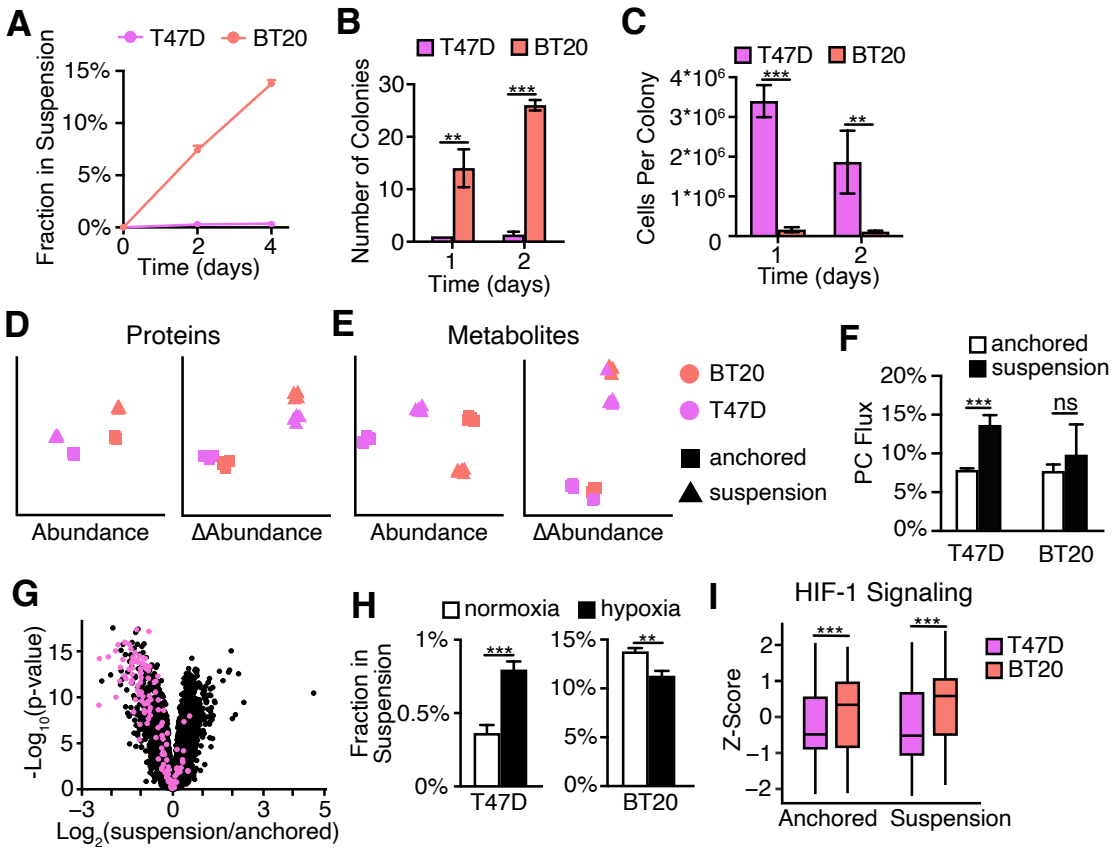


Figure 4.4. PC flux and proteome profiles of anchored and suspension cells. (A) Fraction of spontaneously detached cells among all viable cells across three time points. (B) Number of colonies formed when cells were grown in forced-detachment condition. (C) Number of viable cells in a colony under forced-detachment condition. (D-E) t-SNE dimensional reduction plots for proteins (D) and metabolites (E). (F) PC flux tabulated with the MFA method. (G) Volcano plot showing fold changes of protein abundances upon detachment in T47D. Quantified proteins are colored in black and members of the protein network regulating PC flux are colored in pink. (H) Fraction of suspension cells among all viable cells cultured under normoxia or hypoxia. (I) Boxplot showing hypoxic proteomic profile. Abundances of proteins in HIF-1 signaling KEGG pathway are transformed to z-scores. Error bars represent standard deviation (A-C, F, and H-I). *p<0.05, **p<0.01, and ***p<0.001 by a two-tailed, equal variance, Student's t test (B-C, F, and H-I).

OXPHOS and forced-suspension network responses are components of the metastatic response

Because PC activity and expression are linked to metastatic potential (8-10), relative

flux through PC was measured in response to forced-suspension for both BT20 and T47D cells. Flux through PC is elevated nearly two-fold in T47D cells and unchanged in BT20 cells (Figure 4.4F); the opposite of the PC-flux response pattern observed due to inhibited OXPHOS (Figure 4.2B). Because PC flux is regulated upon forced suspension of T47D cells, the response of the protein network associated with PC flux upon inhibition of OXPHOS (Figures 4.3C and 4.3D) was examined. Indeed, the same protein abundance shifts are highly regulated when T47D cells are forced into suspension (Figure 4.4G), but surprisingly in the opposite direction as PC flux (Figure 4.4F). Because the examined protein-network response is positively correlated with the PC-flux response when OXPHOS is inhibited (Figure 4.3C) and PC flux is dramatically upregulated in BT20 and T47D cells upon OXPHOS inhibition (Figure 4.2B) and forced suspension (Figure 4.4F), respectively, T47D cells were cultured in hypoxic conditions (1% O_2) to determine if simultaneous exposure to PC flux induced by detachment and the protein network responses linked to PC flux in conditions of OXPHOS inhibition would lead T47D cells to acquire a spontaneous detachment profile more similar to that of BT20 cells. Indeed, hypoxia induces nearly a 100% increase in the rate at which T47D cell spontaneously detach without substantially altering the high basal rate of BT20 cells (Figure 4.4H). BT20 cells cultured in normoxic conditions have a more hypoxic proteome profile relative to T47D cells, in anchored or suspension conditions (Figure 4.4I), potentially predisposing them to spontaneous detachment. Taken together, these findings indicate that components of metabolic and proteome network responses to individual stresses in the metastatic process, *i.e.* detachment and hypoxia, can interact to drive metastatic phenotypes in heterogenous tumors types.

Discussion

As the gateway to a major anaplerotic source, PC regulation/dysregulation, are established contributors to metabolic manifestations of chronic diseases and cancers (38). Aberrant hepatic PC activity can drive post-prandial gluconeogenesis under conditions of insulin resistance (3). Furthermore, PC is believed to be the major anaplerotic substrate supplier in non-small cell lung cancer (39) and allows certain gliomas and renal carcinomas to bypass TCA cycle defects while supplying anabolic substrate (6, 7, 40). Amongst studies of breast cancer, activation of pyruvate carboxylase is commonly implicated as a mechanism supporting metastasis. Fendt and colleagues report elevated PC activity in lung metastases of primary breast tumors (13) whereas Wendt and colleagues find PC to promote formation of lung metastases without impacting growth of the primary tumor (12). Furthermore, PC is reported as a driver of migration capacity and invasiveness (11).

As presented in the current work, the established methodology of carbon tracing with [3-^{13}C]glucose provides a fairly reliable readout of *in situ* PC activity. Though active flux through the PPP could be a confounding factor yielding a modest overestimation, its greatest shortcoming is not accuracy or precision, but its apparent interchangeability with interpretation of labeling patterns resulting from culture with [U-$^{13}C_6$]glucose. Because tracing with [U-$^{13}C_6$]glucose is generally informative of metabolic pathway activities, including glycolysis, the PPP, glycosylation, the TCA cycle, fatty acid synthesis, cholesterol synthesis, and amino acid synthesis; it is commonly applied to investigate metabolic adaptations. Tracing with [U-$^{13}C_6$]glucose also provides an apparent readout of relative flux through PC, prompting most

researchers to forgo a resource intensive follow-up investigation with [3-^{13}C]glucose. The current work highlights overestimation of PC flux by readouts of tracing with [U-$^{13}C_6$]glucose and inaccuracies associated with currently applied correction methods (18). The robust MFA approach presented herein, capable of harnessing data produced in experiments with [U-$^{13}C_6$]glucose, eliminates the motivation to present other problematic interpretations of glucose anaplerosis regulation based on labeling patterns derived from [U-$^{13}C_6$]glucose metabolism. Given the importance of measuring PC activity, the current work provides a foundation for fundamental investigations of metabolic regulation and a potential avenue for more accurate re-interpretations of prior works.

By defining intracellular pathway utilization and proteomic responses to inhibited OXPHOS across five breast cancer cell lines, results presented herein highlight that environmental and nutrient sensing regulatory networks (41, 42) can give rise to heterogenous metabolic responses to identical cellular stimuli. Integrated analyses highlight that metabolic responses are co-regulated with cellular functions. Several findings are of established relations between protein expression and metabolic function, including the PDH activity response positively correlating with OXPHOS protein expression shifts and the proliferation response positively correlating with expression of DNA replication proteins. Affirmation of well-established functional relationships lends confidence to the putative functional relationships between metabolism and the proteome that are seemingly more distant. For instance, the reductive carboxylation response to inhibited OXPHOS is highly correlated with proteins of the RNA polymerase complex. A precise mechanistic explanation for this relationship is elusive,

but it suggests transcription is sensitive to NAD(P)+/NAD(P)H ratios as reductive carboxylation produces NAD(P)+ by consumption of NAD(P)H (34-36). Similarly, mechanisms justifying the findings of co-regulation of PC flux, the cell cycle, and apoptosis suppression are not well-defined. However, a study of MDAMB231 breast cancer cells finds PC mRNA expression to prevent cell cycle arrest and be protective against apoptosis (43). Thus, their reported observations, and others linking PC to tumor cell survival (38), may be explained by PC activity co-regulation with apoptosis suppressors, BIRC5 and GTSE1, amongst regulators of the cell cycle.

The initial steps of the metastatic process include survival in environments with limited access to oxygen and cellular detachment (37) As described in the introduction, upregulation of PC activity is implicated as a driver of metastasis in breast cancer; though at which point in the process and the regulatory mechanisms involved are unclear. Herein, PC activity is found to be induced exclusively by inhibited OXPHOS or culture in suspension, dependent on the cell line, and is associated with different stimuli-specific proteome program shifts. Thus, the initial stages of the metastatic process can be thought of as a sum of network adaptations to the individual components. The most striking heterogeneity observed was the active downregulation of the PC-correlated, inhibited-OXPHOS, proteome program shift in T47D cells when cultured in suspension; even though suspension culture induced a 100% up-regulation of PC activity. This led to the correct prediction that T47D cells in suspension are primed to respond to hypoxia, illustrating the resolving power of network analyses on the metastatic process in a heterogeneous disease such as breast cancer.

In an era of wide application of high throughput molecular profiling of the genome,

epigenome, transcriptome, translatome, proteome, phosphoproteome, and metabolome; the network-based nature of cellular regulation is becoming clearer. However, the assumption of correlations of metabolic activity with high throughput measurements of enzyme expression is not always sound. Thus, integration of molecular profiling with measurements of pathway activity, such as stable-isotope tracing and MFA, can traverse the gap and link regulation of energetic and biosynthetic metabolism to cellular and molecular network responses. Furthermore, heterogenous diseases share common processes, such as the metastatic process in breast cancer, and progression through the metastatic process involves common metabolic adaptations. However, as the current work reveals, metabolic changes may be activated at different stages. Thus, resolving the metabolic regulations and molecular profile shifts responding to the component stresses of metastasis brings us closer to understanding the mechanisms driving similar metabolic adaptations across a heterogeneous disease like breast cancer.

Methods

Cell Culture

Growth medium for all cells was RPMI supplemented with 10% v/v fetal bovine serum. Cells were cultured in T75 flasks and passaged while subconfluent. Cells were passaged by rinsing with PBS and lifting by incubation with a trypsin EDTA solution at 37^{o}C.

Proliferation Measurements

30,000 cells were plated in each well of 24-well plates. After 48, 96, and 144 hours,

cells were stained with 1 mL PBS, 0.5 µL Hoechst, 2 µL propidium iodide and incubated at 37°C for 15 minutes. Numbers of total and dead cells were counted in triplicates using a Celigo S Imaging Cytometer (Nexcelom). Blue and red channels in viability application were used to count the numbers of cells stained with Hoechst and propidium iodide, respectively. Exposure time was set at 50,000 milliseconds. Number of live cells were quantified by subtracting the number of dead cells from the total number of cells in each well. Growth was modeled as exponential (Equation 1), and growth rate determined as the slope of the linear regression fit to the log-transformed growth equation.

$$n = n_0 e^{\mu t} \quad [1]$$

where n is the number of cells, n_0 is the initial number of cells, µ is the growth rate, and t is time.

Counting Numbers of Spontaneously Detached Cells in Normoxia and Hypoxia

Cells were cultured in hypoxia (1% O_2) for 48 hours to pre-condition the cells in hypoxia before quantification. In both normoxic (20% O_2) or hypoxic conditions, 30,000 cells were plated with growth medium in each well of 24-well plates. After 48 and 96 hours, the medium was moved to separate wells. 1 µL Hoechst and 4 µL propidium iodide were added to the medium and incubated at 37°C for 15 minutes to stain spontaneously detached cells. The cells remained anchored to the plates were stained with 1 mL PBS, 0.5 µL Hoechst, 2 µL propidium iodide and incubated at 37°C

for 15 minutes.

