Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Kristine K. Dennis Date

Health Benefits of Plant-based Diets:
Effects of Dietary Phytochelatins on Toxic Metal Absorption

By

Kristine K. Dennis
Doctor of Philosophy

Nutrition and Health Sciences

_____
Dean P. Jones, Ph.D.
Advisor

_____
Jessica A. Alvarez, Ph.D., R.D.
Committee Member

_____
Terryl J. Hartman, Ph.D., R.D.
Committee Member

_____
Karan Uppal, Ph.D.
Committee Member

_____
Miriam B. Vos, M.D., M.S.P.H.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Health Benefits of Plant-based Diets:
Effects of Dietary Phytochelatins on Toxic Metal Absorption

By

Kristine K. Dennis
B.A., Luther College, 2009
M.P.H., Emory University, 2012

Advisor: Dean P. Jones, Ph.D.

An abstract of
a dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Nutrition and Health Sciences

2020

Abstract

Health Benefits of Plant-based Diets:
Effects of Dietary Phytochelatins on Toxic Metal Absorption

By Kristine K. Dennis

Plant-based dietary patterns reduce risks of chronic disease, including cardiovascular disease, type 2 diabetes, and chronic kidney disease. The mechanisms through which plant-based diets protect health are not fully understood. Phytochelatins (PyCs), metal-binding compounds produced by plants, may be beneficial through reducing absorption of toxic dietary metals, such as cadmium (Cd). Cd accumulates in the body and can lead to increased chronic disease risk at current dietary levels. The interaction of PyCs with Cd may be one mechanism through which plant-based diets confer health benefits. However, dietary PyC types and concentrations are unknown, and the functional effects of these compounds are minimally understood. The purpose of this dissertation was to establish foundational understanding of dietary PyCs and how these compounds may protect from Cd exposures.

The present body of work examines the role of PyCs in human health using molecular and population-based approaches. In Aim 1, we created a database of over 46,000 exact masses of PyC and PyC-metal complexes, establishing a tool to facilitate PyC research in biological systems using high-resolution mass spectrometry. The second aim examined PyC concentrations in twenty commonly consumed plant foods. PyCs were quantified in all food types with PyC_2-Gly the most common. In Aim 3, PyC_2-Gly absorption and function was investigated in an intestinal epithelial cell model by co-treating cells with PyC_2-Gly and Cd at expected dietary concentrations. We discovered PyC_2-Gly is absorbed into and through the intestinal cell monolayer and may reduce Cd absorption. The fourth aim investigated if intake of metal-binding plant compounds, as measured via plant food score (PFS), was associated with lower Cd body burden. In a cross-sectional study in the REGARDS cohort, higher PFS was associated with lower Cd body burdens in middle-aged but not older adults. In the closing chapter, I provide my perspective on the potential implications of this research and explore remaining open questions and future directions.

This dissertation demonstrates PyCs are widely distributed in the human diet and may protect from toxic metal exposures, adding to our understanding of the potential mechanisms through which plant-based diets promote health.

Health Benefits of Plant-based Diets:
Effects of Dietary Phytochelatins on Toxic Metal Absorption

By

Kristine K. Dennis
B.A., Luther College, 2009
M.P.H., Emory University, 2012

Advisor: Dean P. Jones, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Nutrition and Health Sciences

2020

ACKNOWLEDGEMENTS

The journey to completing this dissertation began long before it officially started. To my parents, Terry and Janet, thank you for encouraging my curiosity in science and supporting my ability to see that I could be a part of it. I am grateful for your support throughout my educational endeavors. To my sister, Ashley, you always see the best in me. Your cheerleading has helped me through this marathon. Thanks for your perspective, support, and guidance throughout this PhD.

My scientific development is the product of so many contributions. To my committee, Drs. Jessica Alvarez, Terry Hartman, Karan Uppal, and Miriam Vos, thanks for your thoughtful comments and conversations that have made this dissertation stronger and made me a better scientist. To my advisor, Dr. Dean Jones, thank you for your mentorship starting from our first time working together within HERCULES. You always challenge me to think broadly and reach into the unknown. Thanks for sharing your perspectives on science and life and your enthusiasm for research.

My thanks to everyone in the NHS program. To all the faculty, your feedback at seminars, your breadth of nutrition knowledge, and your commitment to this program have made me a more well-rounded nutrition scientist. To the NHS students, thanks for creating a community for this PhD journey. To my cohort, you all are amazing! Thanks for your friendship and support through these four plus years.

To the family and friends in my life, near and far, thank you for believing in me. Mackenzie and Sharia, you have been along this journey since I considered going “back to school” for my PhD. Thanks for listening, providing perspectives, and keeping me sane.

Perry, I am so lucky to have you as my partner in life. It is impossible to capture my gratitude for your support throughout this PhD. Thanks for your support in the day-to-day (e.g. taking on the majority of the travel burden when we lived in two different cities, making sure I have excellent coffee and delicious food to keep me going, etc.), for your endless enthusiasm and curiosity about my ideas, and for the light and joy you bring to everyday life. And finally, to our ray of sunshine, Finn the cat, thanks for literally being by my side as I write up this dissertation during a very strange year.

For all of you, I am forever grateful for your influence on my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Plant-based dietary patterns and chronic disease prevention

Poor diet is a major modifiable risk factor for chronic disease, including cardiovascular disease, cancer, and type 2 diabetes (1). Individual foods and nutrients impact chronic disease risk. For example, both red and processed meat are associated with increased risk of colon cancer (2-4) and mortality from cardiovascular disease and cancer (5), while processed meat consumption is associated with higher incidence of coronary heart disease and type 2 diabetes (6, 7). In contrast, diets higher in fruits and vegetables are inversely associated with cardiovascular disease and type 2 diabetes risk (8-10).

For chronic disease prevention, evidence supports plant-based diets, defined broadly as diets with high quantities of vegetables, fruits, whole grains, nuts, seeds, and legumes and low quantities of animal foods, for reducing risks of cardiovascular disease, type 2 diabetes, chronic kidney disease, and some cancers (11-14). Plant-based dietary patterns include vegetarian and vegan diets along with diets which have limited intake of animal products. For example, the Mediterranean dietary pattern emphasizes plant foods while including moderate to high fish intake and low dairy and meat intake. People with greater adherence to Mediterranean diets and other plant-based diets have reduced risks of a wide range of chronic diseases (12, 13, 15).

Some of the mechanisms through which plant-based diets confer health benefits have been explored, but a full understanding of the health benefits of plant-based diets is not available. Specific compounds such as lignans, plant sterols, and polyphenols have been linked to disease pathways. For example, dietary lignans can modulate inflammatory and apoptotic pathways and cell proliferation due to their antioxidant, anti-estrogenic, and anti-cancer properties (16, 17). In a controlled study, women consuming a lignan commonly found in flaxseed, secoisolariciresinol diglucoside, had reduced cell proliferation in benign breast tissue (18). In patients recently diagnosed with breast cancer, women consuming 25 g/day of flaxseed had reduced tumor biomarkers (19). Other studies focus on the protective mechanisms linked to overall dietary patterns higher in plant foods. In studies of chronic kidney disease (CKD), research suggests lower dietary acid load through proportionally higher consumption of fruits and vegetables relative to animal products reduces incident CKD and CKD progression (20-22). Other studies have found higher fiber intake, particularly from legumes and vegetables, reduces CKD risk (23-25). Plant-based dietary patterns bring together a range of phytochemicals which contribute to reduced chronic disease risks through specific biological mechanisms.

Dietary exposome – impact of interacting dietary exposures

Challenges to understanding mechanisms underlying the health benefits of plant-based diets exist due to the human diet being a complex exposure, consisting of nutrient and non-nutrient components. The interaction of these components and the related biological response ultimately contributes to health or disease across the lifespan. The dietary exposome is a concept encompassing the cumulative impact of exposures occurring through the human diet across the lifetime (**Figure 1.1**) (26, 27). Dietary patterns capture this complexity by inherently

encompassing the interactions of individual food components, such as phytochemicals, fiber, and minerals. Additionally, environmental exposures occur via diet, such as mercury (Hg) in fish, polycyclic aromatic hydrocarbons (PAHs) in charbroiled foods, and cadmium (Cd) in leafy greens and shellfish. The interaction of these nutritional and environmental exposures in the diet influences the resulting impact on health and disease processes.

Importantly, foods derived from both plants and animals can contain dietary toxic metals, such as lead, Hg, and Cd, which contribute to increased disease risks. Cd is a particular metal of interest as dietary Cd is the greatest source of Cd exposure in non-smokers in the U.S. population. Dietary Cd exposures occur from a wide range of foods, including shellfish, organ meats, leafy greens, cereals, and bread products (28, 29). Additionally, Cd has a long half-life of 10 to 30 years, resulting in Cd accumulation in the body over the lifespan (29). Cd body burden is associated with increased risks of kidney dysfunction, type 2 diabetes, and cancer, among others (30). As chronic disease continues to increase in the U.S., Cd-related disease processes may contribute to increased disease burdens. For example, individuals experiencing reduced kidney function related to type 2 diabetes may suffer more rapid decline if they also have a high Cd body burden. Strategies to mitigate dietary Cd absorption could reduce risks of long-term health consequences.

Plants contain naturally occurring metal-binding compounds which could protect against dietary toxic metals such as Cd. Phytochelatins (PyCs) are a class of metal-chelating plant compounds produced widely in plants which could contribute to the health benefits of plant-based diets

through their interactions with toxic dietary metals (31-33). PyCs have high metal-binding capacity due to γ-glutamyl-cysteine peptides which repeat from 2-11 times (34, 35). However, the PyC types and concentrations in the human diet are unknown.

In this dissertation, we focus on characterizing dietary PyCs and understanding how PyCs may prevent absorption of Cd from the diet. We centralize our research questions on preventing risks related to Cd absorption due to its wide distribution in the human diet, the health risks of lifelong exposure, and the potential to mitigate absorption through interactions with phytochemical compounds concentrated in plant-based dietary patterns.

Overall goals and specific aims

The goal of this dissertation was to gain insight into how plant-based dietary patterns may protect from dietary toxic metal absorption. We approached this through investigating PyCs, metal-binding plant compounds found in the human diet, using molecular and population-based approaches to address the following four aims, considered in successive chapters.

Specific Aim 1

To create a database of predicted PyC and PyC-metal complexes of common metals and metalloids for use with mass spectrometry analyses.

In Chapter 2, titled *Phytochelatin database (PyCDB) –a resource for phytochelatin complexes of nutritional and environmental metals*, we created a database of 46,260 PyCs and

PyC-metal complexes using 13 metals of nutritional and toxicological significance. Due to the multiple PyC forms which exist and chemical properties of PyCs and metals, PyCs have a wide range of possible structures. Previous research of PyCs focus on a few PyC or PyC-metal complexes. With advances in ultra-high resolution mass spectrometry techniques, characterization of a large number of PyCs and PyC-metal complexes in a single sample analysis is possible. However, no resource was available for high-throughput assessment of mass spectrometry data for PyCs. Here we created a database to directly query plant and food mass spectrometry data for accurate mass annotation of PyCs. Additionally, we established a workflow for using the database, including an online querying system for the PyCDB.

Specific Aim 2

Determine the relative abundance of phytochelatins in five groups of commonly consumed plant foods.

In Chapter 3, *Distribution of phytochelatins, metal-binding compounds, in plant foods: a survey of commonly consumed fruits, vegetables, grains, and legumes*, we investigated the PyC types and concentrations in commonly consumed plant foods by food group, food type, processing level, and growing conditions using liquid chromatography-mass spectrometry. Prior research demonstrated PyCs are widely produced across the plant kingdom (35, 36). However, PyCs have not been studied in commonly consumed foods. Given the metal-binding properties of PyCs, dietary PyCs could have implications for metal absorption, redox systems, and metal homeostasis in humans (32, 33, 37, 38). Through this project, we establish an understanding of PyC concentrations in the human diet for the first time, providing a foundation for future research on the biological mechanisms and consequences of PyC-metal interactions.

Specific Aim 3

Investigate the impact of PyC_2-Gly concentrations on Cd and PyC_2-Gly uptake and transcellular transport in a human intestinal cell model.

In Chapter 4, *Transport and function of a common dietary phytochelatin, PyC_2-Gly, in intestinal epithelial cells*, we used a cell model system to study PyC_2-Gly absorption through an intestinal epithelial cell monolayer at dietary concentrations established in Chapter 3. Additionally, we investigated if Cd uptake into cells was reduced when co-exposed to dietary PyC_2-Gly and Cd concentrations. Previous studies demonstrated transport of PyC_3-Gly into intestinal epithelial cells and reduced Cd accumulation in cells co-treated with PyC_3-Ser (33, 39). However, PyC_2-Gly intestinal epithelial transport and function remains a knowledge gap in the literature. This chapter addresses this gap by examining the absorption and impact of two doses of PyC_2-Gly and Cd in combination and alone in Caco-2 epithelial cell monolayers.

Specific Aim 4

Determine if intake of metal-binding plant compounds, as measured by a plant food score, is inversely associated with cadmium body burden in a cross-sectional sample of adults.

In Chapter 5, *Plant food intake is associated with lower cadmium body burden in middle-aged adults*, we examined the association between a plant food score and urinary Cd in the REasons for Geographic and Racial Differences in Stroke (REGARDS) cohort. Prior studies have found variation in Cd bioavailability is not fully explained by known factors (40, 41). Metal-binding plant compounds could reduce absorption of Cd but cannot be directly assessed with dietary recall information in population studies (42-45). We created a plant food score to

serve as a proxy for consumption of metal-binding phytochemicals via plant-derived beverages and foods. Using food frequency questionnaire information and an established measure of Cd body burden, this chapter describes the creation and use of our plant food score to investigate another potential benefit of plant-based dietary patterns – reduced Cd body burden.

In the final chapter, I summarize these findings to provide new understanding of the potential of PyCs in plant-based diets to provide a protective mechanism against disease processes involving toxic metals. I further present a perspective on critical knowledge gaps which must be addressed to confirm the mechanistic insights provided by these findings and address directions for developing new resources to translate knowledge of PyC distributions and abundances in foods into improved dietary guidance to reduce risks of toxic metal exposure. The information obtained from this dissertation research provides new understanding of the role of PyCs and metal-binding plant compounds in human nutrition and establishes a foundation for future research investigating novel mechanisms through which plant-based diets confer health benefits.

Chapter 1 References

1. Collaborators GBDD. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958-72. Epub 2019/04/08. doi: 10.1016/S0140-6736(19)30041-8. PubMed PMID: 30954305; PMCID: PMC6899507.

2. Willett WC, Stampfer MJ, Colditz GA, Rosner BA, Speizer FE. Relation of meat, fat, and fiber intake to the risk of colon cancer in a prospective study among women. N Engl J Med. 1990;323(24):1664-72. Epub 1990/12/13. doi: 10.1056/NEJM199012133232404. PubMed PMID: 2172820.

3. Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A, Willett WC. Intake of fat, meat, and fiber in relation to risk of colon cancer in men. Cancer Res. 1994;54(9):2390-7. Epub 1994/05/01. PubMed PMID: 8162586.