Counting Number of Viable Cells and Colonies in Forced-Detachment Condition

1 g of agarose in 100 mL of water was sterilized. 1 mL of the sterile agarose solution was used to coat each well of 6-well plates. 300,000 cells were plated in each agarose-coated well. After 48 and 96 hours, numbers of colonies were counted under the light microscope by putting grids (10 mm in 100 square) under the plate. Cells suspending as single cells or forming colonies were separated using cell strainers (pore size of 40 μm). The cells suspending as single cells passed through the strainer together with the medium. 1 μL Hoechst and 4 μL propidium iodide were added to the medium containing the single suspension cells in 6-well plates and incubated at 37°C for 15 minutes. The cells forming colonies were caught on the cell strainers which were moved back to 6-well plates by washing the strainer with 1 mL PBS. The colonies were broken up by pipetting multiple times. 0.5 μL Hoechst and 2 μL propidium iodide were added to each well and incubated at 37°C for 15 minutes.

Protein Extraction

Proteins were extracted from cells in sub-confluent T75 culture flasks. Growth media was removed and cells rinsed with PBS prior to being lifted off the flask with 1 mL of a trypsin+EDTA solution. The cell/trypsin suspension was diluted in 10 mL of PBS and pelleted by centrifugation at 400 g. The supernatant was removed, cells resuspended in 1 mL of PBS, and pelleted again by centrifugation at 400 g. The resulting cell pellet was stored at -80°C. Pellets were resuspended in 150 μL lysis buffer (4% w/v

SDS, 25mM HEPES, 1 mM dithiothreitol, pH 7.6), heated to 95°C for 5 minutes, and sonicated with a probe sonicator until achieving a viscosity similar to water as determined by visual inspection. Samples were centrifuged at 14,000 g for 15 minutes and the supernatant was collected. Protein was quantified using the BioRad DC Protein assay and protein solutions were stored at -80°C.

Protein Digestion

Buffer exchange and protein digestion was performed using the Sp3 magnetic bead clean-up protocol. Sp3 bead mixtures (Cytiva 45152105050250 and 65152105050250, 50 mg/mL) were removed from cold storage (4°C) and mixed for 5 minutes by swirling. Aliquots, 200 µL, from each mixture were pooled and rinsed twice with 900 µL water. Sp3 beads were separated from solution by incubation on a magnetic rack and removal of supernatant. Sp3 beads were then resuspended in 400 µL water and set aside for later use. Chloroacetamide (20 µL at 0.4 M) and the Sp3 bead solution (40 µL) were added to each protein solution, followed by acetonitrile to a final composition of 70% (v/v). Protein mixtures were incubated at room temperature in a rotating rack for 20 min. The Sp3 bead and protein mixture was incubated on a magnetic rack for 2 min, and the supernatant removed and discarded. The Sp3 beads were washed twice by resuspension and removal of 900 µL of 70% (v/v) ethanol, then allowed to dry at room temperature for 1 min. A LysC (Wako Pure Chemical Industries 129-02541) solution in water, 20 µL at 2 µg/uL, was diluted in 980 µL of a 1M urea, 50 mM HEPES (pH 7.8) solution. Sp3 beads were resuspended in 100 µL of the LysC/urea/HEPES solution and incubated at 37°C overnight with gentle swirling. Trypsin (Thermo 90057) was resuspended in

500 µL of 50 mM HEPES, pH 7.8. The trypsin solution was added to the digestion mixture, 100 µL, followed by another overnight incubation period at 37°C with gentle swirling. Reaction mixtures were then incubated on magnetic racks for 2 min and the supernatant (containing proteins digested to peptides) collected and stored at -80°C.

TMT Labeling

TMT10plex or TMTpro reagents (collectively TMT) were removed from the freezer and allowed to come to room temperature. Each reagent tag was resuspended in 40 µL of acetonitrile and set aside. From each peptide sample solution, 80 µL were aliquoted followed by addition of 8 µL of 1M TEAB. TMT reagent tag solutions were each added to a sample aliquot and allowed to incubate at room temperature for 2 hours. A single TMT10plex set was used to label each of the five cell lines, untreated and treated with rotenone (10 total samples). A single TMTpro set was used to label each of 4 replicates of BT20 and T47D cultured in attached or forced-suspension conditions (16 total samples).

All sample volumes were then pooled and cleaned by strong cation exchange (SCX). Formic acid, 10% (v/v) in water, was added to the pooled sample mixture to bring its pH below 3. The strong cation exchange column (Strata-X-C 8B-S029-TAK) was washed with 600 µL of ethanol followed by 600 µL of water. The pooled sample mixture was then added to the column. The column was then washed with 1 mL of a 30% (v/v) MeOH, 0.1% (v/v) formic acid, 69.9% (v/v) water solution. The sample was then eluted from the column with 1 mL of a 30% (v/v) MeOH, 5% (v/v) ammonia, 65% (v/v) water solution. The eluted sample was then dried under vacuum and stored at -80°C.

Prefractionation by HiRIEF

The pooled sample solution eluted from SCX was resuspended in 250 µL of 99% (v/v) 8M urea and 1% (v/v) ampholyte solution (Cytiva). A 24 cm-long immobilized pH gradient (IPG) strip, pH 3-10 (Cytiva) was swollen with the sample solution overnight. Isoelectric focusing (IEF) was run on an Ettan IPGphor system (GE Healthcare) until 150 kV-h was reached. After the focusing, peptides from the strip were extracted into 72 fractions. Peptides from each fraction were extracted by incubating with 60 µL of water for 1 hour and 60 µL of 35% acetonitrile, 64.9% water, 0.1% formic acid for 1 hour. The peptide solutions were dried under vacuum and stored at -80ºC.

Peptide Analysis by LC-MS/MS

Each HiRIEF fraction, or pooled fraction sample, was loaded onto a C18 trap column (Thermo 164535) by a capillary flow of a 97% water 3% ACN solution and separated by nano flow LC on 60 or 90 minute gradients ramping the polar component, A (95% water, 5% DMSO), to organic component, B (90% ACN, 5% DMSO, 5% water), ratio from 97:3 to 65:35. Data dependent acquisition isolated the top 5 abundant ions with a +2 to +7 charge over a 1 m/z window offset by 0.5 m/z for HCD fractionation and MS2 analysis. Ion exclusion time was set to 70 seconds. MS1 and MS2 spectra were obtained at resolutions of 70,000 and 35,000 respectively. All m/z filtering and spectral collections were performed with a Thermo QE tandem quadrupole Orbitrap MS/MS.

Spectral Search and Protein Quantification

Spectra were searched against the human subset of the Ensembl proteome database (v91) by application of MSGF+ (44) and Percolator (45). Search parameters included fixed TMT10plex modifications at N-termini and lysine residues, variable methionine oxidation, fixed carbamidomethylation of cysteine residues, a precursor mass tolerance of 10 ppm, and a precursor charge range of +2 to +6. Peptide spectral matches (PSMs) found at 1% false discovery rate (FDR) were used to infer gene identities. Protein FDR was constrained to 1% by the picked-FDR method using gene symbols as protein groups. Quantifications are thus condensed to a one-to-one mapping with gene symbols, presented as normalized to the average of all samples for every gene, followed by normalization to the resulting median of all genes for every sample, and then $\log_2$-transformation. Where protein abundance changes or adaptations due to rotenone treatment are presented, normalized quantifications are further normalized to the cell-line average for each gene prior to $\log_2$-transformation (46, 47).

Stable Isotope Tracing

Growth medium was removed from subconfluent cells plated in 6-well plates. Cells were rinsed with PBS. Tracer medium was placed on all wells and cells were incubated for 48 hours. Tracer medium was an RPMI formulation without glucose, glutamine, and pyruvate, and supplemented with 10% v/v dialyzed fetal bovine serum. Glucose and glutamine were added back at 10 mM and 2.5 mM respectively. In glucose tracing experiments, glucose was added back as either [U-$^{13}C_6$]glucose or [3-^{13}C]glucose. In glutamine tracing experiments, glutamine was added back [U-$^{13}C_5$]glutamine. After 48 hours, metabolites were extracted. Tracer medium was removed and all wells rinsed

with a saline solution. Cold methanol, 400 uL on dry ice, was charged to each well along with 160 μL of ice cold water containing 50 nmol of the internal standard, norvaline. Cells were scraped in the MeOH/water mixture and moved to a 1.5 mL microcentrifuge tube. Chloroform, 400 μL containing 0.5 nmol of the internal standard heptadecanoate, was added to each metabolite extraction solution. MeOH/water/$CHCl_3$ mixtures were shaken for 5 minutes at 1500 rpm and centrifuged at 16,000 g. The polar (top) layer was separated and dried under vacuum. The dried polar metabolites were derivatized with a methoxyamine hydrochloride solution, 20 uL at 20 mg/mL, at 40°C for 30 minutes. Reaction mixtures were then each charged with 20 μL of N-methyl-n-(tert-butyldimethylsilyl)trifluoroacetamide + 1% tert-butyldimethylchlorosilane and incubated for an additional 1 hour at 40°C. Derivatized metabolite mixtures were then analyzed with gas chromatography mass spectrometry (GCMS). From each derivatization reaction volumne, 1 uL was injected at the GCMS inlet. The carrier gas (He) flow-rate was set to 1 mL/min with a septum purge flow of 3 mL/min and a purge flow to the split vent at 15 mL/min. Injections were splitless and performed into an inert splitless liner with an inlet temperature of 270°C. Metabolites were separated in a 0.25 mm inner diameter by 30 m long column with a 0.25 um thick silica film lining. Prior to injection, the column oven equilibrated at 100°C for 30 s. The temperature ramp program was as follows: hold at 100°C for 1 min, ramp to 255°C at 3.5°C/min, ramp to 320°C at 15°C/min, hold for 4 min. GC eluants were fragmented and ionized by electron ionization. Mass to charge ratios were measured by a single quadrupole mass spectrometer. Resulting spectra were searched and integrated using the PolyMID software package. Metabolite abundances are recorded as peak areas and normalized

analogously to protein quantities. When calculating mass isotopologue distributions, incorporation of naturally occurring heavy isotopes was corrected for by PolyMID-Correct (48).

Stable Isotope Tracing on Suspension Cells

1,000,000 cells were cultured in agarose-coated 35 mm culture dishes with growth medium for 24 hours. Cells in the growth medium were centrifuged at 600 g for 5 minutes. The growth medium was removed and the cells were re-suspended with PBS which was centrifuged again at 600 g for 5 minutes. After removing PBS, the cells were re-suspended in tracer medium and moved to new agarose-coated 35 mm culture dishes. The cells were incubated with the tracer medium for 24 hours. The tracer medium was an RPMI formulation without glucose, glutamine, and pyruvate, and supplemented with 10% v/v dialyzed fetal bovine serum. 10 mM of [U-$^{13}C_6$]glucose and 2.5 mM of glutamine were added back to the medium. After 24 hours of incubation, the tracer medium containing the cells were centrifuged at 600 g for 4 minutes. After removing the medium, the cells were re-suspended with saline and centrifuged again at 600 g for 4 minutes. The saline solution was removed and the cells were re-suspended with a mixture of 160 µL of ice cold water containing 50 nmol of the internal standard, norvaline, and 400 uL of cold methanol on dry ice. Cells re-suspended in the mixture were pipetted several times. Chloroform, 400 µL containing 0.5 nmol of the internal standard heptadecanoate, was added to the cells in the mixture. MeOH/water/$CHCl_3$ mixtures were shaken for 5 minutes at 1500 rpm and centrifuged at 16,000 g to separate polar (top) and non-polar (bottom) layers. After separating and drying the polar (top)

layer under vacuum, derivatization was performed following the same method used for anchored cells.

Metabolic Flux Analysis

Metabolic Flux Analysis (MFA) was performed using the INCA software suite (49, 50). The following reactions (Equations 2-14) are included in the model:

Glc (abcdef) -> Pyr (abc) + Pyr (def) [2]

Pyr (abc) -> AcCoA (bc) + CO_2 (a) [3]

Pyr (abc) + CO_2 (d) -> Mal.pc (abcd) [4]

Mal.pc (abcd) -> Pyr (abc) + CO2 (d) [5]

Mal (abcd) -> Pyr (abc) + CO2 (d) [6]

Mal (abcd) + Mal (efgh) + AcCoA (ij) + AcCoA (kl) -> Mal (ijbc) + Mal (gflk) + CO_2 (a) + CO_2 (d) + CO_2 (e) + CO_2 (h) [7]

Mal.pc (abcd) + Mal.pc (efgh) + AcCoA (ij) + AcCoA (kl) -> Mal (ijbc) + Mal (gflk) + CO_2 (a) + CO_2 (d) + CO_2 (e) + CO_2 (h) [8]

FA (ab) -> AcCoA (ab) [9]

Mal.dilution (abcd) -> Mal (abcd) [10]

0*Mal (abcd) -> Mal.measured (abcd) [11]

0*Mal.pc (abcd) -> Mal.measured (abcd) [12]

Mal.measured (abcd) -> Mal.sink (abcd) [13]

Mal (abcd) -> Mal.sink (abcd) [14]

where lower-case letters represent the order of carbon atoms of the molecular species in each reaction. Reactions represented by Equations 2 and 13 have flux values fixed to

1.0. All other flux values were determined by INCA's solver by minimizing the sum of squared error between measured and model output mass isotopologue distributions. The flux of Equation 12 is taken as the relative contribution of PC to the malate pool. PC flux values are reported as the model output value +/- the sample standard deviation of three or more replicates. Continuation was performed on each model output to ensure PC flux was bounded, though only variances and degrees of freedom from biological replicates were used to perform downstream statistical analyses.

t-distributed stochastic neighbor embedding

Dimensionality reduction was performed by t-distributed stochastic neighbor embedding using the R package, Rtsne (51), setting output dimensionality to 2, perplexity to 3, and with 10,000 iterations.

Cross-referenced enrichment-corrected correlation distance distribution analysis (CREC-CDDA)

Pearson correlations are converted to correlation distances by Equation 15. Cross-referenced correlation distances between proteins and metabolic fluxes were computed between *ΔFluxes* and *ΔAbundances* by averaging the correlation distance of a *ΔFlux* / *ΔAbundance* pair with that of the protein of interest and its highest correlating protein (determined by examination of *ΔAbundance*'s in the present data set) in an independent data set of breast tumor proteomes (52). If the protein of interest and/or the highest correlating protein are not in the independent data set, then the next highest correlating protein(s) with the protein of interest (in the current data set) were used. Enrichment-

corrected correlation distances were then adjusted by Equation 16.

$$d_c = 0.5 * (1 - r) \quad [15]$$

$$d_{ec} = d_{cr} + (0.5 - d_{cr}) * p_e \quad [16]$$

where d_c is correlation distance, r is Pearson correlation, d_{cr} is cross-referenced correlation distance, d_{ec} is enrichment-corrected cross-referenced correlation distance, p_e is the lowest adjusted p-value of a KEGG 2021 pathway (53) enrichment, computed by the Enrichr tool (54), of the block of 100 genes containing the protein in the list of genes ranked by lowest d_{cr}. Decoy correlations were computed by correlating vectors of identical sizes populated with random numbers. Decoy p_e values for each correlation pair were randomly drawn from the p_e distribution resulting from 10,000 enrichment queries of the KEGG 2021 pathways with 100 randomly selected gene symbols of proteins quantified in the current study. A d_{ec} cut-off of 0.125 defined protein regulators of metabolic flux and resulted in less than 8-15% density contributed by the decoy distribution (*i.e.* 8-15% false discovery rate).

Correlation network of protein regulators of metabolic flux

A Pearson correlation coefficient of protein abundances adaptations greater than 0.96 amongst protein regulators of metabolism (selected as described above) was used as criteria to connect nodes and define edges. The network topology was generated by application of the Yifan Hu algorithm (32) in Gephi 0.9.7 (55).

Acknowledgements

This work was supported in part by a Cancerfonden (Swedish Cancer Society)

postdoctoral fellowship award (number CAN/2016/1219, to NMV), a Wenner-Gren Foundations postdoctoral fellowship award (number UPD2016-0099, to NMV and JL), a Karolinska Institutet Foundation Research Grant (number 2016fobi50574, to NMV), a Cornell University College of Human Ecology Data Sciences Teaching Initiative grant (to NMV), a Cornell University Center for Vertebrate Genomics Seed Grant award (to NMV), a Cornell University Center for Vertebrate Genomics Scholars award (to HJ), and Cornell University Division of Nutritional Sciences faculty start-up funds (to NMV).

Author Contributions

This study was conceived by NMV. Oversight was provided by HJJ, JL, and NMV. Experiments and analyses were performed by HJ and NMV. HJ and NMV wrote the original draft manuscript and created the figures. All authors contributed to editing and revisions of the text and figures.

Declaration of Interests

The authors declare no competing interests.

REFERENCES

1. Raimundo N, Baysal BE, Shadel GS. Revisiting the TCA cycle: signaling to tumor formation. Trends Mol Med. 2011;17(11):641-9.

2. Owen OE, Kalhan SC, Hanson RW. The key role of anaplerosis and cataplerosis for citric acid cycle function. J Biol Chem. 2002;277(34):30409-12.

3. Perry RJ, Camporez JG, Kursawe R, Titchenell PM, Zhang D, Perry CJ, et al. Hepatic acetyl CoA links adipose tissue inflammation to hepatic insulin resistance and type 2 diabetes. Cell. 2015;160(4):745-58.

4. Li X, Cheng KKY, Liu Z, Yang JK, Wang B, Jiang X, et al. The MDM2-p53-pyruvate carboxylase signalling axis couples mitochondrial metabolism to glucose-stimulated insulin secretion in pancreatic β-cells. Nat Commun. 2016;7:11740.

5. Cappel DA, Deja S, Duarte JAG, Kucejova B, Iñigo M, Fletcher JA, et al. Pyruvate-Carboxylase-Mediated Anaplerosis Promotes Antioxidant Capacity by Sustaining TCA Cycle and Redox Metabolism in Liver. Cell Metab. 2019;29(6):1291-305.e8.

6. Lussey-Lepoutre C, Hollinshead KE, Ludwig C, Menara M, Morin A, Castro-Vega LJ, et al. Loss of succinate dehydrogenase activity results in dependency on pyruvate carboxylation for cellular anabolism. Nat Commun. 2015;6:8784.

7. Cardaci S, Zheng L, MacKay G, van den Broek NJ, MacKenzie ED, Nixon C, et al. Pyruvate carboxylation enables growth of SDH-deficient cells by supporting aspartate biosynthesis. Nat Cell Biol. 2015;17(10):1317-26.

8. Schwörer S, Pavlova NN, Cimino FV, King B, Cai X, Sizemore GM, et al. Fibroblast pyruvate carboxylase is required for collagen production in the tumour microenvironment. Nat Metab. 2021;3(11):1484-99.

9. Liu C, Zhou X, Pan Y, Liu Y, Zhang Y. Pyruvate carboxylase promotes thyroid cancer aggressiveness through fatty acid synthesis. BMC Cancer. 2021;21(1):722.

10. Elia I, Doglioni G, Fendt SM. Metabolic Hallmarks of Metastasis Formation. Trends Cell Biol. 2018;28(8):673-84.

11. Phannasil P, Thuwajit C, Warnnissorn M, Wallace JC, MacDonald MJ, Jitrapakdee S. Pyruvate Carboxylase Is Up-Regulated in Breast Cancer and Essential to Support Growth and Invasion of MDA-MB-231 Cells. PLoS One. 2015;10(6):e0129848.

12. Shinde A, Wilmanski T, Chen H, Teegarden D, Wendt MK. Pyruvate carboxylase supports the pulmonary tropism of metastatic breast cancer. Breast Cancer Res. 2018;20(1):76.

13. Christen S, Lorendeau D, Schmieder R, Broekaert D, Metzger K, Veys K, et al. Breast Cancer-Derived Lung Metastases Show Increased Pyruvate Carboxylase-Dependent Anaplerosis. Cell Rep. 2016;17(3):837-48.

14. Phannasil P, Ansari IH, El Azzouny M, Longacre MJ, Rattanapornsompong K, Burant CF, et al. Mass spectrometry analysis shows the biosynthetic pathways supported by pyruvate carboxylase in highly invasive breast cancer cells. Biochim Biophys Acta Mol Basis Dis. 2017;1863(2):537-51.

15. Waks AG, Winer EP. Breast Cancer Treatment: A Review. Jama. 2019;321(3):288-300.

16. de Falco B, Giannino F, Carteni F, Mazzoleni S, Kim DH. Metabolic flux analysis: a comprehensive review on sample preparation, analytical techniques, data analysis, computational modelling, and main application areas. RSC Adv. 2022;12(39):25528-48.

17. Jang C, Chen L, Rabinowitz JD. Metabolomics and Isotope Tracing. Cell. 2018;173(4):822-37.

18. Buescher JM, Antoniewicz MR, Boros LG, Burgess SC, Brunengraber H, Clish CB, et al. A roadmap for interpreting (13)C metabolite labeling patterns from cells. Curr Opin Biotechnol. 2015;34:189-201.

19. Martinez-Reyes I, Chandel NS. Cancer metabolism: looking forward. Nat Rev Cancer. 2021;21(10):669-80.

20. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. Cell. 2011;147(2):275-92.

21. Vacanti NM, Divakaruni AS, Green CR, Parker SJ, Henry RR, Ciaraldi TP, et al. Regulation of substrate utilization by the mitochondrial pyruvate carrier. Mol Cell. 2014;56(3):425-35.

22. Lewis CA, Parker SJ, Fiske BP, McCloskey D, Gui DY, Green CR, et al. Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. Mol Cell. 2014;55(2):253-63.

23. Zhang H, Badur MG, Divakaruni AS, Parker SJ, Jager C, Hiller K, et al. Distinct Metabolic States Can Support Self-Renewal and Lipogenesis in Human Pluripotent Stem Cells under Different Culture Conditions. Cell Rep. 2016;16(6):1536-47.

24. Badur MG, Zhang H, Metallo CM. Enzymatic passaging of human embryonic stem cells alters central carbon metabolism and glycan abundance. Biotechnol J.

2015;10(10):1600-11.

25. Linares JF, Cordes T, Duran A, Reina-Campos M, Valencia T, Ahn CS, et al. ATF4-Induced Metabolic Reprograming Is a Synthetic Vulnerability of the p62-Deficient Tumor Stroma. Cell Metab. 2017;26(6):817-29.e6.

26. Krycer JR, Yugi K, Hirayama A, Fazakerley DJ, Quek LE, Scalzo R, et al. Dynamic Metabolomics Reveals that Insulin Primes the Adipocyte for Glucose Metabolism. Cell Rep. 2017;21(12):3536-47.

27. Leone RD, Zhao L, Englert JM, Sun IM, Oh MH, Sun IH, et al. Glutamine blockade induces divergent metabolic programs to overcome tumor immune evasion. Science. 2019;366(6468):1013-21.

28. Davis JC, Alves TC, Helman A, Chen JC, Kenty JH, Cardone RL, et al. Glucose Response by Stem Cell-Derived β Cells In Vitro Is Inhibited by a Bottleneck in Glycolysis. Cell Rep. 2020;31(6):107623.

29. Faubert B, Li KY, Cai L, Hensley CT, Kim J, Zacharias LG, et al. Lactate Metabolism in Human Lung Tumors. Cell. 2017;171(2):358-71.e9.

30. Tomitsuka E, Kita K, Esumi H. The NADH-fumarate reductase system, a novel mitochondrial energy metabolism, is a new target for anticancer therapy in tumor microenvironments. Ann N Y Acad Sci. 2010;1201:44-9.

31. Chinopoulos C. Which way does the citric acid cycle turn during hypoxia? The critical role of α-ketoglutarate dehydrogenase complex. J Neurosci Res. 2013;91(8):1030-43.

32. Hu YF. Algorithms for Visualizing Large Networks. Ch Crc Comp Sci Ser. 2012:525-49.

33. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 1999;27(1):29-34.

34. Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, et al. Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature. 2011;481(7381):380-4.

35. Mullen AR, Wheaton WW, Jin ES, Chen PH, Sullivan LB, Cheng T, et al. Reductive carboxylation supports growth in tumour cells with defective mitochondria. Nature. 2011;481(7381):385-8.

36. Wise DR, Ward PS, Shay JE, Cross JR, Gruber JJ, Sachdeva UM, et al. Hypoxia promotes isocitrate dehydrogenase-dependent carboxylation of alpha-ketoglutarate

to citrate to support cell growth and viability. Proc Natl Acad Sci U S A. 2011;108(49):19611-6.

37. Rankin EB, Nam JM, Giaccia AJ. Hypoxia: Signaling the Metastatic Cascade. Trends Cancer. 2016;2(6):295-304.

38. Lao-On U, Attwood PV, Jitrapakdee S. Roles of pyruvate carboxylase in human diseases: from diabetes to cancers and infection. J Mol Med (Berl). 2018;96(3-4):237-47.

39. Fan TW, Lane AN, Higashi RM, Farag MA, Gao H, Bousamra M, et al. Altered regulation of metabolic pathways in human lung cancer discerned by (13)C stable isotope-resolved metabolomics (SIRM). Mol Cancer. 2009;8:41.