4. WCRF. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. 2018.

5. Wang X, Lin X, Ouyang YY, Liu J, Zhao G, Pan A, Hu FB. Red and processed meat consumption and mortality: dose-response meta-analysis of prospective cohort studies. Public Health Nutr. 2016;19(5):893-905. Epub 2015/07/07. doi: 10.1017/S1368980015002062. PubMed PMID: 26143683.

6. Micha R, Wallace SK, Mozaffarian D. Red and processed meat consumption and risk of incident coronary heart disease, stroke, and diabetes mellitus: a systematic review and meta-analysis. Circulation. 2010;121(21):2271-83. Epub 2010/05/19. doi: 10.1161/CIRCULATIONAHA.109.924977. PubMed PMID: 20479151; PMCID: PMC2885952.

7. Fretts AM, Howard BV, McKnight B, Duncan GE, Beresford SA, Mete M, Eilat-Adar S, Zhang Y, Siscovick DS. Associations of processed meat and unprocessed red meat intake with incident diabetes: the Strong Heart Family Study. Am J Clin Nutr. 2012;95(3):752-8. Epub 2012/01/27. doi: 10.3945/ajcn.111.029942. PubMed PMID: 22277554; PMCID: PMC3278249.

8. Dauchet L, Amouyel P, Hercberg S, Dallongeville J. Fruit and vegetable consumption and risk of coronary heart disease: a meta-analysis of cohort studies. J Nutr. 2006;136(10):2588-93. Epub 2006/09/22. doi: 10.1093/jn/136.10.2588. PubMed PMID: 16988131.

9. He FJ, Nowson CA, MacGregor GA. Fruit and vegetable consumption and stroke: meta-analysis of cohort studies. Lancet. 2006;367(9507):320-6. Epub 2006/01/31. doi: 10.1016/S0140-6736(06)68069-0. PubMed PMID: 16443039.

10. He FJ, Nowson CA, Lucas M, MacGregor GA. Increased consumption of fruit and vegetables is related to a reduced risk of coronary heart disease: meta-analysis of cohort studies. J Hum Hypertens. 2007;21(9):717-28. Epub 2007/04/20. doi: 10.1038/sj.jhh.1002212. PubMed PMID: 17443205.

11. Kim H, Caulfield LE, Garcia-Larsen V, Steffen LM, Grams ME, Coresh J, Rebholz CM. Plant-Based Diets and Incident CKD and Kidney Function. Clin J Am Soc Nephrol. 2019;14(5):682-91. Epub 2019/04/27. doi: 10.2215/CJN.12391018. PubMed PMID: 31023928; PMCID: PMC6500948.

12. Kim H, Caulfield LE, Garcia-Larsen V, Steffen LM, Coresh J, Rebholz CM. Plant-Based Diets Are Associated With a Lower Risk of Incident Cardiovascular Disease, Cardiovascular Disease Mortality, and All-Cause Mortality in a General Population of Middle-Aged Adults. J Am Heart Assoc. 2019;8(16):e012865. Epub 2019/08/08. doi: 10.1161/JAHA.119.012865. PubMed PMID: 31387433; PMCID: PMC6759882.

13. Satija A, Bhupathiraju SN, Rimm EB, Spiegelman D, Chiuve SE, Borgi L, Willett WC, Manson JE, Sun Q, Hu FB. Plant-Based Dietary Patterns and Incidence of Type 2 Diabetes in US Men and Women: Results from Three Prospective Cohort Studies. PLoS Med. 2016;13(6):e1002039. Epub 2016/06/15. doi: 10.1371/journal.pmed.1002039. PubMed PMID: 27299701; PMCID: PMC4907448.

14. Kahleova H, Levin S, Barnard N. Cardio-Metabolic Benefits of Plant-Based Diets. Nutrients. 2017;9(8). doi: 10.3390/nu9080848. PubMed PMID: 28792455; PMCID: PMC5579641.

15. Sofi F, Cesari F, Abbate R, Gensini GF, Casini A. Adherence to Mediterranean diet and health status: meta-analysis. BMJ. 2008;337:a1344. Epub 2008/09/13. doi: 10.1136/bmj.a1344. PubMed PMID: 18786971; PMCID: PMC2533524.

16. Rodriguez-Garcia C, Sanchez-Quesada C, Toledo E, Delgado-Rodriguez M, Gaforio JJ. Naturally Lignan-Rich Foods: A Dietary Tool for Health Promotion? Molecules. 2019;24(5). Epub 2019/03/09. doi: 10.3390/molecules24050917. PubMed PMID: 30845651; PMCID: PMC6429205.

17. Mason JK, Thompson LU. Flaxseed and its lignan and oil components: can they play a role in reducing the risk of and improving the treatment of breast cancer? Appl Physiol Nutr Metab. 2014;39(6):663-78. Epub 2014/05/30. doi: 10.1139/apnm-2013-0420. PubMed PMID: 24869971.

18. Fabian CJ, Kimler BF, Zalles CM, Klemp JR, Petroff BK, Khan QJ, Sharma P, Setchell KD, Zhao X, Phillips TA, Metheny T, Hughes JR, Yeh HW, Johnson KA. Reduction in Ki-67 in benign breast tissue of high-risk women with the lignan secoisolariciresinol diglycoside. Cancer Prev Res (Phila). 2010;3(10):1342-50. Epub 2010/08/21. doi: 10.1158/1940-6207.CAPR-10-0022. PubMed PMID: 20724470; PMCID: PMC2955777.

19. Thompson LU, Chen JM, Li T, Strasser-Weippl K, Goss PE. Dietary flaxseed alters tumor biological markers in postmenopausal breast cancer. Clin Cancer Res. 2005;11(10):3828-35. Epub 2005/05/18. doi: 10.1158/1078-0432.CCR-04-2326. PubMed PMID: 15897583.

20. Rebholz CM, Coresh J, Grams ME, Steffen LM, Anderson CA, Appel LJ, Crews DC. Dietary Acid Load and Incident Chronic Kidney Disease: Results from the ARIC Study. Am J Nephrol. 2015;42(6):427-35. Epub 2016/01/21. doi: 10.1159/000443746. PubMed PMID: 26789417; PMCID: PMC4733508.

21. Banerjee T, Crews DC, Wesson DE, Tilea AM, Saran R, Rios-Burrows N, Williams DE, Powe NR, Centers for Disease C, Prevention Chronic Kidney Disease Surveillance T. High Dietary Acid Load Predicts ESRD among Adults with CKD. J Am Soc Nephrol. 2015;26(7):1693-700. Epub 2015/02/14. doi: 10.1681/ASN.2014040332. PubMed PMID: 25677388; PMCID: PMC4483581.

22. Goraya N, Simoni J, Jo CH, Wesson DE. A comparison of treating metabolic acidosis in CKD stage 4 hypertensive kidney disease with fruits and vegetables or sodium bicarbonate. Clin J Am Soc Nephrol. 2013;8(3):371-81. Epub 2013/02/09. doi: 10.2215/CJN.02430312. PubMed PMID: 23393104; PMCID: PMC3586961.

23. Mirmiran P, Yuzbashian E, Asghari G, Sarverzadeh S, Azizi F. Dietary fibre intake in relation to the risk of incident chronic kidney disease. Br J Nutr. 2018;119(5):479-85. Epub 2018/01/22. doi: 10.1017/S0007114517003671. PubMed PMID: 29352819.

24. Diaz-Lopez A, Bullo M, Basora J, Martinez-Gonzalez MA, Guasch-Ferre M, Estruch R, Warnberg J, Serra-Majem L, Aros F, Lapetra J, Ros E, Pinto X, Covas MI, Salas-Salvado J. Cross-sectional associations between macronutrient intake and chronic kidney disease in a population at high cardiovascular risk. Clin Nutr. 2013;32(4):606-12. Epub 2012/11/13. doi: 10.1016/j.clnu.2012.10.013. PubMed PMID: 23141101.

25. Gopinath B, Harris DC, Flood VM, Burlutsky G, Brand-Miller J, Mitchell P. Carbohydrate nutrition is associated with the 5-year incidence of chronic kidney disease. J Nutr. 2011;141(3):433-9. Epub 2011/01/14. doi: 10.3945/jn.110.134304. PubMed PMID: 21228263.

26. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Biomarkers Prev. 2005;14(8):1847-50. Epub 2005/08/17. doi: 10.1158/1055-9965.EPI-05-0456. PubMed PMID: 16103423.

27. Miller GW, Jones DP. The nature of nurture: refining the definition of the exposome. Toxicol Sci. 2014;137(1):1-2. Epub 2013/11/12. doi: 10.1093/toxsci/kft251. PubMed PMID: 24213143; PMCID: PMC3871934.

28. Kim K, Melough MM, Vance TM, Noh H, Koo SI, Chun OK. Dietary Cadmium Intake and Sources in the US. Nutrients. 2018;11(1). Epub 2018/12/24. doi: 10.3390/nu11010002. PubMed PMID: 30577418; PMCID: PMC6356330.

29. European Food Safety Authority. Cadmium dietary exposure in the European population. EFSA Journal. 2012;10(1).

30. Satarug S, Vesey DA, Gobe GC. Health Risk Assessment of Dietary Cadmium Intake: Do Current Guidelines Indicate How Much is Safe? Environ Health Perspect. 2017;125(3):284-8. doi: 10.1289/EHP108. PubMed PMID: 28248635; PMCID: PMC5332171.

31. Pal R, Rai JP. Phytochelatins: peptides involved in heavy metal detoxification. Appl Biochem Biotechnol. 2010;160(3):945-63. doi: 10.1007/s12010-009-8565-4. PubMed PMID: 19224399.

32. Fujita Y, el Belbasi HI, Min KS, Onosaka S, Okada Y, Matsumoto Y, Mutoh N, Tanaka K. Fate of cadmium bound to phytochelatin in rats. Res Commun Chem Pathol Pharmacol. 1993;82(3):357-65. PubMed PMID: 8122036.

33. Jumarie C, Fortin C, Houde M, Campbell PG, Denizeau F. Cadmium uptake by Caco-2 cells: effects of Cd complexation by chloride, glutathione, and phytochelatins. Toxicol Appl Pharmacol. 2001;170(1):29-38. doi: 10.1006/taap.2000.9075. PubMed PMID: 11141353.

34. Cobbett C, Goldsbrough P. Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. Annu Rev Plant Biol. 2002;53:159-82. doi: 10.1146/annurev.arplant.53.100301.135154. PubMed PMID: 12221971.

35. Seregin IV, Kozhevnikova AD. Low-molecular-weight ligands in plants: role in metal homeostasis and hyperaccumulation. Photosynth Res. 2020. Epub 2020/07/13. doi: 10.1007/s11120-020-00768-1. PubMed PMID: 32653983.

36. Mou RX, Cao ZY, Lin XY, Wu L, Cao ZZ, Zhu ZW, Chen MX. Characterization of the phytochelatins and their derivatives in rice exposed to cadmium based on high-performance liquid chromatography coupled with data-dependent hybrid linear ion trap orbitrap mass spectrometry. Rapid Commun Mass Spectrom. 2016;30(16):1891-900. doi: 10.1002/rcm.7669. PubMed PMID: 27426698.

37. Mehra RK, Kodati VR, Abdullah R. Chain length-dependent Pb(II)-coordination in phytochelatins. Biochem Biophys Res Commun. 1995;215(2):730-6. doi: 10.1006/bbrc.1995.2524. PubMed PMID: 7488015.

38. Dennis KK, Go YM, Jones DP. Redox Systems Biology of Nutrition and Oxidative Stress. J Nutr. 2019;149(4):553-65. Epub 2019/04/06. doi: 10.1093/jn/nxy306. PubMed PMID: 30949678; PMCID: PMC6461723.

39. Langelueddecke C, Lee WK, Thevenod F. Differential transcytosis and toxicity of the hNGAL receptor ligands cadmium-metallothionein and cadmium-phytochelatin in colon-like Caco-2 cells: implications for in vivo cadmium toxicity. Toxicol Lett. 2014;226(2):228-35. doi: 10.1016/j.toxlet.2014.01.049. PubMed PMID: 24518829.

40. Horiguchi H, Oguma E, Sasaki S, Miyamoto K, Ikeda Y, Machida M, Kayama F. Comprehensive study of the effects of age, iron deficiency, diabetes mellitus, and cadmium burden on dietary cadmium absorption in cadmium-exposed female Japanese farmers. Toxicol Appl Pharmacol. 2004;196(1):114-23. doi: 10.1016/j.taap.2003.11.024. PubMed PMID: 15050413.

41. ATSDR, Faroon O, Ashizawa A, Wright S, Tucker P, Jenkins K, Ingerman L, Rudisill C. Toxicological Profile for Cadmium. In: ATSDR, editor. Atlanta, GA 2012.

42. Anjum NA, Hasanuzzaman M, Hossain MA, Thangavel P, Roychoudhury A, Gill SS, Rodrigo MA, Adam V, Fujita M, Kizek R, Duarte AC, Pereira E, Ahmad I. Jacks of metal/metalloid chelation trade in plants-an overview. Front Plant Sci. 2015;6:192. Epub 2015/04/18. doi: 10.3389/fpls.2015.00192. PubMed PMID: 25883598; PMCID: PMC4382971.

43. Schlemmer U, Frolich W, Prieto RM, Grases F. Phytate in foods and significance for humans: food sources, intake, processing, bioavailability, protective role and analysis. Mol Nutr Food Res. 2009;53 Suppl 2:S330-75. Epub 2009/09/24. doi: 10.1002/mnfr.200900099. PubMed PMID: 19774556.

44. Olszowy M. What is responsible for antioxidant properties of polyphenolic compounds from plants? Plant Physiol Biochem. 2019;144:135-43. Epub 2019/09/30. doi: 10.1016/j.plaphy.2019.09.039. PubMed PMID: 31563754.

45. Jadan-Piedra C, Chiocchetti GM, Clemente MJ, Velez D, Devesa V. Dietary compounds as modulators of metals and metalloids toxicity. Crit Rev Food Sci Nutr. 2018;58(12):2055-67. Epub 2017/07/08. doi: 10.1080/10408398.2017.1302407. PubMed PMID: 28686469.