40. Izquierdo-Garcia JL, Cai LM, Chaumeil MM, Eriksson P, Robinson AE, Pieper RO, et al. Glioma cells with the IDH1 mutation modulate metabolic fractional flux through pyruvate carboxylase. PLoS One. 2014;9(9):e108289.

41. Jeong H, Vacanti NM. Systemic vitamin intake impacting tissue proteomes. Nutr Metab (Lond). 2020;17:73.

42. Metallo CM, Vander Heiden MG. Understanding metabolic regulation and its influence on cell physiology. Mol Cell. 2013;49(3):388-98.

43. Rattanapornsompong K, Khattiya J, Phannasil P, Phaonakrop N, Roytrakul S, Jitrapakdee S, et al. Impaired G2/M cell cycle arrest induces apoptosis in pyruvate carboxylase knockdown MDA-MB-231 cells. Biochem Biophys Rep. 2021;25:100903.

44. Kim S, Pevzner PA. MS-GF+ makes progress towards a universal database search tool for proteomics. Nat Commun. 2014;5:5277.

45. Granholm V, Kim S, Navarro JC, Sjölund E, Smith RD, Käll L. Fast and accurate database searches with MS-GF+Percolator. J Proteome Res. 2014;13(2):890-7.

46. Panizza E, Regalado BD, Wang F, Nakano I, Vacanti NM, Cerione RA, et al. Proteomic analysis reveals microvesicles containing NAMPT as mediators of radioresistance in glioma. Life Sci Alliance. 2023;6(6).

47. Vacanti NM. The Fundamentals of Constructing and Interpreting Heat Maps. Methods Mol Biol. 2019;1862:279-91.

48. Jeong H, Yu Y, Johansson HJ, Schroeder FC, Lehtiö J, Vacanti NM. Correcting for Naturally Occurring Mass Isotopologue Abundances in Stable-Isotope Tracing Experiments with PolyMID. Metabolites. 2021;11(5).

49. Young JD. INCA: a computational platform for isotopically non-stationary metabolic flux analysis. Bioinformatics. 2014;30(9):1333-5.

50. Rahim M, Ragavan M, Deja S, Merritt ME, Burgess SC, Young JD. INCA 2.0: A tool for integrated, dynamic modeling of NMR- and MS-based isotopomer measurements and rigorous metabolic flux analysis. Metab Eng. 2022;69:275-85.

51. L vdM. Accelerating t-SNE using Tree-Based Algorithms. Journal of Machine Learning Research. 2014;15:3221-45.

52. Johansson HJ, Socciarelli F, Vacanti NM, Haugen MH, Zhu Y, Siavelis I, et al. Breast cancer quantitative proteome and proteogenomic landscape. Nat Commun. 2019;10(1):1600.

53. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023;51(D1):D587-d92.

54. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128.

55. Bastian M. HS, Jacomy M. Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media. 2009.

Appendix to Chapter 1

Abbreviations

AcCoA	Acetyl Coenzyme A
AKT	Serine/Threonine Kinases
amu	Atomic Mass Unit
DNA	Deoxyribonucleic Acid
ECM	Extracellular Matrix
ER	Estrogen Receptor
ETC	Electron Transport Chain
HER2	Human Epidermal Growth Factor Receptor 2
MFA	Metabolic Flux Analysis
MID	Mass Isotopologue Distribution
PC	Pyruvate Carboxylase
PDH	Pyruvate Dehydrogenase
PI3K	Phosphoinositide 3-Kinases
PR	Progesterone Receptor
RAR	Retinoic Acid Receptor
RXR	Retinoid X Receptor
TCA	Tricarboxylic Acid
TN	Triple-Negative
US	United States
[U-$^{13}C_6$]glucose	Uniformly Labeled Glucose

Appendix to Chapter 2

Abbreviations

AcAc	Acetoacetate
ACACA	Acetyl-CoA Carboxylase
ACAD	Acyl-CoA Dehydrogenase
AcCoA	Acetyl coenzyme A
ADAMTS13	ADAM Metallopeptidase with Thrombospondin Type I Motif 13
ADCY	Adenylate Cyclase
ADIPOQ	Adiponectin
ADP	Adenine Dinucleotide Phosphate
aKG	Alpha ketoglutarate
AKT	Serine/threonine kinases
ALAS1-2	5'-Aminolevulinate Synthase 1-2
ALB	Albumin
ALDH	Aldehyde dehydrogenases
Ala	Alanine
amu	Atomic mass unit
APO	Apolipoproteins
APOA1	Apolipoprotein A1
APP	amyloid beta precursor protein
ART	ADP-Ribosyl Transferase
Asp	Aspartate
ATP	Adenosine Triphosphate
ATPB	ATP Synthase Beta
BCKDHA	Branched Chain Keto Acid Dehydrogenase E1 Subunit Alpha
BCO1	Beta Carotene Oxygenase 1
BCO2	Beta Carotene Oxygenase 2
BDH2	3-Hydroxybutyrate Dehydrogenase 2
BGLAP	Bone Gamma-Carboxyglutamate Protein

CALM1	Calmodulin 1
cAMP	Cyclic Adenosine Monophosphate
Cit	Citrate
CoA	Coenzyme A
CFL1	Cofilin 1
CRABP1	Cellular retinoic acid binding protein 1
CRABP2	Cellular retinoic acid binding protein 2
CRP	C-Reactive Protein
CTH	Cystathionine Gamma-Lyase
CTP	Cytidine Triphosphate
CNRIP1	Cannabinoid Receptor-Interacting Protein 1
CPT1A	Carnitine Palmitoyltransferase 1A
CPT1B	Carnitine Palmitoyltransferase 1B
CPT1C	Carnitine Palmitoyltransferase 1C
CYP2R1	Cytochrome P450 Family 2 Subfamily R Member 1
CYP27B1	Cytochrome P450 Family 27 Subfamily B Member 1
DBT	Dihydrolipoamide Branched Chain Transacylase E2
DDC	Dopa Decarboxylase
DHCR7	7-dehydrocholesterol reductase
DHO	Dihydroorotate
DHRS	Dehydrogenase/Reductase
DLAT	Dihydrolipoamide S-Acetyltransferase
DLD	Dihydrolipoamide Dehydrogenase
DLST	Dihydrolipoamide S-Succinyltransferase
DNA	Deoxyribonucleic acid
dTMP	Deoxythymidine Monophosphate
dUDP	Deoxyuridine Diphosphate
dUMP	Deoxyuridine Monophosphate
ECM	Extracellular matrix
ER	Estrogen receptor
EGLN1	Egl-9 Family Hypoxia Inducible Factor 1

EGLN2	Egl-9 Family Hypoxia Inducible Factor 2
EGLN3	Egl-9 Family Hypoxia Inducible Factor 3
ENO2	Enolase 2
FAD	Flavin Adenine Dinucleotide
FASN	Fatty Acid Synthase
FETUB	Fetuin-B
FLAD1	Flavin Adenine Dinucleotide Synthetase 1
FMN	Flavin Mononucleotide
Fum	Fumarate
FXN	Frataxin
F2	Coagulation Factor II, Thrombin
F6P	Fructose 6-Phosphate
F7	Coagulation Factor VII
F9	Coagulation Factor IX
F10	Coagulation Factor X
GAD 1-2	Gamma-Aminobutyric Acid 1-2
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GC	GC Vitamin D Binding Protein
GGCX	Gamma-Glutamyl Carboxylase
Glc	Glucose
GlcN	Glucosamine
GlcNAc	N-Acetylglucosamine
Glu	Glutamate
GLUD1	Glutamate Dehydrogenase 1
GLUT1	Glucose Transporter 1
GLUT3	Glucose Transporter 3
GLUT4	Glucose Transporter 4
Gly	Glycine
GOT 1-2	Glutamic-Oxaloacetic Transaminase 1-2
GPD2	Glycerol-3-Phosphate Dehydrogenase 2
GPT	Glutamic-Pyruvic Transaminase

GSTA3	Glutathione S-Transferase Alpha 3
GSTP1	Glutathione S-Transferase Pi 1
GPX1	Glutathione Peroxidase 1
G3P	Glyceraldehyde 3-Phosphate
G6P	Glucose 6-Phosphate
G6PD	Glucose 6-Phosphate Dehydrogenase
Hcy	Homocysteine
HER2	Human epidermal growth factor receptor 2
HGNS	HUGO Gene Nomenclature Committee
HIF1A	Hypoxia Inducible Factor 1 Subunit Alpha
HLCS	Holocarboxylase Synthetase
HMGCoA	3-Hydroxy-3-Methylglutaryl-CoA
IC	Isocitrate
IDH1-2	Isocitrate Dehydrogenase (NADP(+)) 1-2
IDH3A	Isocitrate Dehydrogenase (NAD(+)) 3 Catalytic Subunit Alpha
Ile	Isoleucine
INPP1	Inositol Polyphosphate-1-Phosphatase
KEGG	Kyoto Encyclopedia of Genes and Genomes
KEAP	Kelch Like ECH Associated Protein
KYNU	Kynureinase
LDHB	Lactate Dehydrogenase B
Leu	Leucine
LRAT	Lecithin retinol acyltransferase
Mal	Malate
MCCC1	Methylcrotonyl-CoA Carboxylase Subunit 1
MCCC2	Methylcrotonyl-CoA Carboxylase Subunit 2
mDNA	Methylated DNA
Met	Methionine
ME1	Malic Enzyme 1
ME3	Malic Enzyme 3
MFA	Metabolic flux analysis

MGP	Matrix Gla Protein
MID	Mass isotopologue distribution
MMACHC	Methylmalonic Acidemia with Homocystinuria, Cobalamin Deficiency Type C
MMCoA	Methylmalonyl-CoA
MTHFR	Methylene Tetrahydrofolate Reductase
MTR	Methionine Synthase
MUT	Methylmalonyl-CoA Mutase
NAD	Nicotinamide adenine dinucleotide
NADK	NAD Kinase
NADSYN1	NAD Synthetase 1
NAMPT	Nicotinamide Phosphoribosyltransferase
NAPRT	Nicotinate Phosphoribosyltransferase
NDUFS1	NADH:Ubiquinone Oxidoreductase Core Subunit S1
NMNAT	Nicotinamide Nucleotide Adenylyltransferase
NRF2	NFE2 Like BZIP Transcription Factor 2
Oac	Oxaloacetate
OGDH	Oxoglutarate Dehydrogenase
PARP	ADP Ribose Polymerase
PANK1-4	Pantothenate Kinase 1-4
PC	Pyruvate carboxylase
PCCA	Propionyl-CoA Carboxylase
PDH	Pyruvate dehydrogenase
PDHA	Pyruvate Dehydrogenase E1 Subunit Alpha
PDHB	Pyruvate Dehydrogenase E1 Subunit Beta
PDHX	Pyruvate Dehydrogenase Complex Component X
PDIA3	Protein Disulfide Isomerase Family A Member 3
PDXK	Pyridoxal Kinase
PEP	Phosphoenolpyruvate
PFKP	Phosphofructokinase, Platelet
PGAM1	Phosphoglycerate Mutase 1

PGD	Phosphogluconate Dehydrogenase
PHD1	Prolyl Hydroxylase 1
PHD2	Prolyl Hydroxylase 2
PHD3	Prolyl Hydroxylase 3
PI3K	Phosphoinositide 3-kinases
PKM	Pyruvate Kinase, Muscle
PLOD1	Procollagen-Lysine, 2-Oxoglutarate 5-Dioxygenase 1
PLOD2	Procollagen-Lysine, 2-Oxoglutarate 5-Dioxygenase 2
PLOD3	Procollagen-Lysine, 2-Oxoglutarate 5-Dioxygenase 3
PNPO	Pyridoxine Phosphate Oxidase
PPARA	Peroxisome Proliferator Activated Receptor Alpha
PPCoA	Propionyl-CoA
PR	Progesterone receptor
PROC	Protein C, Inactivator of Coagulation Factors Va and VIIIa
PROS1	Protein S
PROZ	Protein Z, Vitamin K Dependent Plasma Glycoprotein
PYGL	Glycogen Phosphorylase L
Pyr	Pyruvate
PYGM	Glycogen Phosphorylase, Muscle-Associated
QTOF	Quadrupole time-of-flight
QQQ	Triple-quadrupole
RAR	Retinoic acid receptor
RARA	Retinoic acid receptor alpha
RARB	Retinoic acid receptor beta
RARG	Retinoic acid receptor gamma
RBP	Retinol-binding proteins
RDH	Retinol Dehydrogenases
reps	Replicates
RFK	Riboflavin Kinase
RHO	Rhodopsin
Ru5P	Ribulose 5-Phosphate