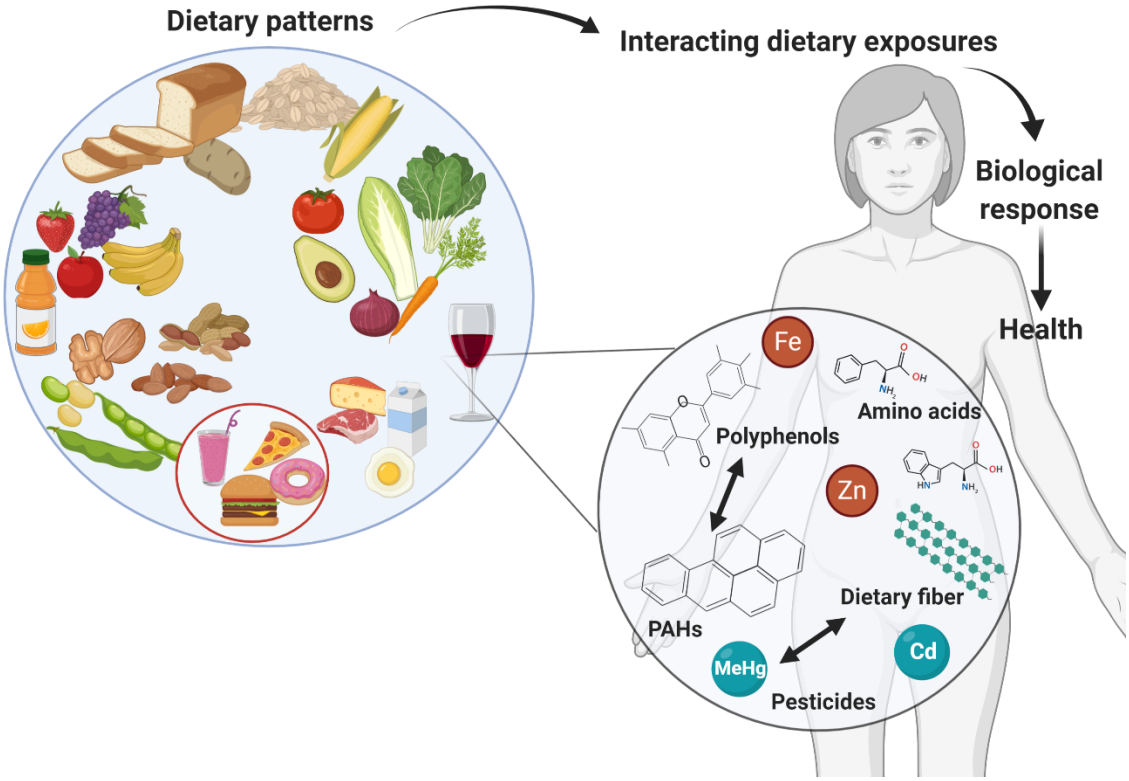

Figure 1.1. The dietary exposome. Dietary patterns capture the entire diet, including the synergistic and antagonistic effects of interacting dietary exposures. The interactions of nutritional exposures such as polyphenols, iron, and dietary fiber with non-nutritional exposures such as polycyclic aromatic hydrocarbons (PAHs), cadmium (Cd), and methylmercury (MeHg) ultimately determines the biological effects of the diet on human health. Created in BioRender.com.

CHAPTER 2

Phytochelatin database (PyCDB) –a resource for phytochelatin complexes of nutritional and environmental metals

Kristine K. Dennis[a,b], Karan Uppal[b], Ken H. Liu[b], Chunyu Ma[b], Bill Liang[b], Young-Mi Go[b], Dean P. Jones[a,b]

[a] Nutrition and Health Sciences, Laney Graduate School, Emory University, Atlanta, GA

[b] Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, Emory University, Atlanta, GA.

Published in *Database* (Oxford). 2019.

doi:10.1093/database/baz083

Format adapted for this dissertation

Abstract

Phytochelatins (PyC) are a diverse set of plant compounds that chelate metals, protect against metal toxicity and function in metal homeostasis. PyCs are present in plants consumed as food by humans and could, in principle, impact absorption and utilization of essential and toxic metals such as selenium and cadmium, respectively. PyCs vary in terminal amino acid composition and chain length, exist in multiple oxidation states, and reversibly bind multiple metals; consequently, PyCs include a large set of possible structures. Although individual PyC-metal complexes have been studied, no resource exists to characterize the diversity of PyCs and PyC-metal complexes. We used the scientific literature to develop a database of elemental formulas for polymer forms varying in chain length from 2 to 11 glutamyl-cysteine repeats. Using elemental formulas, we calculated monoisotopic masses using the most abundant isotopes of each element and calculated masses for complexes with 13 metals of nutritional and toxicological significance. The resulting phytochelatin database (PyCDB) contains 46,260 unique elemental formulas for PyC and PyC-metal complexes. The database is available online for download as well as for direct mass queries for mass spectrometry using an accurate mass annotation tool for user-selected PyC types, metals, and adducts of interest. We performed studies of a commonly consumed food, onion, to validate the database and test utility of the tool. Onion samples were analyzed using ultra-high resolution mass spectrometry-based metabolomics. Mass spectral features were annotated using the PyCDB web tool and the R package, xMSannotator; annotated features were further validated by collision-induced dissociation mass spectrometry. The results establish use and a workflow for PyCDB as a resource for characterization of PyCs and PyC-metal complexes.

Keywords: phytochelatin, metabolomics, metals, nutrition

Database URL: https://kuppal.shinyapps.io/pycdb/

Introduction

Phytochelatins (PyCs) function as key mediators of metal detoxification and homeostasis in plants. PyC-metal complexes protect plants from metal toxicity through chelating heavy metals and metalloids such as cadmium (Cd) and arsenic (As). They also bind required nutrients such as zinc (Zn), selenium (Se) and copper (Cu) (1). Due to their essential role, PyCs are extensively studied in agriculture and soil bioremediation (2-4). However, studies focus on analyzing only specific PyCs and a few metals of interest. More comprehensive characterization of PyC-metal complexes would allow greater understanding of metal sequestration and management in plants and additionally, a role for PyCs in metal bioavailability and toxicity in humans and other animal species consuming PyC-containing foods.

PyCs are glutathione (GSH)-derived polypeptides that are formed enzymatically by dipeptidyl transfer of a donor γ-glutamyl-cysteine (γ-Glu-Cys) to GSH or related peptide. The first PyC form identified was $(\gamma\text{-Glu-Cys})_n\text{-Gly}$ where n = 2-11 (5). Other forms exist in which glycine (Gly) can be substituted with β-alanine (β-Ala), Ala, glutamine (Gln), serine (Ser), or glutamate (Glu) or with no additional amino acid (4, 6-8). Because of their high thiol (-SH) content due to Cys residues, PyCs have strong metal-binding abilities and increasing metal capacity with increasing PyC size (3, 9).

PyC synthesis is increased in response to metal exposure, allowing plants to bind and mediate risk of heavy metal toxicity. PyCs have been observed in a wide range of plant species (1), accumulating in different tissues depending on the plant (10). For example, in rice (*Oryza sativa*)

seedlings exposed to a high dose of Cd, PyC_2-Gly was at the highest concentration in leaves followed by roots and shoots (6). In a different study of wild basil (*Clinopodium vulgare*) grown with excess Cd, the roots had 4-fold and 10-fold higher concentrations of PyC_2-Gly compared to leaves and shoots, respectively (11). PyC concentration and lengths also vary with exposure to different metal types and concentrations (11). Functionally within plants, PyC-metal complexes are sequestered in the vacuoles of plant cells but also occur at lower proportions in cytosol and phloem sap (1, 12). However, comprehensive characterization of types, concentration and tissue-location of PyCs and PyC-metal complexes has not been completed across common plants grown for human consumption.

With the development of high-resolution metabolomics (HRM) (**Figure 2.1**) (13-16), characterization of multiple forms of metal-free PyC and PyC-metal complexes within a single analysis is possible. However, no database exists to allow characterization of PyCs in plants and food. Here we create a phytochelatin database (PyCDB) for use with metabolomics analyses that contains elemental compositions and monoisotopic masses for a wide range of probable and experimentally detected PyC and PyC-metal complexes. We also provide the PyCDB through an accompanying web-based metabolite annotation tool. We include a metabolomics analysis of a common plant food, onion, as an example for identifying PyCs using the PyCDB in a HRM workflow. Additionally, we provide validation of PyC and PyC-metal complexes included in the PyCDB using collision-induced dissociation mass spectrometry. This database provides a resource for research of PyC and PyC-metal complexes in studies of metals with agricultural, nutritional, and toxicological significance.

Materials and Methods

Selection of Phytochelatins

Many metals bind with high affinity to thiols (-SH) in Cys residues, and thousands of potential variations of PyC-metal complexes are possible due to the variation in PyC lengths and terminal amino acids. In addition, sulfur can exist in different oxidation states (e.g. thiol versus disulfide form), and metals can interact with multiple functional groups in the peptides (e.g. O- and N-containing groups). Because generation of non-biologically relevant structures in a database can introduce an obstacle to understanding biology, we limited the range of predicted structures to those for which evidence indicates existence or likelihood of existence under relevant biologic conditions. Therefore, to define the scope of the database, we started with thiol (-SH) forms of PyCs. As most metals have the greatest binding affinity to the ionized thiolate ($-S^-$) form of thiol groups, preferential binding to thiol groups (over O- and N-containing groups) was assumed for this database iteration. At biological pH (near neutral or slightly basic), a proportion of the thiol groups is expected to be in the metal-binding thiolate form, and experimental evidence shows thiolate binding in PyC-metal complexes (17-19). Based upon this, the PyC-metal complex elemental compositions were calculated accordingly, with two protons removed for every divalent (2+) metal ion bound. For monoisotopic mass calculations based on PyC elemental compositions, we used the most abundant isotope for each element. Finally, predicted compounds were based on likely chemical interactions for common metal oxidation states at neutral pH, room temperature and room air. As knowledge of PyC-metal complexes increases, future versions will be updated to include other PyC-metal complexes.

Using the seven PyC forms (see "Base Phytochelatins" in **Table 2.1**) based upon C-terminal amino acid and repeating γ-Glu-Cys peptide units of 2 to 11, we calculated base elemental formulas from which predicted forms of PyC-metal complexes were generated (1) (**Figure 2.2**). As $(\gamma\text{-Glu-Cys})_n$-Ala and $(\gamma\text{-Glu-Cys})_n$-β-Ala are identical in elemental composition and monoisotopic mass, duplicate elemental compositions were not included. Disulfide (S-S) bonds are formed in molecules with two or more Cys under oxidizing conditions. Disulfide bonds form during food storage and preparation (20). Additionally, disulfide bonds will form during sample preparation unless specific anaerobic conditions or reducing agents are used. Disulfide PyC forms likely also exist in vivo due to normal reduction-oxidation reactions occurring as part of cellular signaling and metabolic processes. Single disulfide forms for PyC_2 to PyC_{11} were calculated by subtracting 2H from the elemental formulas, accounting for the two protons lost from thiol groups during disulfide bond formation. For each additional disulfide, an additional 2H were subtracted. For metal-bound forms, an additional two Cys (-SH groups) will be required for binding of a metal in a 2+ oxidation state. For example, a two disulfide form can only occur in PyC_4 or longer, and 2+ metal (Me^{2+})-binding can only occur with a two disulfide form in PyC_6 or longer. Up to five disulfide forms were included corresponding to known PyC lengths (**Table 2.1**). PyC-metal complexes were calculated using the elemental formulas described above as base units for construction of other forms. If two thiols were not available for binding, the PyC-metal complex was not included.

Selection of metals

Database metals include selected metals of either nutritional or toxicological importance. Other factors considered for metal selection included the common oxidation states, most abundant isotopes, and expected Lewis acid-base chemistry. The Lewis acid-base chemistry influences the ability of metals to form stable complexes. This relates to whether the metal ion is a hard acid or soft acid. Soft acids have greater tendency to form stable complexes with thiolates ($-S^-$) (21, 22). Of note but beyond the scope of the current discussion, the molecular environment of the thiol group will also impact the affinity to which it binds the metal. The molecular environment will be influenced by factors such as the biological matrix and sample preparation. For example, the binding affinity of metals to the thiolate will change depending on pH. These factors must be considered when designing experiments and interpreting results for studies of PyC-metal complexes.

Many of the soft metals of nutritional and toxicological interest have a common oxidation state of 2+: Cd, cobalt (Co), Cu, iron (Fe), lead (Pb), manganese (Mn), mercury (Hg), nickel (Ni), and Zn. Some have common oxidation states of 1+: Cu, Hg, and silver (Ag). Initial calculations were made for the 2+ ions of the most abundant isotope, including ^{114}Cd, ^{59}Co, ^{63}Cu, ^{56}Fe, ^{208}Pb, ^{55}Mn, ^{202}Hg, ^{58}Ni, ^{64}Zn, for the 1+ ion (^{107}Ag), and for negatively charged (2-) forms of selenium (^{80}Se). This includes 5 nutritionally important minerals (Fe, Cu, Mn, Se, and Zn) and 8 metals of environmental health concern (Cd, Co, Pb, Mn, Hg, Ni, Ag, and Se) (Supplementary Information, **Figure S2.2**). Some metals, such as Se and Mn, are both nutritional and toxicological metals due to tolerable upper intake limits (i.e. Mn, adults, 11 mg/day; Se, adults, 400 µg/day) being relatively low compared to some exposure levels in humans. Although

calcium (^{40}Ca, 2+) and magnesium (^{24}Mg, 2+) are hard acids and may not bind effectively to thiols, they are abundant in plants and could be present as PyC complexes (22). These Ca and Mg forms were included in the current database and additional forms for potassium (^{39}K, 1+) and sodium (^{23}Na, 1+) could be included in future database iterations.

Phytochelatin-metal complexes

PyC-metal complexes including one or two metals for 2+ ions and Se (2-) were calculated by subtracting 2H and 4H, respectively, from base elemental formulas to account for the loss of thiol protons during metal binding. For Se, the calculations were based upon experimental results demonstrating that selenite reacts with two PyC_2 to create a bound PyC_2-Se and an oxidized (S-S)PyC_2. The product was shown to be the selenotrisulfide, -S-Se-S- (23), where the Se of selenite was reduced to the formal 2+ oxidation state in binding to two thiolates. Ag (1+) one-metal and two-metal forms were calculated by subtracting 1H per metal addition. For two metal (2+) bound forms, at least 4 thiols would be needed. Therefore, calculations for two metal (2+) forms were completed for PyC_4 and larger, one disulfide forms for PyC_6 and larger, two disulfide forms for PyC_8 and larger, and three disulfide forms for PyC_{10} and PyC_{11}. Two metal complexes were calculated for all possible combinations of included metals, recognizing that the precise structure of the complexes cannot be predicted based upon these calculations. A larger number of metals and metal combinations could bind to longer chain lengths (up to 5 in PyC_{10} or PyC_{11} forms). However, these were not included in this PyCDB version because of the large number of combinations and limited evidence for such complex forms.

For all two metal PyC-metal complexes, bridging sulfurs, as occur in iron-sulfur clusters, are possible for most of the metals included in this database iteration (24). We referred to these as sulfido and disulfido forms. The elemental formulas were calculated by adding either one or two sulfurs to elemental formulas for two metal complexes. As selenide may form similar complexes, these were calculated by addition of one Se to the elemental compositions, referred to as the selenido form. Diselenido forms will be included in future database versions if disulfido forms are found.

Calculation of monoisotopic masses

Monoisotopic masses were calculated using the elemental formulas of the predicted compounds. Monoisotopic mass values were generated for each PyC and PyC-metal complex using a modification of the R package, OrgMassSpecR. This R package allows for automated calculations of monoisotopic masses using elemental formulas. Modifications were made to the functions *ListFormula* and *MonoisotopicMass* to include all elements of interest. Monoisotopic mass calculations were completed using the National Institute of Standards and Technology exact masses (rounded to eight decimal places) for the most abundant isotopes (25).

PyCDB web interface

The phytochelatin database (PyCDB) web tool was developed using the shiny package, shinyBS package and DT package in R. The web interface is maintained in the shiny server. Users can enter their experimental masses directly in a text box or upload them with ".csv" format or ".txt" format to perform accurate mass matching within the PyCDB web tool. Using the

get_mz_by_monoisotopicmass function in xMSannotator, the mass-to-charge ratio (*m/z*) for each PyC and PyC-metal complex for adducts of interest were calculated by adding the mass of the respective adduct (e.g. M+H, M+2H) and dividing by the charge state (z = 1, 2, 3) of the adduct (26). For accurate mass matching with the web tool, there is the option to include annotation only for elemental formulas which meet the nitrogen (N), oxygen (O), phosphorus (P), and sulfur (S) to carbon (NOPS) ratio check (27). The NOPS check allows filtering out elemental formulas which do not include the most common ratios of N, O, P, and S atoms to carbon. The user-defined input parameters can be specified and used by the back-end R function to find PyC/PyC-metal matches. Once the processing is complete, a table of PyC/PyC-metal matches is available in the web tool or for download as a .csv file.

Using the PyCDB locally

The user has the option of using the full database locally with existing tools or R packages, such as xMSannotator (26). With the xMSannotator *multilevelannotation* function, annotation criteria are available in addition to those available via the PyCDB web interface. The output of *multilevelannotation* includes a confidence level (i.e. none, low, medium, high) for annotated metabolites. Additional criteria include retention time clustering, hydrogen/carbon ratio checks, adduct requirements specified by the user for high confidence scores (e.g. M+H for high confidence), and abundance ratio checks for isotopes and multiply charged adducts.

Results and application

PyCDB content

The current database includes 46,260 unique elemental formulas for 240 PyC and 46,020 PyC-metal complexes. Information available for each complex includes the molecular formula, monoisotopic mass, isotope, PyC type (see **Table 2.1**), number of repeat units (e.g. PyC_2 to PyC_{11}), and information on bound metals (including type and number). With increasing PyC length, the number of compounds in the database increases due to the higher number of metal-binding sites in longer peptides (Supplementary Information, **Figure S2.2**).

PyCDB website implementation

The database is available in a user-friendly form from https://kuppal.shinyapps.io/pycdb/. As seen in **Figure 2.3**, users can select specific search criteria after uploading their experimental masses. Search options include selecting a subset of the 13 metals, adducts of interest, and specific PyC lengths or types (i.e. terminal amino acid). The "NOPS check" option can help users further filter less likely masses based on the ratio of nitrogen, oxygen, phosphorus and sulfur (NOPS) atoms to carbon in the chemical formula (27). The user can also define the mass error in parts per million (ppm; e.g. 5 ppm) to perform the accurate mass search. As described above, the user-defined input parameters are used by the back-end R function to find PyC matches. The full database is also available for download at https://s3.amazonaws.com/phytochelatindatabase/full_version_PyCDB_20180821.csv, which can be used with other annotation tools like xMSannotator for users interested in additional annotation criteria (26).

Example: phytochelatin detection in onions

Samples from eight onions were analyzed using HRM. Detected metabolites were defined by accurate mass (*m/z*), retention time, and intensity profiles. Details of the sample preparation and analysis are described in the Supplementary Information. 19,270 features were detected on the C18 column. The feature table was analyzed two ways, targeting a subset of the database (Supplementary Information). First, the feature table was annotated using xMSannotator with the custom database option. The xMSannotator code used for the analysis is provided in the Supplementary Information. 628 features were annotated using the PyCDB subset. The feature table was also analyzed in the PyCDB web tool using the same parameters as xMSannotator but without the NOPS check. The PyCDB web tool uses a simple annotation function which does not provide confidence scores. This resulted in 845 annotated features.

Using selection criteria based on high confidence matches from xMSannotator and high intensity features, features were selected for further validation with collision-induced dissociation mass spectrometry (MS/MS). The feature *m/z* 538.1270 was matched to (S-S)PyC_2-Gly ($C_{18}H_{27}N_5O_{10}S_2$) [M+H]. MS/MS of this *m/z* target was completed to confirm the identity (see **Figure S2.1**). Many annotated PyC-metal complexes from the onion data were at too low intensity for collision-induced dissociation analysis. PyC-metal complexes formed using chemical standards and selected metals (see Supplementary Information) showed that experimental masses were consistent with predicted masses. Additional details of metabolite validation and PyC-metal complex collision-induced dissociation are provided in the Supplementary Information (**Figure S2.1** and **Figure S2.3** to **S2.8**). Examples are provided along with literature data for PyCs and PyC-metal complexes in **Table 2.2**.

Future Directions

The database is open source and can be refined as plant and food metabolomics data are compared to the predicted complexes. Such modifications can include additional metals of nutritional and toxicological significance such as molybdenum and arsenic. Although selenocysteine (Sec) is not common in plants, Sec may form in the PyC in high Se conditions and be of interest to explore. Other metals and metal oxidation forms are possible. The complex coordination chemistries of the metals will need to be considered for inclusion. To aid in confirmation of identities, additional methods, such as metal removal or study by ion mobility spectrometry-mass spectrometry, may be needed to address low abundance of the PyC-metal complexes in biologic materials. Although up to two metals were only considered in the current database iteration, PyC binding of up to five metals for longer forms is chemically possible and could be explored. Finally, a future iteration of the PyCDB could include an in silico fragmentation tool to account for predicted MS/MS spectra of PyC-metal complexes. Although peptide fragmentation tools are available, predicted fragmentation for PyC-metal complexes will need to consider metal interactions with O- and N-containing carboxy and amino groups as well as sulfido, disulfido, and selenido PyC forms included in the database.

In addition to the potential applications of the PyCDB for understanding metal homeostasis in plants and absorption of nutritional and toxicological metals in plant-derived foods, the PyCDB could also be useful for understanding metal-dependent processes such as nutritional immunity. Nutritional immunity is the process by which a host controls access to micronutrients to protect from bacterial infections (28, 29), and PyCs could contribute to mechanisms of nutritional

immunity due to their metal-binding characteristics. Thus, the PyCDB has the potential to be a useful resource for diverse in vivo and in vitro investigations of PyCs and PyC-metal complexes.

Conclusions

The range of potential PyCs and PyC-metal complexes in plants and food products is extensive and diverse. The PyCDB provides a foundational resource for research efforts to characterize PyC profiles. Here we demonstrate PyCs can be detected and validated in the edible portion of a commonly consumed plant food, onion. Additionally, PyCs and PyC-metal complexes formed in vitro are detected at the predicted masses in the database. Future database versions can include additional PyC forms and validated compounds. Given the significant role of PyCs in binding metals of toxicological and nutritional significance, this database provides a resource to improve understanding of PyCs in metal homeostasis and metal bioavailability in plants and plant-derived foods consumed by animals.

Funding

This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases [T32 DK007734]; National Institute of Environmental Health Sciences [R01 ES023485, U2C ES030163]; and National Institutes of Health, Office of the Director [S10 OD018006].

Conflict of interest. None declared.

Chapter 2 References

1. Cobbett C, Goldsbrough P. Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. Annu Rev Plant Biol. 2002;53:159-82. doi: 10.1146/annurev.arplant.53.100301.135154. PubMed PMID: 12221971.

2. Cobbett CS. Phytochelatins and their roles in heavy metal detoxification. Plant Physiol. 2000;123(3):825-32. PubMed PMID: 10889232; PMCID: PMC1539264.

3. Hirata K, Tsuji N, Miyamoto K. Biosynthetic regulation of phytochelatins, heavy metal-binding peptides. J Biosci Bioeng. 2005;100(6):593-9. doi: 10.1263/jbb.100.593. PubMed PMID: 16473766.

4. Rauser WE. Phytochelatins and related peptides. Structure, biosynthesis, and function. Plant Physiol. 1995;109(4):1141-9. PubMed PMID: 8539285; PMCID: PMC157644.

5. Rauser WE. Phytochelatins. Annu Rev Biochem. 1990;59:61-86. doi: 10.1146/annurev.bi.59.070190.000425. PubMed PMID: 2197985.

6. Mou RX, Cao ZY, Lin XY, Wu L, Cao ZZ, Zhu ZW, Chen MX. Characterization of the phytochelatins and their derivatives in rice exposed to cadmium based on high-performance liquid chromatography coupled with data-dependent hybrid linear ion trap orbitrap mass spectrometry. Rapid Commun Mass Spectrom. 2016;30(16):1891-900. doi: 10.1002/rcm.7669. PubMed PMID: 27426698.

7. Wood BA, Feldmann J. Quantification of phytochelatins and their metal(loid) complexes: critical assessment of current analytical methodology. Anal Bioanal Chem. 2012;402(10):3299-309. PubMed PMID: 22514800.

8. Kubota H, Sato K, Yamada T, Maitani T. Phytochelatin homologs induced in hairy roots of horseradish. Phytochemistry. 2000;53(2):239-45. PubMed PMID: 10680177.

9. Mehra RK, Kodati VR, Abdullah R. Chain length-dependent Pb(II)-coordination in phytochelatins. Biochem Biophys Res Commun. 1995;215(2):730-6. doi: 10.1006/bbrc.1995.2524. PubMed PMID: 7488015.

10. Marentes E, Rauser WE. Different proportions of cadmium occur as Cd-binding phytochelatin complexes in plants. Physiol Plant. 2007;131(2):291-301. doi: 10.1111/j.1399-3054.2007.00960.x. PubMed PMID: 18251900.

11. Bardarov K, Naydenov M, Djingova R. HPLC-HRMS method for fast phytochelatins determination in plants. Application to analysis of Clinopodium vulgare L. Talanta. 2015;142:20-7. Epub 2015/05/25. doi: 10.1016/j.talanta.2015.04.014. PubMed PMID: 26003687.

12. Mendoza-Cozatl DG, Butko E, Springer F, Torpey JW, Komives EA, Kehr J, Schroeder JI. Identification of high levels of phytochelatins, glutathione and cadmium in the phloem sap of Brassica napus. A role for thiol-peptides in the long-distance transport of cadmium and the effect of cadmium on iron translocation. Plant J. 2008;54(2):249-59. doi: 10.1111/j.1365-313X.2008.03410.x. PubMed PMID: WOS:000254792700006.

13. Soltow QA, Strobel FH, Mansfield KG, Wachtman L, Park Y, Jones DP. High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. Metabolomics. 2013;9(1 Suppl):S132-S43. doi: 10.1007/s11306-011-0332-1. PubMed PMID: 26229523; PMCID: PMC4517297.

14. Go YM, Walker DI, Liang Y, Uppal K, Soltow QA, Tran V, Strobel F, Quyyumi AA, Ziegler TR, Pennell KD, Miller GW, Jones DP. Reference Standardization for Mass Spectrometry and High-resolution Metabolomics Applications to Exposome Research. Toxicol Sci. 2015;148(2):531-43. doi: 10.1093/toxsci/kfv198. PubMed PMID: 26358001; PMCID: PMC4675836.

15. Go YM, Uppal K, Walker DI, Tran V, Dury L, Strobel FH, Baubichon-Cortay H, Pennell KD, Roede JR, Jones DP. Mitochondrial metabolomics using high-resolution Fourier-transform mass spectrometry. Methods Mol Biol. 2014;1198:43-73. doi: 10.1007/978-1-4939-1258-2_4. PubMed PMID: 25270922; PMCID: PMC4318503.

16. Walker DI, Mallon CT, Hopke PK, Uppal K, Go YM, Rohrbeck P, Pennell KD, Jones DP. Deployment-Associated Exposure Surveillance With High-Resolution Metabolomics. J Occup Environ Med. 2016;58(8 Suppl 1):S12-21. doi: 10.1097/JOM.0000000000000768. PubMed PMID: 27501099; PMCID: PMC4978191.

17. Chen LQ, Yang LM, Wang QQ. In vivo phytochelatins and Hg-phytochelatin complexes in Hg-stressed Brassica chinensis L. Metallomics. 2009;1(1):101-6. doi: 10.1039/b815477e. PubMed PMID: WOS:000268924300011.

18. Yen TY, Villa JA, DeWitt JG. Analysis of phytochelatin-cadmium complexes from plant tissue culture using nano-electrospray ionization tandem mass spectrometry and capillary liquid chromatography/electrospray ionization tandem mass spectrometry. J Mass Spectrom. 1999;34(9):930-41. Epub 1999/09/24. doi: 10.1002/(SICI)1096-9888(199909)34:9<930::AID-JMS853>3.0.CO;2-E. PubMed PMID: 10491589.

19. Scheidegger C, Suter MJ, Behra R, Sigg L. Characterization of lead-phytochelatin complexes by nano-electrospray ionization mass spectrometry. Front Microbiol. 2012;3:41. Epub 2012/02/22. doi: 10.3389/fmicb.2012.00041. PubMed PMID: 22347879; PMCID: PMC3277950.

20. Jones DP, Coates RJ, Flagg EW, Eley JW, Block G, Greenberg RS, Gunter EW, Jackson B. Glutathione in foods listed in the National Cancer Institute's Health Habits and History Food Frequency Questionnaire. Nutr Cancer. 1992;17(1):57-75. doi: 10.1080/01635589209514173. PubMed PMID: 1574445.

21. Haas KL, Franz KJ. Application of metal coordination chemistry to explore and manipulate cell biology. Chem Rev. 2009;109(10):4921-60. doi: 10.1021/cr900134a. PubMed PMID: 19715312; PMCID: PMC2761982.

22. Pearson RG. Hard and Soft Acids and Bases. Journal of the American Chemical Society. 1963;85(22):3533-9. doi: 10.1021/ja00905a001.

23. Spain SM, Rabenstein DL. Characterization of the selenotrisulfide formed by reaction of selenite with end-capped phytochelatin-2. Anal Bioanal Chem. 2004;378(6):1561-7. PubMed PMID: 15214417.

24. Vahrenkamp H. Sulfur Atoms as Ligands in Metal-Complexes. Angew Chem Int Edit. 1975;14(5):322-9. doi: DOI 10.1002/anie.197503221. PubMed PMID: WOS:A1975AD80600008.

25. Coursey JS, Schwab, D.J., Tsai, J.J., Dragoset, R.A. Atomic Weights and Isotopic Compositions [Online]. Gaithersburg, MD: National Institute of Standards and Technology; 2015 [cited 2017 September 15,]; version 4.1:[Available from: http://physics.nist.gov/Comp

26. Uppal K, Walker DI, Jones DP. xMSannotator: An R Package for Network-Based Annotation of High-Resolution Metabolomics Data. Anal Chem. 2017;89(2):1063-7. doi: 10.1021/acs.analchem.6b01214. PubMed PMID: 27977166; PMCID: PMC5447360.

27. Kind T, Fiehn O. Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. BMC Bioinformatics. 2007;8:105. doi: 10.1186/1471-2105-8-105. PubMed PMID: 17389044; PMCID: PMC1851972.

28. Diaz-Ochoa VE, Jellbauer S, Klaus S, Raffatellu M. Transition metal ions at the crossroads of mucosal immunity and microbial pathogenesis. Front Cell Infect Microbiol. 2014;4:2. Epub 2014/01/31. doi: 10.3389/fcimb.2014.00002. PubMed PMID: 24478990; PMCID: PMC3900919.