RXR	Retinoid X receptor
RXRA	Retinoid X receptor alpha
RXRB	Retinoid X receptor beta
RXRG	Retinoid X receptor gamma
R5P	Ribose 5-Phosphate
SAH	S-Adenosylhomocysteine
SAM	S-Adenosylmethionine
SDHA	Succinate Dehydrogenase Complex Flavoprotein Subunit A
SDHB	Succinate Dehydrogenase Complex Flavoprotein Subunit B
SDHC	Succinate Dehydrogenase Complex Flavoprotein Subunit C
SDHD	Succinate Dehydrogenase Complex Flavoprotein Subunit D
Ser	Serine
SHMT1	Serine Hydroxymethyltransferase 1
SHMT 2	Serine Hydroxymethyltransferase 2
SLC2A1	Solute Carrier Family 2 Member 1
SLC2A3	Solute Carrier Family 2 Member 3
SLC2A4	Solute Carrier Family 2 Member 4
SLC19A2	Solute Carrier Family 19 Member 2
SLC19A3	Solute Carrier Family 19 Member 3
SLC23A1	Solute Carrier Family 23 Member 1
SLC23A2	Solute Carrier Family 23 Member 2
SLC25A20	Solute Carrier Family 25 Member 20
SLC5A6	Solute Carrier Family 5 Member 6
SPTLC 1	Serine Palmitoyltransferase Long Chain Base Subunit 1
SPTLC 2	Serine Palmitoyltransferase Long Chain Base Subunit 2
Suc	Succinate
SucCoA	Succinyl-CoA
TAGLN	Transgelin
TCA	Tricarboxylic acid
TDP	Thiamine Diphosphate
tech	Technical

THF	Tetrahydrofolate
Thr	Threonine
THRSP	Thyroid Hormone Responsive
TKT	Transketolase
TMLH	Trimethyllysine Dioxygenase
TN	Triple negative
TPI1	Triose Phosphate Isomerase 1
TPK1	Thiamine Pyrophosphokinase 1
Triple-TOF	Triple – time of flight
TTPA	Alpha Tocopherol Transfer Protein
TXNRD1	Thioredoxin Reductase 1
TXNRD2	Thioredoxin Reductase 2
TXNRD3	Thioredoxin Reductase 3
UDP	Uridine Diphosphate
UTP	Uridine Triphosphate
UQ	Ubiquinone
Val	Valine
VDR	Vitamin D Receptor
X5P	Xylulose 5-Phosphate
2DGE	Two-Dimensional Gel Electrophoresis
2DGE-MS	Two-Dimensional Gel Electrophoresis–Mass Spectrometry
6PGL	6-Phosphogluconolactone
6PG	6-Phosphogluconate

Appendix Table 2.1: Enzyme, enzyme complexes, or enzyme families requiring vitamin A as a cofactor or substrate

Enzyme	Function
Rhodopsin kinase (*GRK1*)	Phosphorylation of rhodopsin to terminate phototransduction.
Retinoic acid binding proteins (*CRABP1*, *CRABP2*)	Solubilization of retinoic acid and regulation of intracellular retinoic acid metabolism.
Retinol binding protein (*RBP1*, *RBP2*, *RBP3*, *RBP4*, *RBP5*)	Transport of retinol from hepatic parenchymal cell into the blood by forming holo-retinol binding protein (holo-RBP).
Transthyretin (*TTR*)	Transport of holo-RBP in the blood by forming holo-RBP-transthyretin ternary complex.
β-carotene 9'10' dioxygenase (*BCO2*)	Cleavage of β-carotene into β-apo-10'-carotenal and β-ionine.
β-carotene 15,15'-dioxygenase (*BCO1*)	Cleavage of β-carotene into retinal.
Lecithin retinol acyltransferase (*LRAT*)	Esterification of retinol by transferring *sn*-1 fatty acids from phosphatidylcholine to retinol.
Retinol dehydrogenases (*RDH5*, *RDH8*, *RDH10*, *RDH11*, *RDH12*, *RDH13*, *RDH14*, *RDH16*, *DHRS1*, *DHRS2*, *DHRS3*, *DHRS4*, *DHRS4L1*, *DHRS4L2*, *DHRS7*, *DHRS7B*, *DHRS7C*, *DHRS9*, *DHRS11*, *DHRS12*, *DHRS13*, *DHRSX*)	Conversion of retinol to retinal.

Appendix Table 2.2: Enzyme, enzyme complexes, or enzyme families requiring vitamin B1 (thiamine) as a cofactor or substrate

Enzyme	Function
Pyruvate dehydrogenase complex (*PDHA1*, *PDHA2*, *PDHB*, *PDHX*, *DLAT*, *DLD*)	Decarboxylation of pyruvate to acetyl-CoA.
α-Ketoglutarate dehydrogenase complex (*OGDH*, *DLST*, *DLD*)	Decarboxylation of α-ketoglutarate to succinyl-CoA.
Branched-chain α-keto acid dehydrogenase complex (*BCKDHA*, *BCKDHB*, *DBT*, *DLD*)	Transamination of valine, isoleucine and leucine.
Transketolase (*TKT*)	Synthesis of pentoses and NADPH in glycolysis and pentose phosphate pathway.
Thiamine transporters (*SLC19A2*, *SLC19A3*)	Transport of thiamine across the plasma membrane.
Mitochondrial thiamine pyrophosphate carrier (*SLC25A19*)	Transport of thiamine pyrophosphate into mitochondria.

Appendix Table 2.3: Enzyme, enzyme complexes, or enzyme families requiring vitamin B2 (riboflavin) as a cofactor or substrate

Enzyme	Function
Succinate dehydrogenase complex (*SDHA*, *SDHB*, *SDHC*, SDHD)	Complex II in electron transport chain. Conversion of succinate to fumarate in TCA cycle.
Pyruvate dehydrogenase complex (*PDHA1*, *PDHA2*, *PDHB*, *PDHX*, *DLAT*, *DLD*)	Decarboxylation of pyruvate to acetyl-CoA.
sn-Glycerophosphate dehydrogenase (*GPD2*)	NADH shuttle between glycolysis and electron transport chain in mitochondria.
Electron transfer flavoprotein dehydrogenase (*ETFDH*)	Transfer electron from electron transfer flavoprotein to ubiquinone.
Dihydrolipoyl dehydrogenase (*DLD*)	Oxidative decarboxylation of pyruvate by pyruvate dehydrogenase complex.
Acyl-CoA dehydrogenases (*ACADM*, *IVD*, *ACADS*, *ACADL*, *GCDH*, *ACADVL*, *ACADSB*, *ACAD8*, *ACAD9*, *ACAD10*, *ACAD11*)	Dehydrogenation of acyl-CoA esters in fatty acid β-oxidation.
Dimethylglycine dehydrogenase (*DMGDH*)	Catalysis of dimethylglycine to form sarcosine in choline catabolism.
Choline dehydrogenase (*CHDH*)	Choline catabolism.
Sarcosine dehydrogenase (*SARDH*)	Catalysis of sarcosine to form glycine.
NADPH oxidases (*NOX1*, *NOX2*, *NOX3*, *NOX4*, *NOX5*, *DUOX1*, *DUOX2*)	Production of superoxide and hydroperoxide.
Xanthine oxidase (*XDH*)	Oxidation of hypoxanthine to xanthine in purine catabolism.
Aldehyde oxidase (*AOX1*)	Aldehyde oxidation in pyridoxic acid and retinoic acid metabolism.
Pyridoxine phosphate oxidase (*PNPO*)	Conversion of pyridoxamine phosphate to pyridoxal phosphate, an active form of vitamin B_6.
D- and L-amino oxidase (*DAO*, *IL4IA*)	Oxidative deamination of D- and L-amino acids.
Monoamine oxidase (*MAOA*, *MAOB*)	Dopamine, tyramine and histamine metabolism.
Sulfhydryl oxidases (*QSOX1*)	Disulfide bond formation in proteins.
Spermine oxidase (*SMOX*)	Amine catabolism.
Kynurenine 3-monooxygenase (*KMO*)	Conversion of kynurenine to 3-hydroxy-kynurenine in tryptophan metabolism.
Squalene monooxygenase (*SQLE*)	Squalene oxidation in cholesterol synthesis.
Flavin monooxygenases (*FMO1*, *FMO2*, *FMO3*, *FMO4*, *FMO5*, *FMO6*)	Oxidation of amines and sulfides.
NADPH-cytochrome P450 oxidoreductase (*POR*)	Electron donor for cytochrome b_5, heme oxygenase, 3-β-hydroxysterol delta-24-reductase and squalene monooxygenase.
NADH-ubiquinone oxidoreductase (*NDUFV1*, *NDUFV2*, *NDUFV3*)	Complex I in electron transport chain.

Cytochrome b_5 reductase (*CYB5R1*, *CYB5R2*, *CYB5R3*, *CYB5R4*)	Double bond formation in fatty acids.
3-β-Hydroxysterol delta-24-reductase (*DHCR24*)	Conversion of desmosterol to cholesterol.
7-Dehydrocholesterol reductase (*DHCR7*)	Conversion of 7-dehydrocholesterol to cholesterol.
Methylenetetrahydrofolate reductase (*MTHFR*)	Conversion of 5,10-methylene tetrahydrofolate to 5-methyl tetrahydrofolate.
Glutathione reductase (*GSR*)	Reduction of glutathione disulfide (GSSG) to glutathione (GSH).
Thioredoxin reductase (*TXNRD1*, *TXNRD2*, *TXNRD3*)	Thioredoxin reduction in conversion of ribonucleotides to deoxyribonucleotides in DNA synthesis.
3-Ketosphinganine reductase (*KDSR*)	Conversion of 3-ketosphinganine to sphinganine in sphingosine synthesis.
Dihydroceramide desaturase (*DEGS1*, *DEGS2*)	Desaturation of dihydroceramide to maintain balance between sphingolipids and dihydrosphingolipids.
Dihydroxyacetone phosphate acyltransferase (*GNPAT*)	Ether lipids synthesis in peroxisomes.
Alkyl dihydroxyacetone phosphate synthase (*AGPS*)	
FAD synthase (*FLAD1*)	Adenylation of flavin mononucleotide (FMN) for flavin adenine dinucleotide (FAD) generation.
Nitric oxide synthases (*NOS1*, *NOS2*, *NOS3*)	Production of nitric oxides from L-arginine.
Lysine-specific demethylase 1 (*KDM1A*)	Demethylation of mono- and di-methylated lysine 4 of histone H3 (H3K4) and lysine 9 of histone 3 (H3K9).

Appendix Table 2.4: Enzyme, enzyme complexes, or enzyme families requiring vitamin B3 (niacin) as a cofactor or substrate

Pathway	**Enzyme**
Glycolysis	Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*)
Interconversion of pyruvate and lactate	Lactate dehydrogenase (*LDHA*, *LDHB*, *LDHC*)
TCA cycle	Pyruvate dehydrogenase (*PDHA1*, *PDHA2*, *PDHB*), Isocitrate dehydrogenase (*IDH1*, *IDH2*, *IDH3A*, *IDH3B*, *IDH3G*), α-Ketoglutarate dehydrogenase (*OGDH*), Malate dehydrogenase (*MDH1*, *MDH2*)
Fatty acid β-oxidation	β-Hydroxyacyl-CoA dehydrogenase (*HADH*,