29. Hood MI, Skaar EP. Nutritional immunity: transition metals at the pathogen-host interface. Nat Rev Microbiol. 2012;10(8):525-37. Epub 2012/07/17. doi: 10.1038/nrmicro2836. PubMed PMID: 22796883; PMCID: PMC3875331.

30. Vacchina V, Chassaigne H, Oven M, Zenk MH, Lobinski R. Characterisation and determination of phytochelatins in plant extracts by electrospray tandem mass spectrometry. Analyst. 1999;124(10):1425-30. doi: DOI 10.1039/a905163e. PubMed PMID: WOS:000083334700003.

31. Spisso AA, Cerutti S, Silva F, Pacheco PH, Martinez LD. Characterization of Hg-phytochelatins complexes in vines (Vitis vinifera cv Malbec) as defense mechanism against metal stress. Biometals. 2014;27(3):591-9. doi: 10.1007/s10534-014-9732-9. PubMed PMID: WOS:000335729900013.

32. Dago A, Arino C, Diaz-Cruz JM, Esteban M. Analysis of phytochelatins and Hg-phytochelatin complexes in Hordeum vulgare plants stressed with Hg and Cd: HPLC study with amperometric detection. Int J Environ an Ch. 2014;94(7):668-78. doi: 10.1080/03067319.2013.864649. PubMed PMID: WOS:000335843200004.

Table 2.1. PyC structures vary by terminal amino acid, number of repeating peptide units (n = 2-11) and number of disulfide bonds (m = 1-5).

Base PyCs	Elemental formula (e.g. n = 2)	Disulfide form
$(\gamma\text{-Glu-Cys})_n\text{-Gly}$	$C_{18}H_{29}N_5O_{10}S_2$	$(S\text{-}S)_m(\gamma\text{-Glu-Cys})_n\text{-Gly}$
$(\gamma\text{-Glu-Cys})_n\text{-}\beta\text{-Ala}$	$C_{19}H_{31}N_5O_{10}S_2$	$(S\text{-}S)_m(\gamma\text{-Glu-Cys})_n\text{-}\beta\text{-Ala}$
$(\gamma\text{-Glu-Cys})_n\text{-Ala}$	$C_{19}H_{31}N_5O_{10}S_2$	$(S\text{-}S)_m(\gamma\text{-Glu-Cys})_n\text{-Ala}$
$(\gamma\text{-Glu-Cys})_n$	$C_{16}H_{26}N_4O_9S_2$	$(S\text{-}S)_m(\gamma\text{-Glu-Cys})_n$
$(\gamma\text{-Glu-Cys})_n\text{-Gln}$	$C_{21}H_{34}N_6O_{11}S_2$	$(S\text{-}S)_m(\gamma\text{-Glu-Cys})_n\text{-Gln}$
$(\gamma\text{-Glu-Cys})_n\text{-Ser}$	$C_{19}H_{31}N_5O_{11}S_2$	$(S\text{-}S)_m(\gamma\text{-Glu-Cys})_n\text{-Ser}$
$(\gamma\text{-Glu-Cys})_n\text{-Glu}$	$C_{21}H_{33}N_5O_{12}S_2$	$(S\text{-}S)_m(\gamma\text{-Glu-Cys})_n\text{-Glu}$

Table 2.2. Comparisons of predicted masses in database with experimental masses of PyCs identified by mass spectrometry

PyC	Predicted mass (*m/z*)	Experimental mass (*m/z*)	Mass deviation (ppm)	Adduct
		Unbound PyCs		
(S-S)PyC_2-Gly[a,b]	538.1272	538.1254	3.3	M+H
PyC_2-Gly[a,b]	540.1429	540.1437, 540.1, 540.1430, 540.0	1.5, - , 0.1, -	M+H
PyC_3-Gly[a,b]	772.1946	772.1955, 772.2, 772.1948, 772.0	1.2, - , 0.3, -	M+H
PyC_4-Gly[a,b]	1004.2464	1004.2436, 1004.2, 1004.2458, 1004.0	2.8, - , 0.6, -	M+H
PyC_5-Gly[b]	1236.2982	1236.2	-	M+H
PyC_6-Gly[b]	1468.35	1468.2	-	M+H
PyC_2-Ser[b]	570.1534	570.1538	0.7	M+H
PyC_3-Ser[b]	802.2052	802.206	1	M+H
PyC_4-Ser[b]	1034.257	1034.2577	0.7	M+H
PyC_2-Gln[b]	611.18	611.1802	0.3	M+H
PyC_3-Gln[b]	843.2318	843.2328	1.2	M+H
PyC_2-Glu[b]	612.1640	612.1648	1.3	M+H
PyC_3-Glu[b]	844.2158	844.2167	1.1	M+H
PyC_4-Glu[b]	1076.2675	1076.2686	1	M+H
		PyC-metal complexes		
PyC_2-Gly-Hg[b]	740.0979	740, 740.1	-	M+H
PyC_3-Gly-Hg[b]	972.1496	972, 972.1	-	M+H
PyC_4-Gly-Hg[b]	1202.1858	1202, 1202.2	-	M+H
PyC_4-Gly-Hg(2)[b]	1404.1564	1404, 1404.2	-	M+H
PyC_2-Gly-Cd[a]	652.0306	652.0346	6.1	M+H
PyC_3-Gly-Cd[a,b]	884.0824	884.0710, 884.1	12.9, -	M+H
PyC_4-Gly-Cd[a,b]	1116.1341	1116.1392, 1116.1	4.6, -	M+H

PyC_4-Gly-Cd(2)[a]	1228.0218	1228.0279	5	M+H
PyC_5-Gly-Cd[b]	1348.1859	1348.2	-	M+H
PyC_2-Pb[b]	746.1039	746.1034	0.7	M+H
PyC_3-Pb[b]	978.1556	978.1559	0.3	M+H
PyC_4-Pb[b]	1210.2074	1210.1986	7.3	M+H
PyC_4-Pb(2)[b]	1416.1684	1416.1556	9	M+H
PyC_2-Zn[a*,b]	602.0564	602.0560, 602.0544	0.7, 3.3	M+H
PyC_3-Zn[a*]	834.1081	834.1136	6.6	M+H
PyC_4-Zn[a*]	1066.1599	1066.1608	0.8	M+H
PyC_4-Zn(2)[a*]	1128.0734	1128.0763	2.6	M+H
PyC_2-Mn[a]	593.0653	593.0665	2	M+H
PyC_3-Mn[a]	825.1170	825.1251	9.8	M+H
PyC_4-Mn[a]	1057.1688	1057.1672	1.5	M+H
PyC_4-Mn(2)[a]	1110.0912	1110.0913	0.1	M+H

[a]Compared with standards; see supplementary information.

[b]Previously identified PyC derivative with mass as reported (6, 18, 19, 30-32).

*Data not shown.

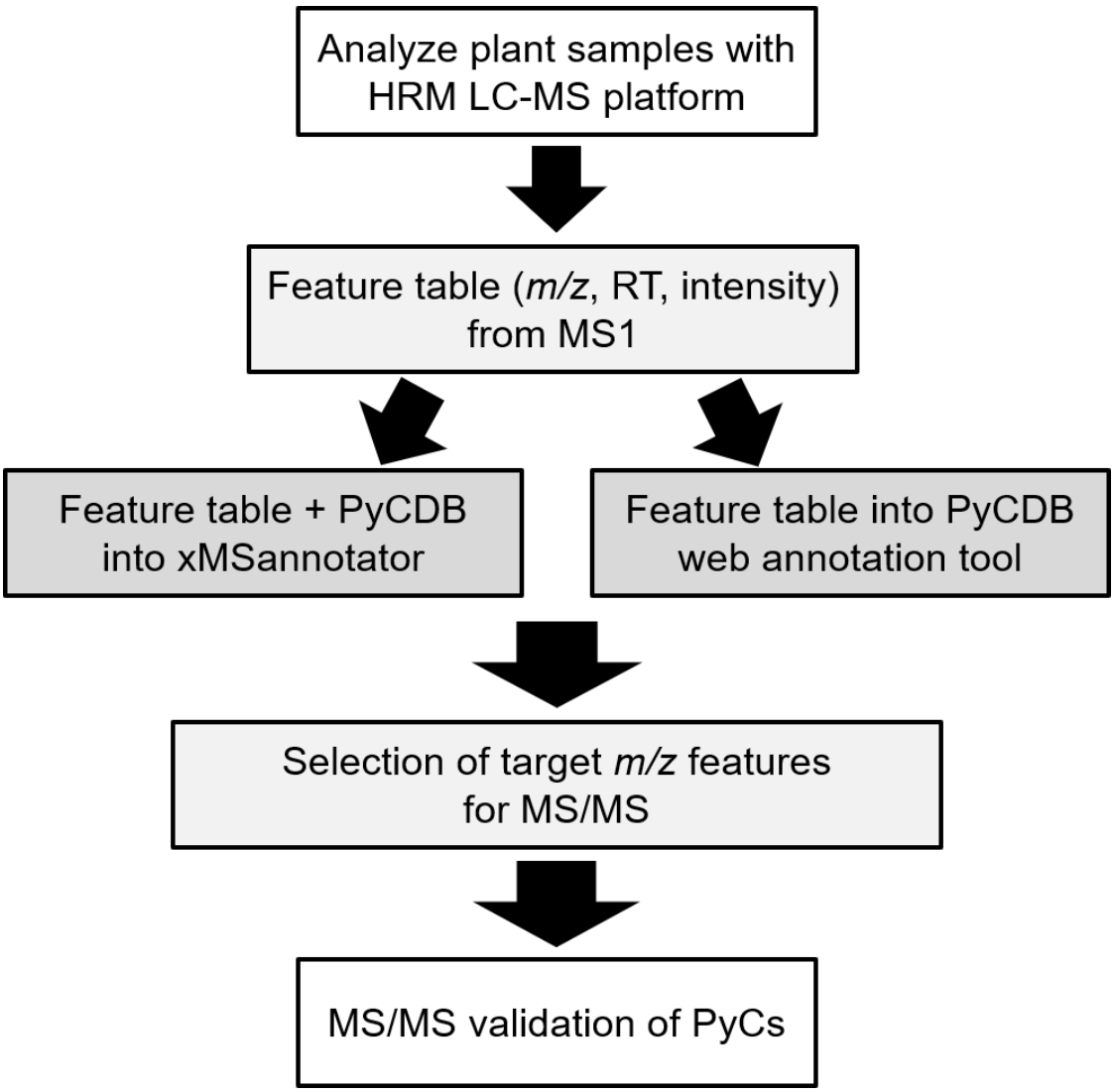

Figure 2.1. HRM workflow for PyC detection and validation. Using liquid chromatography with ultra-high resolution mass spectrometers followed by the application of data extraction algorithms, broad characterization of the metabolites (mass-to-charge, *m/z*; retention time, RT; and relative intensity) can be obtained. The mass spectral feature table is then searched against the database of compounds using matching criteria such as a retention time window and maximum allowable ppm differences. Annotated features of interest are then targeted for validation by collision induced dissociation using tandem mass spectrometry (MS/MS).

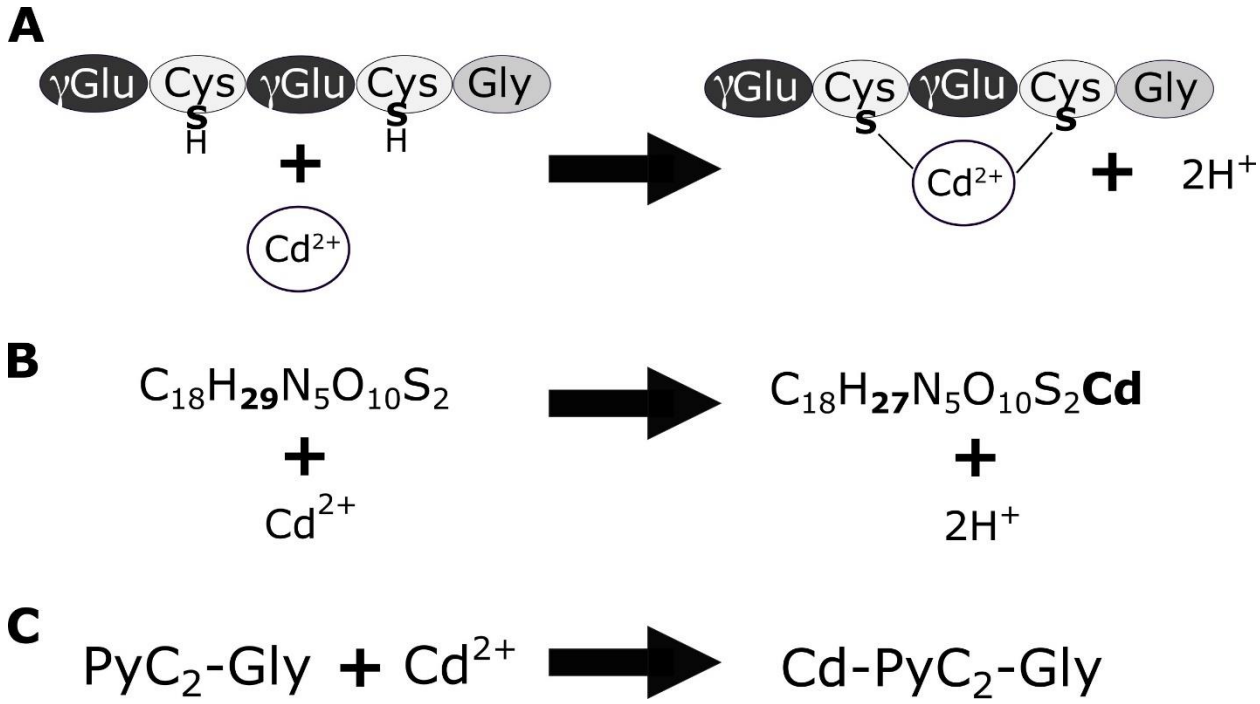

Figure 2.2. Formation of PyC-metal complexes. A, PyCs form complexes with metal ligands. Metal ligands in 2+ oxidation state will bind with the sulfurs of thiol groups on two cysteine residues. B, An example of the elemental formulas in PyCDB for the base PyC, phytochelatin2-glycine (PyC_2-Gly), in metal-bound and unbound forms. C, An example of the abbreviated name for the base PyC form, number of repeating peptide units, and metal (if bound).