	HSD17B10, *HSD17B4*, *EHHADH*)
Vitamin B_6 Catabolism	Aldehyde dehydrogenases (*ALDH3A1*, *ALDH3A2*, *ALDH3B1*, *ALDH3B2*, *ALDH1A1*, *ALDH1A2*, *ALDH1A3*, *ALDH1B1*, *ALDH1L1*, *ALDH1L2*, *ALDH2*, *ALDH4A1*, *ALDH5A1*, *ALDH6A1*, *ALDH7A1*, *ALDH8A1*, *ALDH9A1*, *ALDH16A1*, *ALDH18A1*)
Pentose phosphate pathway	Glucose-6-phosphate dehydrogenase (*G6PD*), 6-Phosphogluconate dehydrogenase (*PGD*)
Malate aspartate shuttle	Malate dehydrogenase (*MDH1*, *MDH2*)
Glycerol-3-phosphate shuttle	Glycerol-3-phosphate dehydrogenase (*GPD1*, *GPD2)*
Fatty acid and steroid synthesis	β-Ketoacyl-acyl carrier protein (ACP) reductase (*HSD17B8*), 3-Hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase (*HMGCR*)
Triacylglycerol synthesis	Diacylglycerol O-transferase 1 (*DGAT1*)
Ketone body formation	3-Hydroxybutyrate dehydrogenase (*BDH1*, *BDH2*)
Folate-dependent one carbon metabolism	Dihydrofolate reductase (*DHFR*, *DHFR2*), Methylene THF reductase (*MTHFR*), Methylene THF dehydrogenase (*MTHFD1*, *MTHFD2*)
Branched chain amino acid catabolism	β-Hydroxyisobutyrate dehydrogenase (*HIBADH*), Methylmalonic semialdehyde dehydrogenase (*ALDH6A1*), Branched chain keto acid dehydrogenase (*BCKDHA*, *BCKDHB*), β-Hydroxyacyl-CoA dehydrogenase (*HADH*, *HSD17B10*, *HSD17B4*, *EHHADH*)
Lysine catabolism	Lysine-α-ketoglutarate reductase (*AASS*), Saccharopine dehydrogenase (*SCCPDH*), L-α-Aminoadipate-δ-semialdehyde dehydrogenase (*AASDH*)
Serine and glycine metabolism	D-Glycerate dehydrogenase (*GRHPR*), D-3-Phosphoglycerate dehydrogenase (*PHGDH*)
Pyrimidine synthesis	Dihydroorotate dehydrogenase (*DHODH*), Ribonucleotide reductase (*RRM1*, *RRM2*, *RRM2B*)
ADP-ribosylation reactions for post translational modifications	ADP ribose polymerases (*PARP1*, *PARP2*, *PARP3*, *PARP4*, *TNKS*, *TNKS2*, *PARP6*, *TIRARP*, *PARP8*, *PARP9*, *PARP10*, *PARP11*, *PARP12*, *PARP13*, *PARP14*, *PARP15*, *PARP16*), ADP-ribosyl transferases (*ART1*, *ART3*, *ART4*, *ART5*)
Synthesis of cyclic adenosine monophosphate (cAMP) for intracellular signal transduction	Adenylate cyclases (*ADCY1*, *ADCY2*, *ADCY3*, *ADCY4*, *ADCY5*, *ADCY6*, *ADCY7*, *ADCY8*, *ADCY9*, *ADCY10*)
Cyclic ADP-ribose and nicotinic acid adenine dinucleotide phosphate (NAADP) formation in calcium	ADP-ribosyl cyclases (*CD38*, *BST1*)

metabolism	
Sirtuin deacetylation in p53 function and genomic stability	NAD-dependent deacetylases sirtuins 1-6 (*SIRT1*, *SIRT2*, *SIRT3*, *SIRT4*, *SIRT5*, *SIRT6*)

Appendix Table 2.5: Enzyme, enzyme complexes, or enzyme families requiring vitamin B5 (pantothenic acid) as a cofactor or substrate

Enzyme	Function
10-Formylteterahydrofolate dehydrogenase (*ALDH1L1*, *ALDH1L2*)	Oxidation of 10-formylteterahydrofolate to form tetrahydrofolate.
Cytosolic fatty acid synthase (*FASN*)	Fatty acid synthesis.
4'-Phosphopantetheine transferase (*AASDHPPT*)	Addition of 4'-phosphopantetheine to proteins.
α-Aminoadipate semialdehyde dehydrogenase (*ALDH7A1*)	Conversion of lysine to α-aminoadipate semialdehyde in lysine degradation.
Pyruvate dehydrogenase complex (*PDHA1*, *PDHA2*, *PDHB*, *PDHX*, *DLAT*, *DLD*)	Decarboxylation of pyruvate to acetyl-CoA.
α-Ketoglutarate dehydrogenase complex (*OGDH*, *DLST*, *DLD*)	Decarboxylation of α-ketoglutarate to succinyl-CoA.
Branched-chain α-keto acid dehydrogenase complex (*BCKDHA*, *BCKDHB*, *DBT*, *DLD*)	Transamination of valine, isoleucine and leucine.

Appendix Table 2.6: Enzyme, enzyme complexes, or enzyme families requiring vitamin B6 as a cofactor or substrate

Enzyme	Function
Aminotransferases	Transfer of an amino group from an amino acid to α-keto acid in nonessential amino acids synthesis. Aspartic amino transferase (*GOT1*, *GOT2*) and alanine aminotransferase (*GPT*, *GPT2*) are the common aminotransferases with vitamin B_6 as coenzymes.
Decarboxylases	Removal of carboxy group from amino acids in the formation of γ-aminobutyric acid (glutamate decarboxylase (*GAD1*, *GAD2*)), serotonin (*DDC*), histamine, dopamine and epinephrine.
Racemases (*SRR*)	Interconversion of D-amino acids and L-amino acid.
Dehydratases (*SDS*)	Removal of an amino group from amino acids.

Cystathionine β-synthase (*CBS*)	Conversion of homocysteine to cystathionine in cysteine synthesis.
Cystathionine γ-lyase (*CTH*)	Conversion of cystathionine to cysteine.
Selenocysteine β-lyase (*SCLY*)	Conversion of selenocysteine to selenide.
Serine hydroxymethyltransferase (*SHMT1*, *SHMT2*)	Transfer of hydroxyl group from serine to tetrahydrofolate (THF) for glycine generation.
δ-Aminolevulinate synthase (*ALAS1*, *ALAS2*)	Labilization of glycine and addition of succinate to generate δ-aminolevulinate, the first and rate-limiting step in heme synthesis.
Glycogen phosphorylase (*PYGM*, *PYGL*)	Degradation of glycogen to glucose-1-phosphate.
Serine palmitoyl transferase (*SPTLC1*, *SPTLC2*, *SPTLC3*)	Condensation of serine with palmitoyl-CoA in sphingolipid synthesis.
δ-6-Desaturase (*FADS2*)	Desaturation of linoleic and γ-linolenic acids.
Kynureninase (*KYNU*)	Conversion of 3-hydroxykynurenine to 3-hydroxyanthranilic acid in niacin synthesis.

Appendix Table 2.7: Enzyme, enzyme complexes, or enzyme families requiring vitamin B7 (biotin) as a cofactor or substrate

Enzyme	Function
Pyruvate carboxylase (*PC*)	Carboxylation of pyruvate to oxaloacetate in anaplerotic reaction.
Acetyl-CoA carboxylase (*ACACA*, *ACACB*)	Carboxylation of acetyl-CoA to malonyl-CoA.
Propionyl-CoA carboxylase (*PCCA*, *PCCB*)	Carboxylation of propionyl-CoA generated from catabolism of odd-chain fatty acids, isoleucine, threonine and methionine to form methylmalonyl-CoA.
β-Methylcrotonyl-CoA carboxylase (*MCCC1*, *MCCC2*)	Carboxylation of β-methylcrotonyl-CoA generated from catabolism of leucine to form β-methylglutaconyl-CoA which can be further catabolized to acetoacetate and acetyl-CoA.
Holocarboxylase synthetase (*HLCS*)	Catalysis of covalent linkage of biotin to lysine residues of carboxylases.
Sodium-dependent solute carriers (*SLC5A6*, *SLC19A3*)	Transport of biotin across the plasma membrane

Appendix Table 2.8: Enzyme, enzyme complexes, or enzyme families requiring vitamin B9 (folic acid) as a cofactor or substrate

Enzyme	Function
Glycinamide ribonucleotide	Addition of formyl groups at C8 position in 10-

formyltransferase (*GART*)	formyl THF in purine synthesis.
5-amino-4-imidazole carboxamide ribonucleotide formyltransferase (*ATIC*)	Addition of formyl groups at C2 position in 10-formyl THF in purine synthesis.
Thymidylate synthetase (*TYMS*)	Conversion of deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP), a rate-limiting step in DNA synthesis.
Serine hydroxymethyltransferase (*SHMT1*, *SHMT2*)	Conversion of THF and serine to 5,10-methylene THF and glycine.
Methionine synthase (*MTR*)	Conversion of 5-methyl THF and homocysteine to THF and methionine.
10-Formyl THF synthetase (*MTHFD1*, *MTHFD1L*)	Formylation of THF to form 10-Formyl-THF.
5,10-Methenyl THF cyclohydrolase (*MTHFD1*, *MTHFD2*)	Conversion of 10-formyl THF to 5,10-methenyl THF.
5,10-Methylene-THF dehydrogenase (*MTHFD1*, *MTHFD2*)	Conversion of 5,10-methenyl THF to 5,10-methylene THF.
Dihydrofolate reductase (*DHFR*, *DHFR2*)	Conversion of folic acid to dihydrofolate. Conversion of dihydrofolate to THF.
Methylene THF reductase (*MTHFR*)	Conversion of 5,10-methylene THF to 5-methyl THF

Appendix Table 2.9: Enzyme, enzyme complexes, or enzyme families requiring vitamin B12 (cobalamin) as a cofactor or substrate

Enzyme	Function
Methionine synthase (*MTR*)	Conversion of homocysteine to methionine in the cytosol with methylcobalamin as a cofactor.
Methylmalonyl-CoA mutase (*MUT*)	Isomerization of methylmalonyl-CoA to succinyl-CoA to catabolize odd-chained fatty acids and cholesterol in the mitochondria with adenosylcobalamin as a cofactor.

Appendix Table 2.10: Enzyme, enzyme complexes, or enzyme families requiring vitamin C (ascorbic acid) as a cofactor or substrate

Enzyme	Function
Dopamine β-hydroxylase (*DBH*)	Hydroxylation of dopamine to form norepinephrine.
Peptidylglycine α-amidating monooxygenase (*PAM*)	Addition of carboxy-terminal α-amide group to activate peptide hormones.
Prolyl 3-hydroxylase (*P3H1*, *P3H2*, *P3H3*)	Hydroxylation of prolyl and lysyl residues in collagen synthesis.
Prolyl 4-hydroxylase (*P4HA1*, *P4HA2*,	

P4HA3, *P4HB*, *P4HTM*)	
Lysyl hydroxylase (*PLOD1*, *PLOD2*, *PLOD3*)	
Hypoxia inducible factor prolyl hydroxylases (*EGLN1*, *EGLN2*, *EGLN3*)	Hydroxylation of hypoxia inducible factor 1 subunit alpha (HIF1A) protein.
Tyrosine hydroxylase (*TH*)	Conversion of tyrosine to L-DOPA in catecholamine biosynthesis.
Trimethyllysine hydroxylase (*TMLH*)	Conversion of trimethyllysine to hydroxy trimethyllysine in carnitine biosynthesis.
γ-Butyrobetaine hydroxylase (*BBOX*)	Conversion of γ-butyrobetaine to carnitine in carnitine biosynthesis.
4-Hydroxyphenylpyruvate dioxygenase (*HPD*)	Conversion of 4-hydroxyphenylpyruvate to homogentisate in tyrosine catabolism.
Sodium-dependent vitamin C transporters (*SLC23A1*, *SLC23A2*)	Absorption of ascorbic acid at the intestinal brush border and distribution to tissues.
Glucose transporters (*SLC2A1*, *SLC2A3*, *SLC2A4*)	Absorption of oxidized ascorbic acid at the gastrointestinal tract and distribution to tissues.
Carnitine-acylcarnitine translocase (*SLC25A20*)	Transport of carnitine fatty acid complexes into inner mitochondria.

Appendix Table 2.11: Enzyme, enzyme complexes, or enzyme families requiring vitamin D as a cofactor or substrate

Enzyme	Function
Vitamin D binding protein (*GC*)	Transport of vitamin D metabolites in the blood to tissues.
Protein disulfide isomerase A3 (*PDIA3*)	Regulation of rapid membrane-associated signaling pathway of vitamin D.
1α-hydroxylase (*CYP27B1*)	Conversion of 25(OH)D to $1,25(OH)_2D_3$ in the kidney.
24-hydroxylase (*CYP24A1*)	Degradation of $1,25(OH)_2D_3$ and 25(OH)D to inactive forms.
25-hydroxylase (*CYP2R1*)	Conversion of vitamin D_3 to $25(OH)D_3$.

Appendix Table 2.12: Enzyme, enzyme complexes, or enzyme families requiring vitamin E as a cofactor or substrate

Enzyme	Function
α-tocopherol transfer protein (*TTPA*)	Transfer of RRR α-tocopherol and to very low-density lipoprotein (VLDL) for distribution to tissues.

Appendix Table 2.13: Enzyme, enzyme complexes, or enzyme families requiring vitamin K as a cofactor or substrate

Enzyme	Function
γ-Glutamyl carboxylase (*GGCX*)	Posttranslational carboxylation of glutamic acid residues to form γ-carboxy-glutamate required for coagulation and bone mineralization.

Appendix to Chapter 3

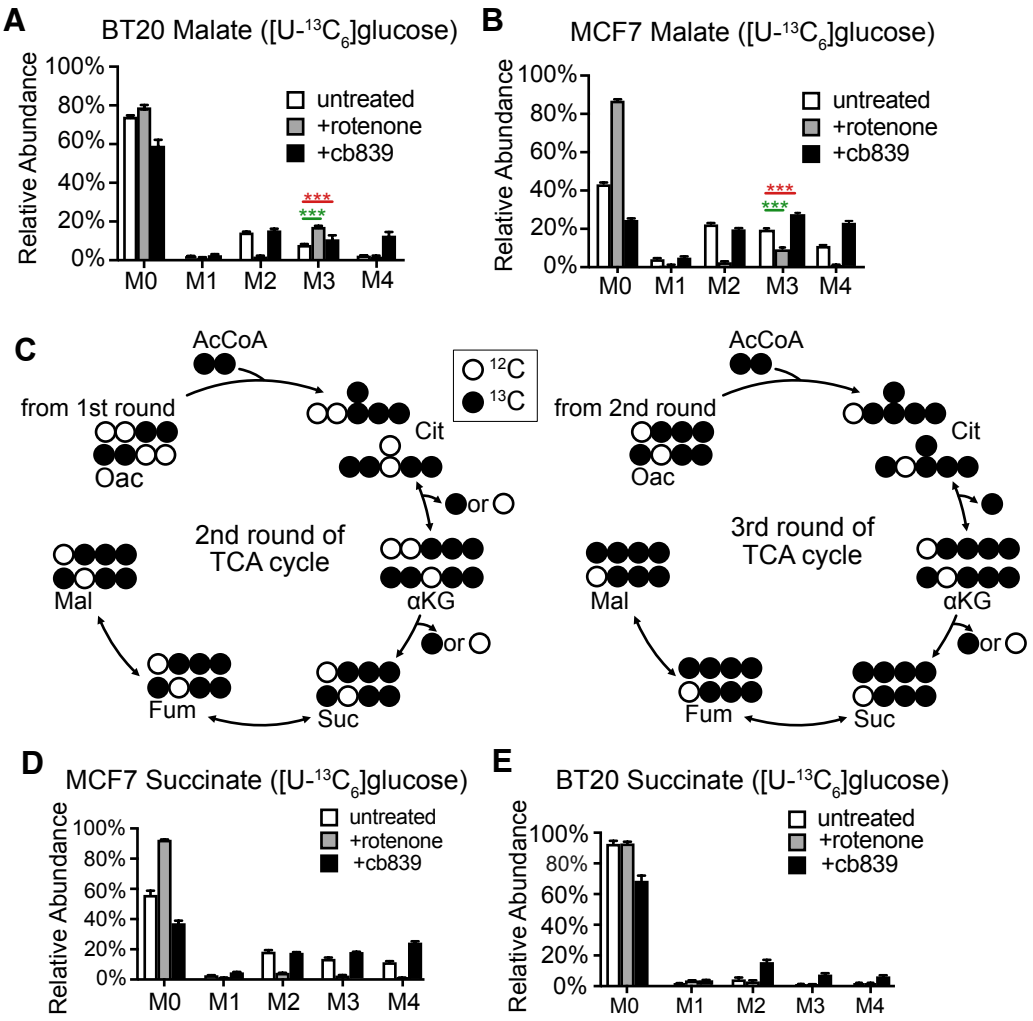
Abbreviations

MFA	Metabolic Flux Analysis
MID	Mass Isotopologue Distribution
TCA	Tricarboxylic Acid
IMD	Isotope Mass Distribution

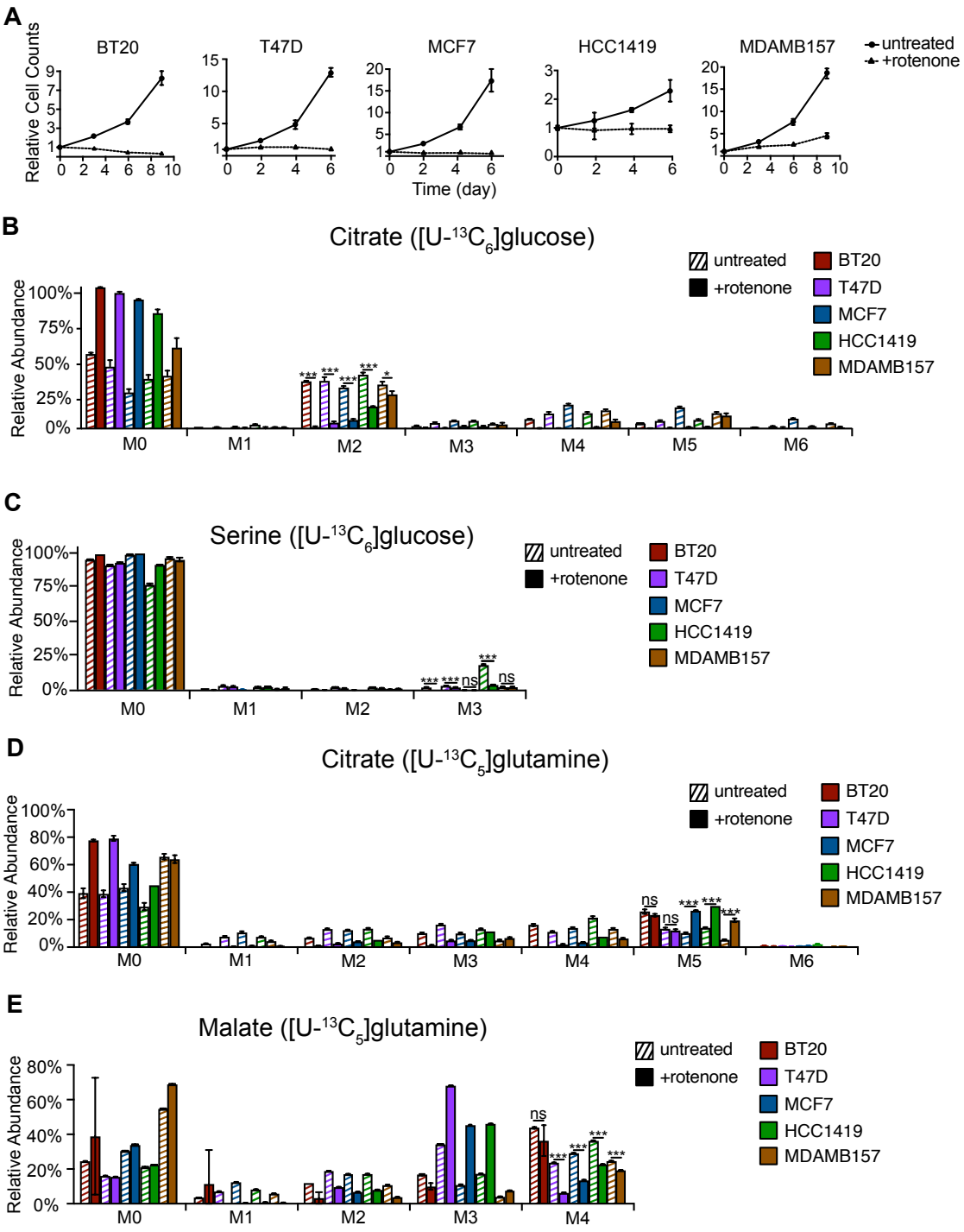
Appendix to Chapter 4

Abbreviations

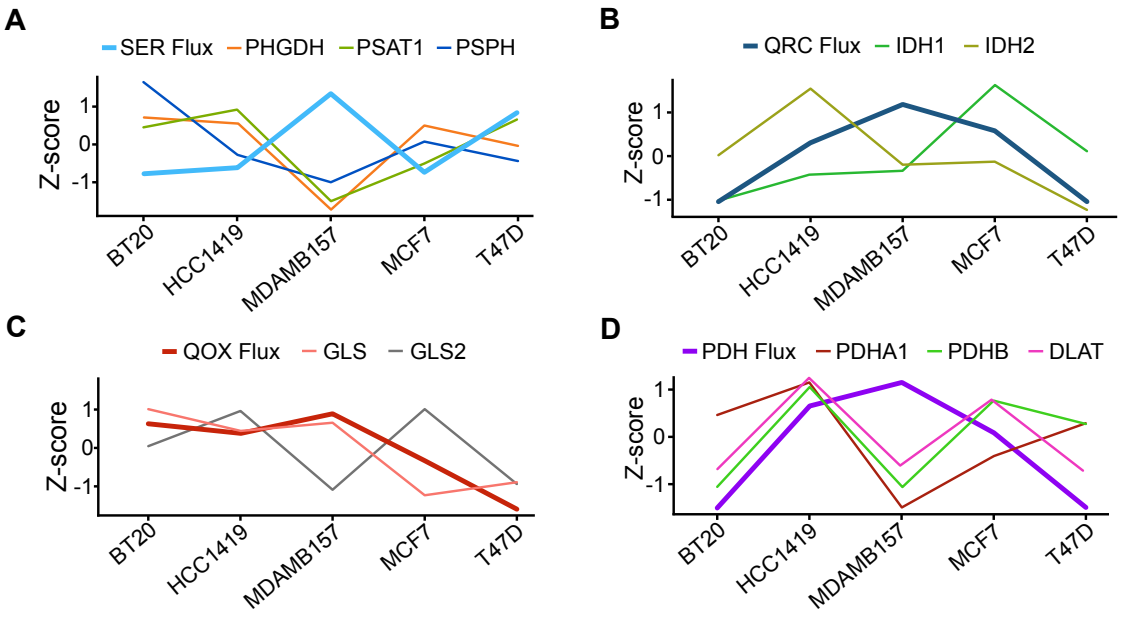
OXPHOS	Oxidative Phosphorylation
TCA	Tricarboxylic Acid
ATP	Adenosine Triphosphate
PC	Pyruvate Carboxylase
MFA	Metabolic Flux Analysis
PDH	Pyruvate Dehydrogenase
SDH	Succinate Dehydrogenase
MIDs	Mass Isotopologue Distributions
PPP	Pentose Phosphate Pathway
t-SNE	t-Distributed Stochastic Neighbor Embedding
BIRC5	Baculoviral IAP Repeat Containing 5
GTSE1	G2 and S-Phase Expressed 1
GLS	Glutaminase
GLS2	Glutaminase 2



Appendix Figure 4.1. Labeling of TCA cycle intermediates by [U-$^{13}C_6$]glucose. (A) Malate MID of BT20 cells cultured with [U-$^{13}C_6$]glucose. (B) Malate MID of MCF7 cells cultured with [U-$^{13}C_6$]glucose. (C) Schematic of [U-$^{13}C_6$]glucose labeling of carbon atoms in the second and third turns of the TCA cycle. (D) Succinate MID of MCF7 cells cultured with [U-$^{13}C_6$]glucose. (E) Succinate MID of BT20 cells cultured with [U-$^{13}C_6$]glucose. Error bars represent standard deviations. *p<0.05, **p<0.01, and ***p<0.001 by ANOVA with Tukey's post hoc test.



Appendix Figure 4.2. Proliferation and mass isotopologue distributions.
(A) Number of viable cells quantified versus time, normalized to initial time point. (B and C) MID of citrate (B) and serine (C) in breast cancer cell lines cultured with [U-$^{13}C_6$]glucose. (D and E) MID of citrate (D) and malate (E) in breast cancer cell lines cultured with [U-$^{13}C_5$]glutamine. Error bars represent standard deviations. *p<0.05, **p<0.01, and ***p<0.001 by a two-tailed, equal variance, Student's t test.



Appendix Figure 4.3. Correlations of flux responses with enzyme abundance responses upon inhibited OXPHOS. (A) Adaptations of *de novo* serine synthesis flux and abundances of enzymes catalyzing steps in the pathway upon rotenone treatment. (B) Adaptations of glutamine reductive carboxylation flux and abundances of isocitrate dehydrogenase isoforms upon rotenone treatment. (C) Adaptations of glutamine oxidation flux and abundances of glutaminase isoforms upon rotenone treatment. (D) Adaptations of pyruvate dehydrogenase flux and pyruvate dehydrogenase complex members upon rotenone treatment. PDH Flux: pyruvate oxidation flux, QOX Flux: glutamine oxidation flux, QRC Flux: glutamine reductive carboxylation flux, SER Flux: de novo serine synthesis flux.

Appendix Table 4.1. Enrichr pathway analysis of protein ΔAbundances correlated with ΔFluxes, +/- rotenone. Enrichment in KEGG pathways. Pathways with adjusted p-values less than 0.01 are displayed.