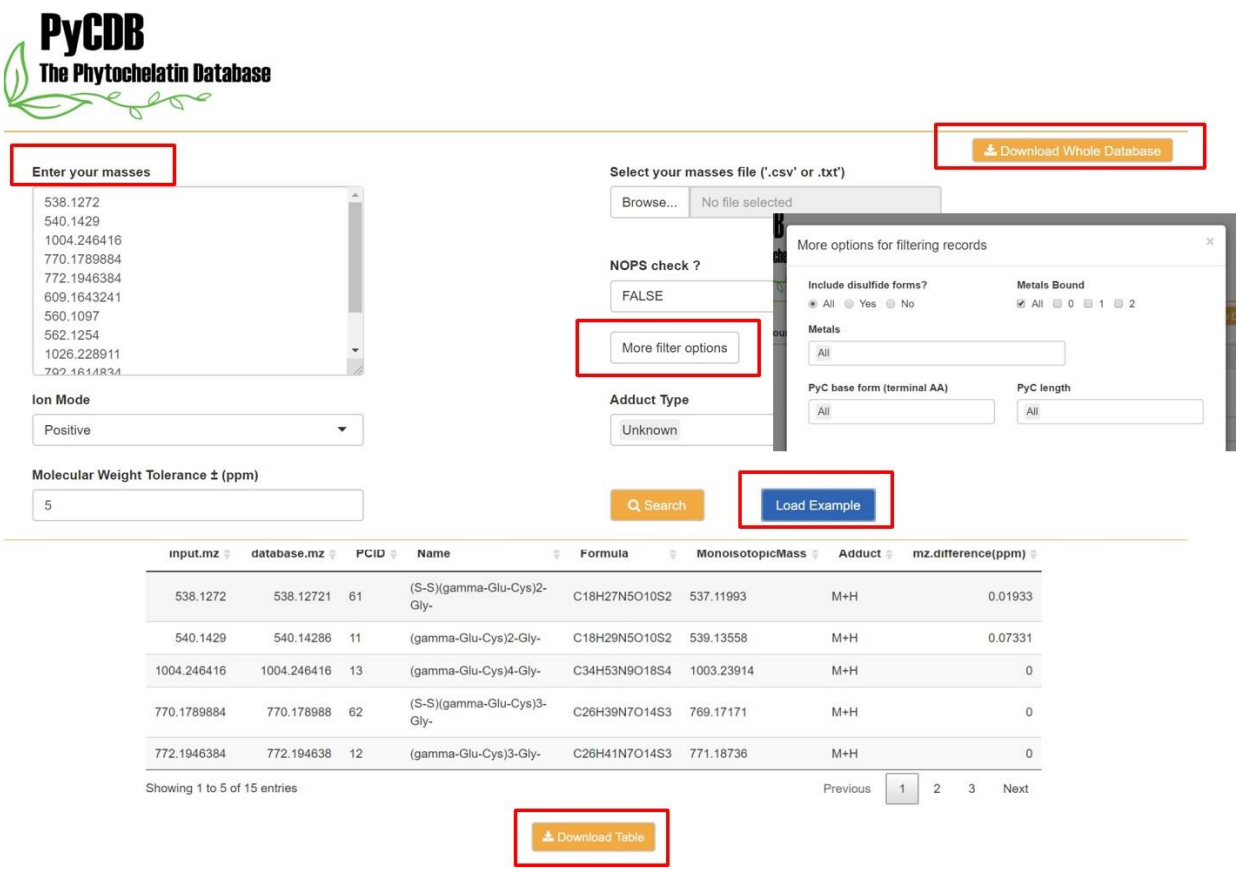

Figure 2.3. Screenshots of the web version of the PyC database. Query masses of interest can be entered manually or uploaded from a .csv or .txt file. Screenshot provided of an example search using the "Load Example" button with the default search settings and the output shown below. The output can be reviewed on the webpage or downloaded as a .csv file for future use. The webpage also provides an option to "Download Whole Database" for use with other annotation tools such as the R package, xMSannotator.

Supplementary Information

Phytochelatin database (PyCDB): A resource for phytochelatin complexes of nutritional and environmental metals

Kristine K. Dennis[a,b], Karan Uppal[b], Ken H. Liu[b], Chunyu Ma[b], Bill Liang[b], Young-Mi Go[b], Dean P. Jones[a,b]

[a] Nutrition and Health Sciences, Laney Graduate School, Emory University, Atlanta, GA

[b] Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, Emory University, Atlanta, GA.

Abstract

The supplementary information includes the methods of onion sample preparation and metabolomics analysis, annotation of onion features with xMSannotator, LC-MS/MS validation of PyC_2-Gly in onions, R script used for annotation with xMSannotator, and MS/MS results for PyC authentic standards (PyC_2-Gly, PyC_3-Gly, PyC_4-Gly) complexed with manganese and cadmium.

Metabolomics of onions

Eight onions were selected for analysis. 100 mg pieces of each onion were frozen at -80^{o}C. Metabolites were extracted from onion samples using a Polytron homogenizer in 400 µL extraction solution (2:1 acetonitrile to water) and homogenates were centrifuged for 10 min at 14,000 rpm at 4^{o}C. Metabolites in the supernatant were collected, loaded into auto-sampler vials,

and run on a high resolution LTQ-Velos Orbitrap mass spectrometer (Thermo Fisher). Samples were analyzed in triplicate with a 10 µl injection volume on a C18 column in positive electrospray ionization mode using a previously described method (1, 2). The data were extracted using the R packages apLCMS and xMSanalyzer to provide a table of detected metabolites (mass spectral features defined by accurate mass (*m/z*), retention time (RT), and intensity profiles) (3, 4).

Annotation of onion data with xMSannotator

Feature tables were analyzed using xMSannotator and the *multilevelannotation* function with the custom database option as defined in the R script below. Based on the expected concentration and frequency of detected PyC lengths from previous research, a subset of the PyCDB was run with only PyC_2 to PyC_6, focusing on forms with one or no metals (5, 6). xMSannotator allows accurate mass matching (Level 5 identification according to Schymanski et al. 2014) based on criteria such as correlation analysis, network modularity analysis, RT-based clustering, and mass defect analysis to assign confidence scores (0-3) to the annotations (7, 8). Multiple adducts were considered (i.e. M+H, M+Na, M+K, M+2H, M+3H, $M+NH_4$) with the M+H adduct required for the highest confidence level to be assigned. Using these parameters, 628 features were annotated using the PyCDB subset defined above.

Validation of phytochelatin with MS/MS

MS/MS of *m/z* 538.1270 was completed on the LTQ-Velos Orbitrap. The identification was confirmed via MS/MS in positive mode using collision-induced dissociation at 35V on the LTQ-

Velos Orbitrap. Validation of (S-S)PyC_2-Gly with MS/MS was performed in two ways. First, MS/MS onion spectra were uploaded into MetFrag for matching (9, 10). MetFrag links with PubChem, which has three phytochelatins in the database (PyC_2-Gly, PyC_3-Gly, PyC_4-Gly) (11). From MS/MS spectra for *m/z* 538.1270, MetFrag returned a top match for (S-S)PyC_2-Gly ($C_{18}H_{27}N_5O_{10}S_2$). An example of a representative MS/MS spectra can be seen in **Figure S2.1B**. As an authentic standard is available for PyC_2-Gly, MS/MS analysis was performed on the LTQ-Velos Orbitrap to generate a reference spectra (**Figure S2.1A**). Validation was also performed via matching of fragmentation patterns of the onion spectra (**Figure S2.1B**) with the spectra of the PyC_2-Gly authentic standard.

Mass fragmentation analysis of PyC and PyC-metal complex standards

Analytical standards of PyC_2-Gly (95% purity), PyC_3-Gly (95% purity), and PyC_4-Gly (95% purity) were obtained from CPC Scientific Inc. $CdCl_2$ and $MnCl_2$ were obtained from Sigma-Aldrich. Individual stock solutions of PyC_2, PyC_3, $CdCl_2$, and $MnCl_2$ were prepared in HPLC-grade water. PyC_2-Gly, PyC_3-Gly and PyC_4-Gly were prepared at 10 µM. PyC-metal complex solutions were prepared in water at 10 µM:10 µM PyC to metal ratios and were analyzed with direct injection mass spectrometry using a LTQ-Velos Orbitrap mass spectrometer (Thermo Fisher).

R script for annotation of onion metabolites with xMSannotator

1) Read in PyCDB as custom database file:

```
PC_full<-read.csv("full_version_PyCDB_20180821.csv")
```

2) Create subset of PyCDB for query (optional):

```
PC_full$PCkeep<-ifelse((PC_full$PC.General %in% c("PC2","PC3","PC4","PC5","PC6"))
& (PC_full$Metal.Form %in% c("0","1")), 1,0)
```

```
PC2_PC6<-subset(PC_full, PC_full$PCkeep ==1)
```

3) Read in data table:

```
dataA<-read.table("~/PyC/FoodFiles/onion/C18/C18_Onion_Stage3b_filter.txt",
header=TRUE)
```

4) Specify search parameters

```
max.mz.diff<-10  #mass search tolerance for DB matching in ppm
max.rt.diff<-10 #retention time tolerance between adducts/isotopes
corthresh<-0.7 #correlation threshold between adducts/isotopes
max_isp=5 #maximum number of isotopes to search for
mass_defect_window=0.01 #mass defect window for isotope search
```

5) Specify output location

```
outloc<-"~/PyC/FoodFiles/C18/onion_PyCDBpaper/"
```

6) Specify number of cores to be used

```
num_nodes<-4
```

7) Specify name of database to search against

```
db_name="Custom"
status=NA
customIDs<-NA
```

8) Specify number of sets the total database entries should be split into for searches

```
num_sets<-300
```

9) Specify ionization mode and adducts for search

```
mode<-"pos"
queryadductlist=c("M+2H","M+H","M+Na","M+K", "M+3H","M+NH4")
```

10) Provide name of custom database to be used for annotation

```
customDB<-PC2_PC6
```

11) xMSannotator multilevel annotation function

```
#############################
dataA<-unique(dataA)
print(dim(dataA))
print(format(Sys.time(), "%a %b %d %X %Y"))

system.time(annotres<-
multilevelannotation(dataA=dataA,max.mz.diff=max.mz.diff,max.rt.diff=max.rt.diff,cormeth
od="pearson",num_nodes=num_nodes,queryadductlist=queryadductlist,
```

```
                    mode=mode,outloc=outloc,db_name=db_name,
adduct_weights=NA,num_sets=num_sets,allsteps=TRUE,

corthresh=corthresh,NOPS_check=TRUE,customIDs=customIDs,missing.value=NA,deepsplit=2,networktype="unsigned",

minclustsize=10,module.merge.dissimilarity=0.2,filter.by=c("M+H"),biofluid.location=NA,origin=NA,status=status,boostIDs=NA,max_isp=max_isp, customDB=customDB,
HMDBselect=NA,mass_defect_window=mass_defect_window,pathwaycheckmode="pm",mass_defect_mode="pos")
)

print(format(Sys.time(), "%a %b %d %X %Y"))
```

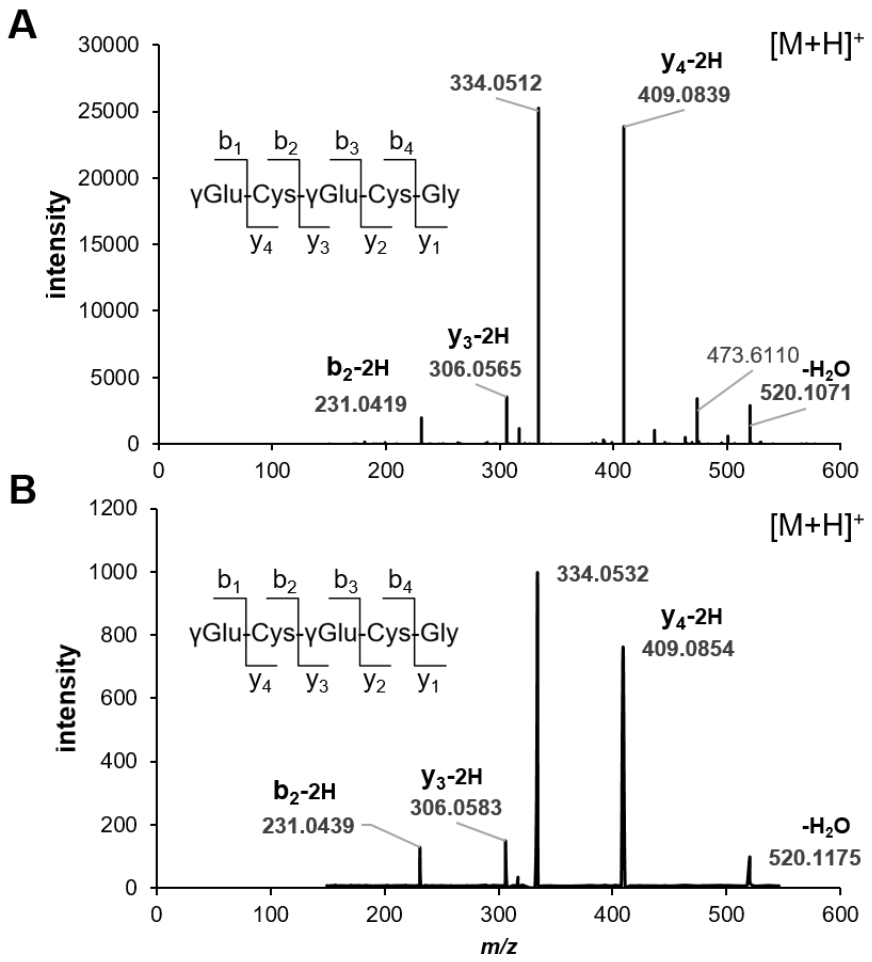

Figure S2.1 (Supplementary). Matching MS/MS spectra of *m/z* 538.1272 [M+H] for PyC_2-Gly authentic standard and onion samples (representative spectra shown). A, Fragmentation pattern for (S-S)PyC_2-Gly [M+H] from the authentic standard sample using collision-induced dissociation (CID) at 25V. B, Fragmentation pattern for (S-S)PyC_2-Gly [M+H] from an onion sample using CID at 35V.

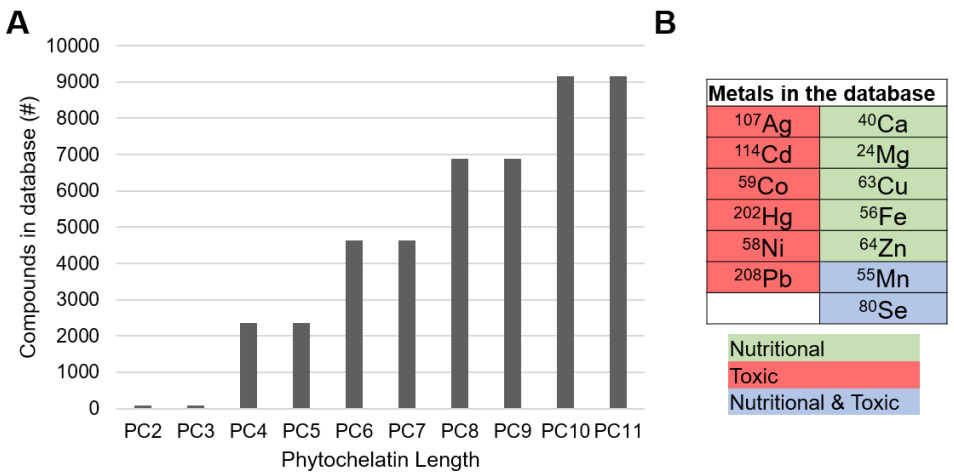

Figure S2.2 (Supplementary). Characteristics of the database. A, Number of compounds in database by phytochelatin length (i.e. repeating peptide units, n = 2-11). With increasing phytochelatin length, the number of possible phytochelatins and phytochelatin-metal complexes increases. B, The most common isotope of metals of nutritional and toxicological significance are included in the database.

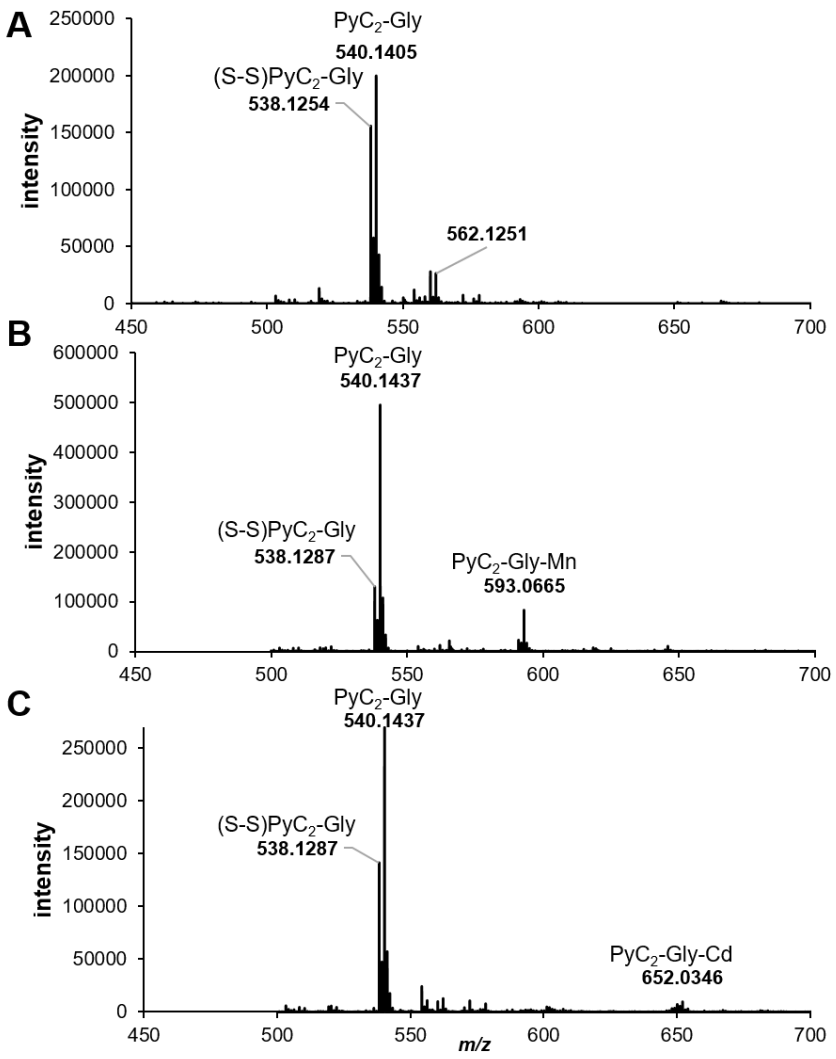

Figure S2.3 (Supplementary). MS1 spectra of PyC_2-Gly authentic standard alone (A) and in combination with equimolar $MnCl_2$ (B) or $CdCl_2$ (C). (S-S)PyC_2-Gly, oxidized form of PyC_2-Gly; PyC_2-Gly, reduced form of PyC_2-Gly; PyC_2-Gly-Mn, PyC_2-Gly complexed with Mn^{2+} ion; PyC_2-Gly-Cd, PyC_2-Gly complexed with Cd^{2+} ion.

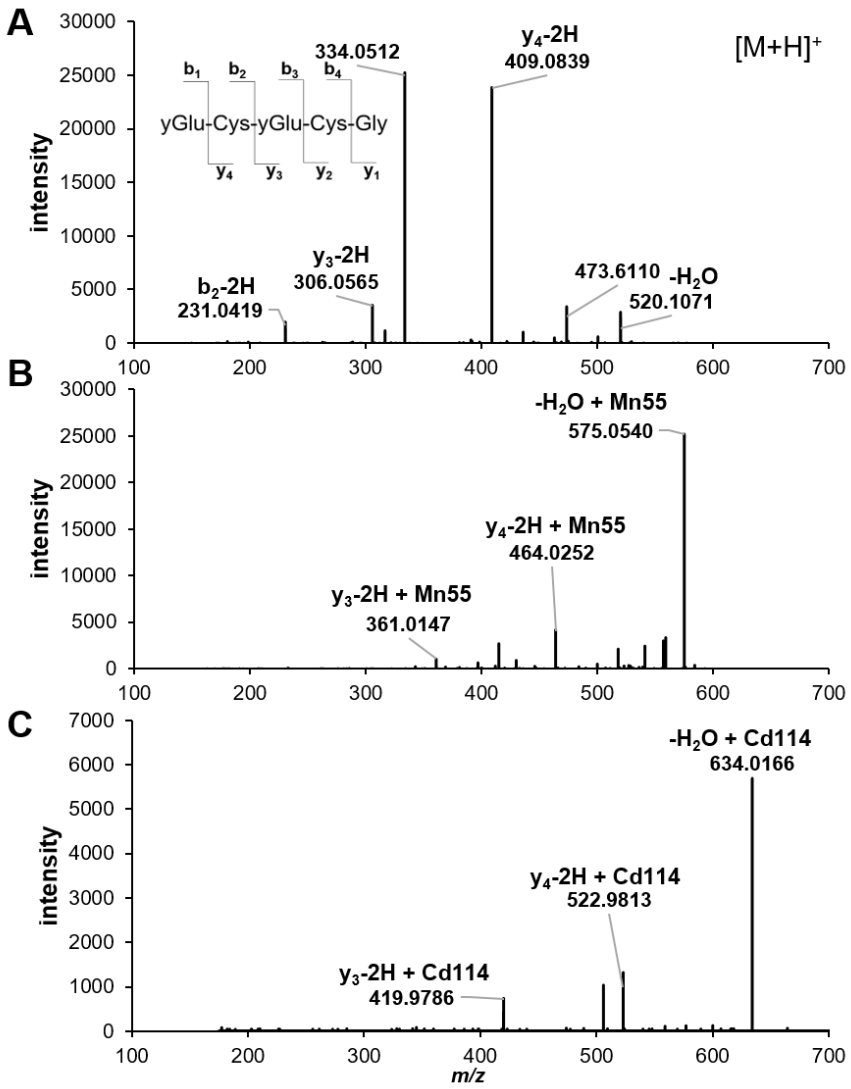

Figure S2.4 (Supplementary). MS/MS of (S-S)PyC$_2$-Gly, PyC$_2$-Gly-Mn, and PyC$_2$-Gly-Cd. A, Fragmentation pattern for (S-S)PyC$_2$-Gly [M+H] (*m/z* 538.13) from the authentic standard sample using collision-induced dissociation (CID) at 25V. B, Fragmentation pattern for PyC$_2$-Gly-Mn [M+H] (*m/z* 593.07) from the authentic standard mixed with equimolar $MnCl_2$ using CID25V. C, Fragmentation pattern for PyC$_2$-Gly-Cd [M+H] (*m/z* 652.03) from the authentic standard mixed with equimolar $CdCl_2$ using CID35V.

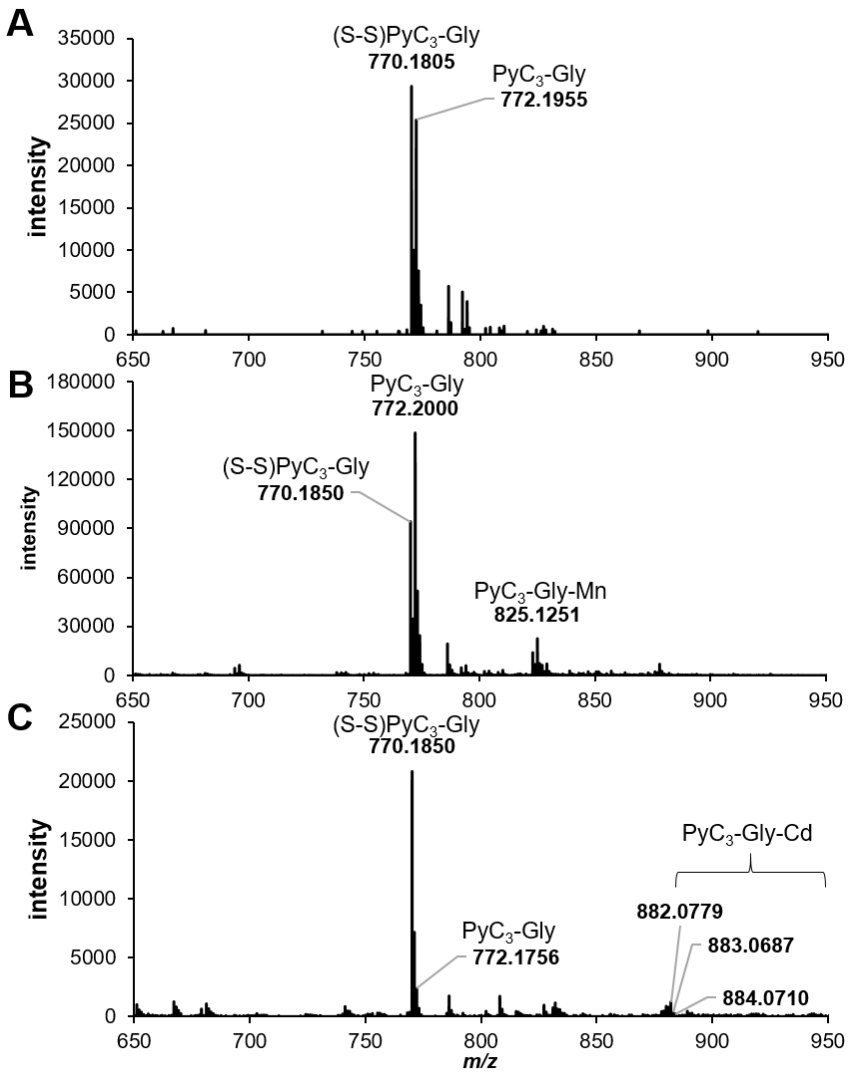

Figure S2.5 (Supplementary). MS1 spectra of PyC_3-Gly authentic standard alone (A) and in combination with equimolar $MnCl_2$ (B) or $CdCl_2$ (C). (S-S)PyC_3-Gly, oxidized form of PyC_3-Gly; PyC_3-Gly, reduced form of PyC_3-Gly; PyC_3-Gly-Mn, PyC_3-Gly complexed with Mn^{2+} ion; PyC_3-Gly-Cd, PyC_3-Gly complexed with Cd^{2+} ion.

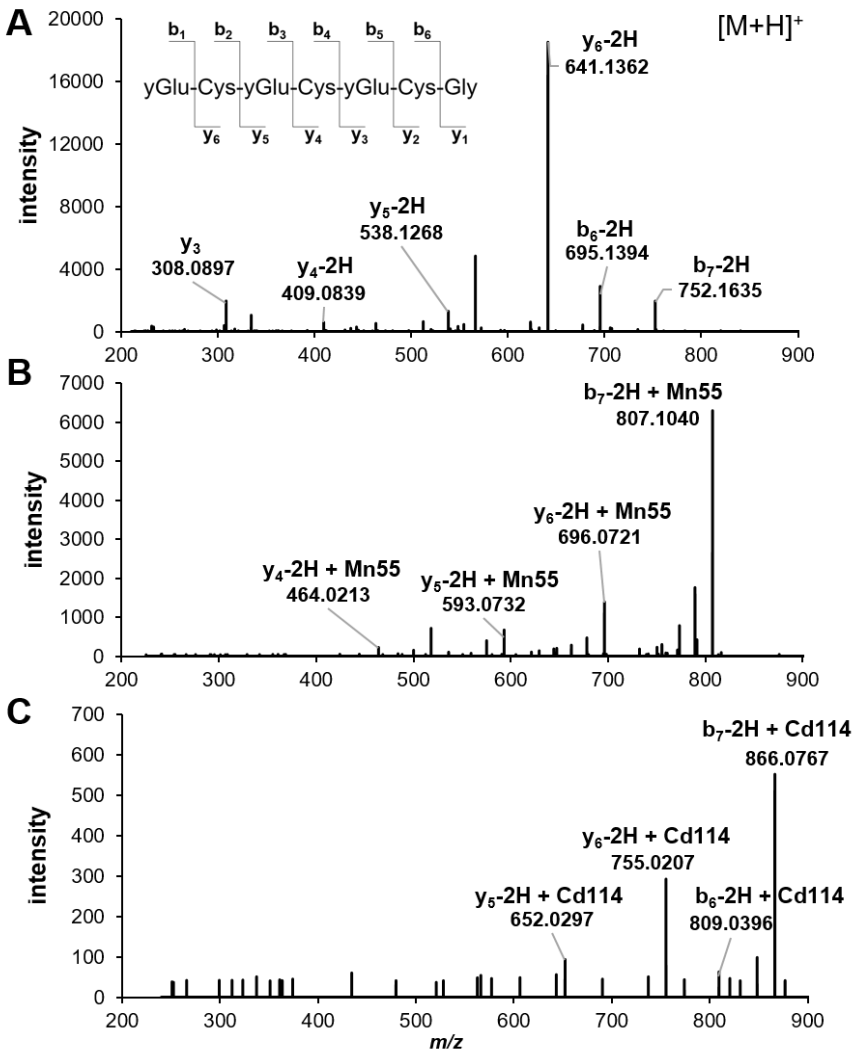

Figure S2.6 (Supplementary). MS/MS of (S-S)PyC$_3$-Gly, PyC$_3$-Gly-Mn, and PyC$_3$-Gly-Cd. A, Fragmentation pattern for (S-S)PyC$_3$-Gly [M+H] (*m/z* 770.18) from the authentic standard sample, (B) for PyC$_3$-Gly-Mn [M+H] (*m/z* 825.12) from the authentic standard mixed with equimolar $MnCl_2$, and (C) for PyC$_3$-Gly-Cd [M+H] (*m/z* 884.09) from the authentic standard mixed with equimolar $CdCl_2$ using collision-induced dissociation (CID) at 25V.