Flux	Correlated Gene Symbols	KEGG Pathway	Correlated Gene Symbols in Pathway	Adjusted P-value	Odds Ratio
QOX	C1orf122, C4orf19, TXLNG, SGMS2, NUDT22, RNF187, PARP1, TFDP1, AZI2, POLG2, DTL, TOP2A, TMEM206, CENPF, GTPBP3, BUB1, UNC45A, PRIM1, TTI1, REST, TELO2, PLCE1, BIRC5, PSMG3, PSMG4, CUL4A, SLC45A4, SOS1, CDC123, CABLES1, ELP3, LARGE2, PRC1, RRP12, RHPN2, TRIP13, MAL2, SPATA5L1, GMNN, COQ7, ZFP1, DLG5, MRPS23, SKA1, DPY19L3, ARMC7, PRIM2, XPNPEP3, CAMK2D, KRI1, POLE2, SMC2, ATM, COQ5, TIMELESS, PPIH, BUB1B, CBX2, NFIB, EI24, KIF20B, MNAT1, NCAPD3, AAAS, PHKA1, CCDC93, PMS1, DBF4, ZWINT, MZT1, HOOK1, CDCA2, TIA1, CDC25C, CNKSR1, C2CD2L, AP3M2, ATP1A3, SUOX, CCDC120, WDR60, ANAPC13, ZYG11B, CLDND1, WRB, DDX11, RAD51, UBE2G2, KIF20A, ARID4A, LIN37, PLK1, POLA1, GTF2H4, MRPS7, ATR, NBEAL2, RFC5, HELLS, SHCBP1	Cell Cycle	ANAPC13, TFDP1, DBF4, PLK1, BUB1B, ATM, CDC25C, BUB1, ATR	1.29E-06	17.02
		DNAReplication	RFC5, PRIM2, POLA1, PRIM1, POLE2	5.09E-05	33.73
		Nucleotide Excision Repair	RFC5, CUL4A, POLE2, MNAT1, GTF2H4	1.32E-04	24.88
QRC	TRIAP1, RPL32, GTF2A2, IST1, PPOX, TRABD, RPUSD4, TIA1, KCTD5, UFC1, NDUFA2, EIF4A3, PSMB1, CXorf38, ZNF766, NUP37, RNF126, SMIM26, CGREF1, CENPI, ZYG11B, PSMA4, NSMCE4A, TTI1, ZNF513, CLCC1, VPS4A, NDUFV1, RUVBL2, SRP9, DDX39B, AAAS, INA, SAE1, GTPBP3, PPARG, POLR1D, COQ7, POLR3E, MAGEA12, APPBP2 SENP6, IFT27, URB2, FAM210B, BABAM1, CNOT2, FGD3, PES1, ACTR6, ZNF84, MGME1, EXOSC6, MRPS7, NMRAL1, PCIF1, SPC25, PSMA7, RAD9A, NSMCE1, RDH10, UBE2G2, POLR3D, TCP1, NDUFV2, METTL21A, CCDC6, SOS1, GTF2H1, CTDP1, SF3B5, LEMD3, RPL7A, NUBP1, MRPL47, DDX18, RBM45, SMC4, FOXA1, TECR, MARCKSL1, AGGF1, PAIP2B, IDH3G, ZNF517, POLR3A, PHLPP2, NDUFS1, IRF3, RNF41 POLR2A, PPP4R1, HARS, DCTN5, PPIG, FARP2, MGA, FAM199X, CLNS1A, UBE3A	RNA Polymerase	POLR3A, POLR2A, POLR3D, POLR3E, POLR1D	4.45E-05	38.74
		Cytosolic DNA-Sensing	POLR3A, IRF3, POLR3D, POLR3E, POLR1D	7.42E-04	17.7
		Pyrimidine Metabolism	POLR3A, POLR2A, POLR3D, POLR3E, POLR1D	5.33E-03	10.42
		Huntington's Disease	POLR2A, NDUFA2, NDUFS1, PPARG, NDUFV2, NDUFV1	9.13E-03	6.73
		Epstein-Barr Virus Infection	POLR3A, IRF3, POLR2A, POLR3D, POLR3E, POLR1D	9.03E-03	6.42
PC Flux	CENPF, BIRC5, TOP2A, GTSE1, PLK1, RAD51, PRC1, ELP4, CDCA8, METTL6, C2CD5, CDCA2, FIBP, B4GAT1, MELTF, FAM83D, COQ5, UBE2C, ANAPC13, KIF20A, SUOX, CHAF1B, ANLN, TOB1, TELO2, SLC6A11, POLQ, SOS1, DST, GIT2, ABCF2, TPX2, LIN37, RACGAP1, TMEM206, DTL, SPATA5L1, ESPL1, KIF1BP, PBK, AURKB, GMNN, SPATA5, ZWINT, TXLNG, GSS, PMS1, CAPN15, FAM120B, DBF4, CCNB2, UBE2S, KIF11, PRIM2, SASS6, ARMC7, NR3C1, CCNA2, PRIM1, PSMG3, WRB, HOOK1, KIF15, CLDND1, TRIP13, SGMS2, SLC19A1, C1orf122, CSF2RA, CENPU, KNTC1, RAD54B, CCNB1, KIF20B, C4orf19, CCDC93, KIF18B, ZNF554, RAB30, WWC3, WSB1, HMMR, POLG2, TONSL, SNX21, INCENP, MZT1, TIMELESS, KIF23, NUDT22, TIPIN, TFDP1, PIMREG ELP3, CENPE, AURKA, BUB1B, NCAPD3, CHAF1A, TK1, PDE3B, PRKCE, NDC80, GMPS, PIGH, CKAP2, NBEAL2, NCK1, FAM98A, PHKA1, RBM23, EIF3K, UBE2T, DNAJC14, HASPIN, NUSAP1, HELLS, ELL3, ARHGAP24, MASTL, PPIH, IQSEC2, GEN1, POLA2, KLHL13, RNF187, TRMT6, PRPSAP1, PARP1, NFKBIL1, USP1, REEP4, ZNF720, MCM10, COQ7, NUF2, ESCO2, TICRR, CCNH, ATR, TUBA8, POLA1, BUB1, UBR7, UNC45A, ZER1, EXOC4, GCLC, LRR1, GREB1, C2CD2L, CASP8AP2, ATAD2, INHBB, SMG5, BUD23, CDC20, KIF22, BCAM, VRK1, BRMS1L, PLCE1, CETN3, CDK2, BTBD2, MYO9A, MFAP3L, CEP55, SMC2, MCM2, SEH1L, RPA2, MNAT1, CABLES1, SPPL2B, MGMT, ZNF580, RPL23, DIAPH3, NCAPH, MCM4, WDHD1, UHRF1, SHCBP1, ELP5,HDAC11, CEP78, MAGEA10	Cell Cycle	ANAPC13, CCNH, PLK1, BUB1B, CDC20, CCNA2, CCNB2, CCNB1, TFDP1, DBF4, ESPL1, CDK2, MCM4, BUB1, ATR, MCM2	4.68E-12	16.97
		DNAReplication	PRIM2, POLA1, POLA2, PRIM1, RPA2, MCM4, MCM2	2.06E-06	26.38
		Oocyte Meiosis	CDC20, ANAPC13, CCNB2, CCNB1, ESPL1, PLK1, CDK2, BUB1, AURKA	8.84E-05	8.69
		Progesterone-Mediated Oocyte Maturation	CCNA2, ANAPC13, CCNB2, CCNB1, PLK1, CDK2, PDE3B, BUB1	1.07E-04	9.74
		Fanconi Anemia Pathway	RAD51, UBE2T, USP1, RPA2, TELO2, ATR	2.17E-04	13.86
		p53 Signaling Pathway	CCNB2, CCNB1, CDK2, GTSE1, ATR	8.57E-03	8.43
SER	CPSF2, ZBTB5, DDX18, OIP5, PHF20, MBOAT2, FOXP4, SYTL2, ING3, KANSL2, CENPH, KDM5A, SMC4, TOPBP1, CHD4, UHRF1, NCAPH, TWISTNB, VWA1, BUD23, CENPI, EEF2KMT, ANKLE2, SIPA1L1, SUPT16H, FANCD2, RACK1, PRPF4B, PPP4R3B, GADD45GIP1, MMS22L, RRP36, NCAPD2, NCAPG, NSD1, ECD, NUP155, SMC2, MDC1, WDR34, PODXL, AES, BRD9, MPHOSPH9, GTF2I, RPS14, MCM7, CHAF1A, EP400, PNO1, POLR2B, SPC25, NCAPH2, EZH2, GID8, ANKRD2, SLC2A10, PARD3, ABCB6, ULBP2, TEX2, MBD3, PRRC2C, KIF4B, TAF9B, EPC1, APOB, MIPOL1, MGA, MCM3, AKT3, CASP8AP2, DPY30, PTPN18, BAG3, SLCO4A1, MAEA, POLR2E, MSH6, POLD3, MASTL, FMN1, DHX16, CIAPIN1, HELLS, LSG1, CTDSPL2, TRIR, NR2F6, CSDE1, DDX3X, SNCB, NRDE2, SF3B3, ASTE1, MOCS1, CASP2, EIF3M, NACA, ATAD2B, MARVELD2, TOP3A, MCM5, MCM2, ARID5B, POMK, SLPI, SARAF, CKLF, CDKL2, ERO1B, MRPS30, POLA1, FANCM, L3MBTL2, WDHD1, BYSL, EPS8L1, GK5, NLK, HASPIN, TSR1, MCM4, TUBA8, GINS2, SMARCD1, TIPIN, SYK, PHF19, WDR5, RPL5, FANCF, TRAF3IP2, FAM50A, SYBU, TRMT61A, PTTG1, ASB7, DHFR, TJP1, SNX19, USP12, VSIG10, SAMD1, ANXA10, PBK, TONSL, CDK19, FGF2, MAP3K11, LRRC26	DNA Replication	POLD3, POLA1, MCM7, MCM3, MCM4, MCM5, MCM2	9.88E-07	33.22
PDH	MRPL47, RPUSD4, MRPL19, PHLPP2, INA, DDX18, URB2, NSMCE1, CENPI, SENP6, SF3B3, CLCC1, MARCKSL1, CNOT2, DNAJA3, ISCA1, SPTBN4, NFU1, MGA, CHD5, ZNF513, DDX3X, TIA1, PIGS, RDH10, NSMCE4A, EIF4A3, NUP155, URB1, SMIM26, NDUFA2, SREK1, APOBEC3C, WDR5, MRPL33, GPRC5B, MAGEA12, SDHA, NDUFS1, VPS4A, APPBP2, STRBP, COX5B, ADGRE5, TRIAP1, IST1, DBP, SAE1, NDUFS7, NUP37, SPC25, RRP36, GTF3C2, RPL32, FGD3, SPPL2A, MRPS7, MGLL, LYRM2, TFB2M, LPCAT4, PARD3, SMC4, GTF2A2, FAM192A, NDUFV2, RFFL, BCL7A, PARP4, RPP30, HNRNPD, CENPH, METTL21A, POLD1, C15orf61, EIF2AK3, RFC4, EGFR, CEP290, EEF2KMT, CHERP, SLIRP, PPP4R1, PYURF, ASXL1, ITGB6, IDH3G, MGME1, FAM208B, RBM3, DAP3, IP6K1, FARP2, RAD9A, RFC5, COPG2, NIPSNAP2, MRPS9, GHITM, TUBB3	Non-Alcoholic Fatty Liver Disease	NDUFS7, NDUFA2, EIF2AK3, NDUFS1, SDHA, NDUFV2, COX5B	9.63E-04	10.33
		Alzheimer's Disease	NDUFS7, NDUFA2, EIF2AK3, NDUFS1, SDHA, NDUFV2, COX5B	9.63E-04	9.23
		Oxidative Phosphorylation	NDUFS7, NDUFA2, NDUFS1, SDHA, NDUFV2, COX5B	1.64E-03	9.94
		Parkinson's Disease	NDUFS7, NDUFA2, NDUFS1, SDHA, NDUFV2, COX5B	1.77E-03	9.28
		Mismatch Repair	RFC5, RFC4, POLD1	3.56E-03	30.74
		Huntington's Disease	NDUFS7, NDUFA2, NDUFS1, SDHA, NDUFV2, COX5B	6.23E-03	6.73
		Ribosome	MRPS9, RPL32, MRPL19, MRPS7, MRPL33	7.97E-03	7.88
		DNA Replication	RFC5, RFC4, POLD1		
Growth	DDX18, CENPI, SF3B3, MRPL19, SENP6, MGA, NUP155, CENPH, DDX3X, SMC4, RRP36, RPUSD4, CHD4, MRPL47, EEF2KMT, SPC25, ZBTB5, INA, SPTBN4, PHLPP2, URB2, MBOAT2, PARD3, CNOT2, WDR5, EIF4A3, CPSF2, SMC2, POLD1, NSMCE1, APOBEC3C, EHMT2, ASXL1, CHERP, POLD3, GINS2, FOXP4, ZDBF2, NSMCE4A, SDHA, MPHOSPH9, TWISTNB, WDR34, LSG1, VWA1, PPP4R3B, SUPT16H, GPRC5B, HNRNPD, MIS12, TSPAN15, SMPD4, SMIM26, TIA1, RBM45, SREBF1, APPBP2, SPPL2A, FAM50A, POMK, MRPL33, PNO1, BRD9, SARAF, RBM3, SLPI, PTTG1, DPY30, FARS2, POLA1, MRPL35, IP6K1, RCL1, DNAJA3, FAM208B, OIP5, PARP4, RSBN1, GK5, GTF2I, AGGF1, TEX2, MGME1, ZNF574, RFC5, FANCD2, CDK19, FAM192A, DHFR, CHD5, WDCP, NDUFS7, ANKLE2, B3GNT3, USP12, MRPS30, SLCO4A1, ARHGAP35, MARCKSL1, ABCG2	DNA Replication	POLD3, RFC5, POLA1, POLD1		
		Mismatch Repair	RFC5, POLD3, POLD1		