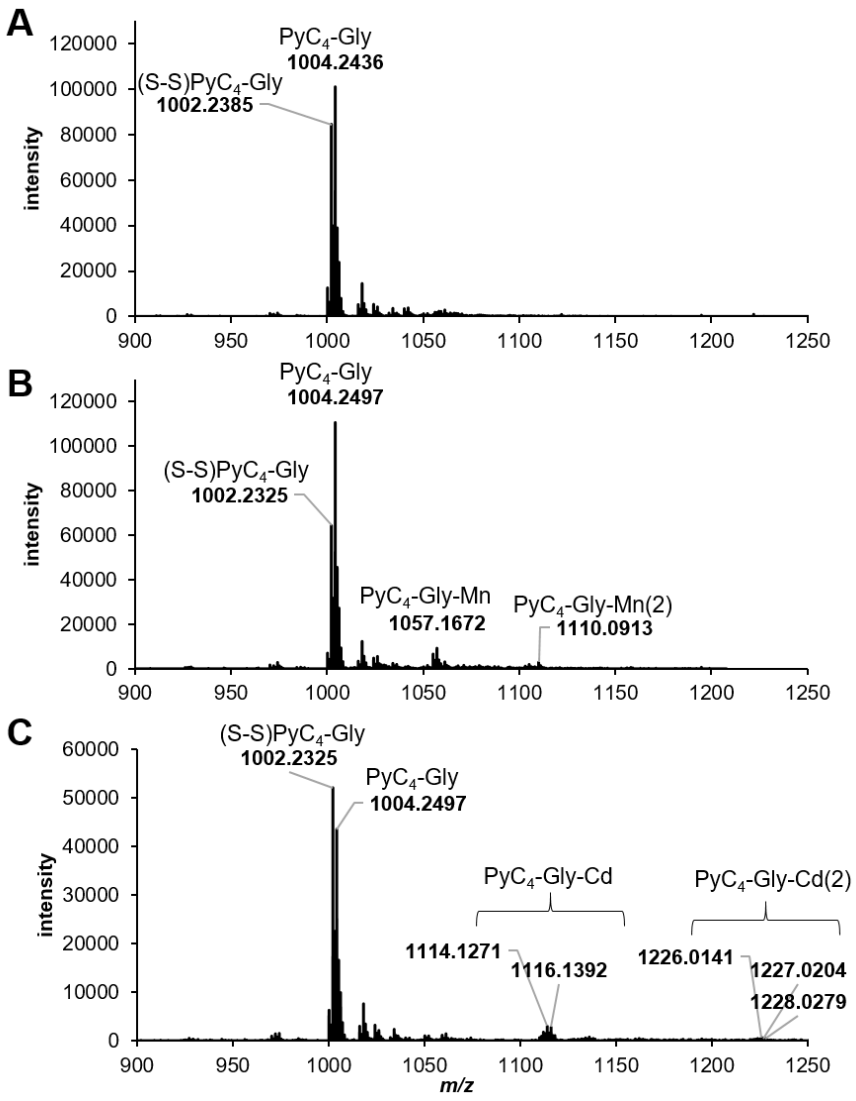

Figure S2.7 (Supplementary). MS1 spectra of PyC_4-Gly authentic standard alone (A) and in combination with equimolar $MnCl_2$ (B) or $CdCl_2$ (C). (S-S)PyC_4-Gly, oxidized form of PyC_4-Gly; PyC_4-Gly, reduced form of PyC_4-Gly; PyC_4-Gly-Mn, PyC_4-Gly complexed with Mn^{2+} ion; PyC_4-Gly-Cd, PyC_4-Gly complexed with Cd^{2+} ion; PyC_4-Gly-Cd(2), PyC_4-Gly complexed with two Cd^{2+} ions.

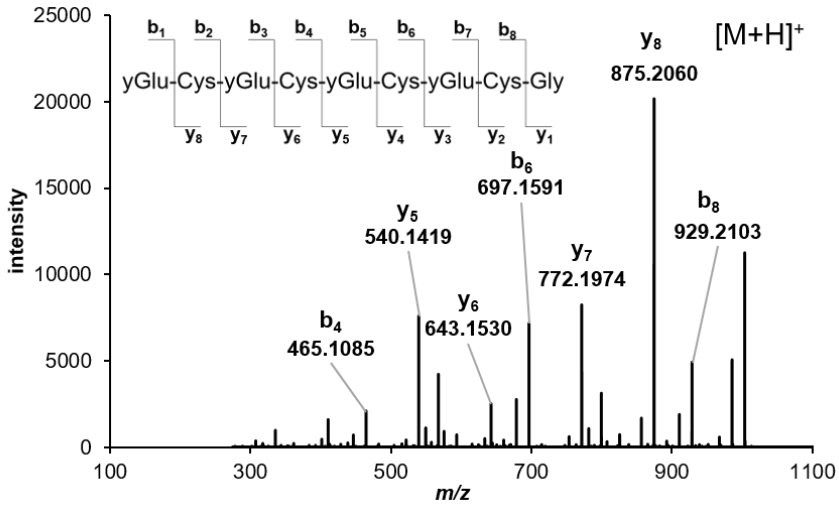

Figure S2.8 (Supplementary). MS/MS of PyC_4-Gly. Fragmentation pattern for PyC_4-Gly [M+H] (*m/z* 1004.25) from the authentic standard sample using collision-induced dissociation (CID) at 25V.

Chapter 2 Supplementary Information References

1. Go YM, Uppal K, Walker DI, Tran V, Dury L, Strobel FH, Baubichon-Cortay H, Pennell KD, Roede JR, Jones DP. Mitochondrial metabolomics using high-resolution Fourier-transform mass spectrometry. Methods Mol Biol. 2014;1198:43-73. doi: 10.1007/978-1-4939-1258-2_4. PubMed PMID: 25270922; PMCID: PMC4318503.

2. Soltow QA, Strobel FH, Mansfield KG, Wachtman L, Park Y, Jones DP. High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. Metabolomics. 2013;9(1 Suppl):S132-S43. doi: 10.1007/s11306-011-0332-1. PubMed PMID: 26229523; PMCID: PMC4517297.

3. Yu T, Park Y, Johnson JM, Jones DP. apLCMS--adaptive processing of high-resolution LC/MS data. Bioinformatics. 2009;25(15):1930-6. doi: 10.1093/bioinformatics/btp291. PubMed PMID: 19414529; PMCID: PMC2712336.

4. Uppal K, Soltow QA, Strobel FH, Pittard WS, Gernert KM, Yu T, Jones DP. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. BMC Bioinformatics. 2013;14:15. doi: 10.1186/1471-2105-14-15. PubMed PMID: 23323971; PMCID: PMC3562220.

5. Cobbett C, Goldsbrough P. Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. Annu Rev Plant Biol. 2002;53:159-82. doi: 10.1146/annurev.arplant.53.100301.135154. PubMed PMID: 12221971.

6. Serrano N, Diaz-Cruz JM, Arino C, Esteban M. Recent contributions to the study of phytochelatins with an analytical approach. Trends in Analytical Chemistry. 2015;73:129-45.

7. Uppal K, Walker DI, Jones DP. xMSannotator: An R Package for Network-Based Annotation of High-Resolution Metabolomics Data. Anal Chem. 2017;89(2):1063-7. doi: 10.1021/acs.analchem.6b01214. PubMed PMID: 27977166; PMCID: PMC5447360.

8. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J. Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ Sci Technol. 2014;48(4):2097-8. doi: 10.1021/es5002105. PubMed PMID: 24476540.

9. Wolf S, Schmidt S, Muller-Hannemann M, Neumann S. In silico fragmentation for computer assisted identification of metabolite mass spectra. BMC Bioinformatics. 2010;11:148. doi: 10.1186/1471-2105-11-148. PubMed PMID: 20307295; PMCID: PMC2853470.

10. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. J Cheminform. 2016;8:3. doi: 10.1186/s13321-016-0115-9. PubMed PMID: 26834843; PMCID: PMC4732001.

11. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH. PubChem Substance and Compound databases. Nucleic Acids Res. 2016;44(D1):D1202-13. doi: 10.1093/nar/gkv951. PubMed PMID: 26400175; PMCID: PMC4702940.

CHAPTER 3

Distribution of phytochelatins, metal-binding compounds, in plant foods: a survey of commonly consumed fruits, vegetables, grains and legumes

Kristine K. Dennis[a], Ken H. Liu[b], Karan Uppal[b], Young-Mi Go[b], Dean P. Jones[a,b]

[a]Nutrition and Health Sciences, Laney Graduate School, Emory University, 615 Michael Street, 225 Whitehead Biomedical Research Building, Atlanta, GA, 30322, USA

[b]Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, Emory University, 615 Michael Street, 225 Whitehead Biomedical Research Building, Atlanta, GA, 30322, USA

Published in *Food Chemistry*. 2020.

doi: 10.1016/j.foodchem.2020.128051

Format adapted for this dissertation

Abstract

Phytochelatins (PyCs) are metal-binding compounds produced by plants. PyCs may reduce bioavailability of dietary toxic metals such as cadmium. However, the PyC concentrations in foods are unknown. The objective of this study was to analyze PyC contents in a subset of commonly consumed plant foods. Foods (20) across five groups were analyzed and PyCs quantified using liquid chromatography-mass spectrometry (LC-MS/MS). The impact of factors such as food processing were also explored. PyCs were in all 20 foods. Five PyC types were detected with PyC_2-Gly, PyC_3-Gly and PyC_2-Ala at quantifiable concentrations. PyC_2-Gly was found at the highest concentrations and most widely distributed. PyC_2-Gly concentrations were highest in fruits and root vegetables. Foods with increased processing tended to have reduced PyC concentrations. This survey of commonly consumed plant foods in the United States demonstrates PyCs are widely distributed and provides a foundation for understanding their concentrations and impact in the human diet.

Keywords: toxic metals, nutritional metals, metal bioavailability, functional foods, plant-based nutrition, phytochemicals

Chemical compounds studied in this article: Phytochelatin (PubChem CID: 20756463); Phytochelatin 2 (PubChem CID: 14704021); Phytochelatin 3 (PubChem CID: 174541); Phytochelatin 4 (PubChem CID: 14704026)

Introduction

Dietary patterns emphasizing plant foods are consistently associated with reduced risks of cardiovascular disease, cancer, type 2 diabetes, and other chronic diseases (1, 2). Plant foods contain health-promoting phytochemicals including flavonoids, carotenoids, and phenolic acids (3). Phytochemicals contribute to a range of protective mechanisms including regulation of inflammatory pathways, antioxidant functions, and cancer-preventing pathways via alterations in gene expression impacting cell proliferation and apoptosis (3). Through eating plant foods, individuals consume a diverse array of phytochemicals which can work synergistically or alone to promote internal biological environments with antioxidant, anticancer, and anti-inflammatory activities (3). However, a major barrier to studying the health impacts of consuming phytochemicals is understanding the distribution of specific compounds across the wide range of plant foods (3).

Phytochelatins (PyCs), are glutathione (GSH)-derived metal-binding polymers produced by plants, and are one class of phytochemicals not characterized in the human diet (4). PyCs consist of 2-11 repeating γ-glutamyl-cysteine (γ-Glu-Cys) peptides and are enzymatically formed by phytochelatin synthase (4). Plants produce PyCs at basal levels and at increased levels as a protective response to metal exposure (5). The peptide repeats typically end with a terminal Gly but can also terminate with Ser, Glu, Gln, Ala, β-Ala, or no amino acid (5, 6). Common names for PyCs include a subscript for the number of repeating γ-Glu-Cys units and the terminal amino acid (7).

A distinctive feature of PyCs is their high Cys content. Metals bind with high affinity to the thiol (-SH) group of Cys at biological pH (close to neutral or slightly basic), resulting in significant capacity to bind and sequester metals in plants (8). For example, PyC_2-Gly [$(\gamma$-Glu-Cys$)_2$-Gly] can bind a metal with an oxidation state of +2. With longer chain lengths, PyCs have increased metal-binding capacity (8). As longer PyCs are formed through the addition of γ-Glu-Cys peptides in response to metal exposure, the most abundant PyCs tend to be the shorter chain lengths (4).

Due to their metal-binding properties, PyCs may have an important role in the absorption of dietary metals. For example, average dietary cadmium (Cd) levels in the U.S. population are linked with poor health outcomes across multiple organ systems due to the long half-life of Cd and its accumulation in the body (9). Dietary factors which reduce Cd bioavailability could have an important role in reducing Cd-related health risks such as kidney dysfunction and cancer (9). Studies of PyCs in mammalian systems have demonstrated PyCs protect against uptake of Cd. Researchers investigating the impact of PyC_3-Gly and PyC_5-Gly on Cd accumulation in rats observed reduced Cd in the liver and kidneys, common target organs for Cd accumulation, compared to rats exposed to only Cd (10). Investigations in gastrointestinal cell models also demonstrated reduced cellular Cd accumulation when cells were exposed to PyC_3-Ser in combination with Cd (11). This evidence suggests consumption of PyCs in the diet may protect against Cd uptake (10, 11).

PyCs have been studied extensively in agricultural and bioremediation in efforts to improve soil and plant health. From this body of research, PyCs are known to occur in agricultural plants grown for human consumption, such as soybeans, corn, and rice (4, 6). However, these studies have focused on roots, stems, and leaves from plant seedlings exposed to heavy metals. There is little known about the amount of PyCs in the edible components of harvested mature plants. Studies show PyCs tend to concentrate in roots as they are produced in response to metals taken up from the soil (6, 12). PyCs are also produced in or transferred to leaves and stems but typically at lower concentrations than roots (13, 14). The PyC types, distributions, and concentrations ultimately depend on a variety of factors including the plant species and growing conditions. Some plant families predominantly have PyCs with terminal amino acids other than Gly, such as Ala in *Fabaceae* (legumes) or Ser in *Poaceae* (cereal grasses/grains), and evidence demonstrates plants can produce multiple PyC types (6, 15, 16). The PyC length and terminal amino acid influence the molecular environment of the metal-binding thiol group and consequently, the affinity with which metals bind to PyCs (7). These could be important factors for understanding the impact of dietary PyCs on metal absorption in humans. For plants, soil heavy metal concentrations have the greatest influence on PyC production. In a study of plant communities along soil pollution gradients, plants had greater PyC content with increasing soil heavy metals (17). However, plants do not require heavy metal exposures to produce PyCs and will produce PyCs in response to essential metals (e.g. zinc, copper) (15).

The human diet is a mixture of exposures, including beneficial components such as fiber and phytochemicals, but also potentially harmful components such as pesticides, herbicides, and toxic metals. As with other phytochemicals, the health impact of PyCs will depend on the

characteristics and concentrations of PyCs, other dietary components consumed, and nutritional status of the individual (3). Additionally, food processing impacts phytochemical concentrations and will likely influence PyCs as well (18). The present study was designed to establish an understanding of PyC concentrations and types in commonly consumed foods in the U.S. diet for the first time. We tested the hypothesis that PyCs most directly related to the precursor GSH, i.e., PyC_2-Gly and PyC_3-Gly, are widespread among plant-derived foods, while other PyCs (PyC_4-Gly, PyC_5-Gly, PyC_6-Gly, PyC_2-Glu, PyC_2-Ala) have more restricted distributions. Additionally, we examined how factors like food group, growing conditions, and food processing may impact PyC concentrations. Here we provide characterization of PyCs in twenty commonly consumed foods across five plant food groups. Through this work, we provide evidence of important food characteristics that impact PyC concentrations, demonstrate the widespread distribution of PyCs in plant foods, and establish relevant dietary PyC concentrations and types for understanding the impact of this relatively unexplored class of phytochemicals on human health.

Material and Methods

Materials

Acetonitrile (HPLC grade), formic acid (LC-MS grade), and water (HPLC grade) were purchased from Sigma Aldrich (St. Louis, MO, USA). HPLC grade water was prepared with 2% formic acid as a solvent for LC-MS/MS. Analytical standards of PyC_2-Gly, PyC_3-Gly, PyC_4-Gly, PyC_5-Gly, PyC_6-Gly, PyC_2-Glu, and PyC_2-Ala at 95% purity were obtained from CPC Scientific Inc (San Jose, CA, USA). Individual stock solutions of PyCs were prepared in HPLC-grade water.

Selection of representative foods

Five plant food groups were selected based on expected distribution of PyCs due to known plant physiology as well as colloquial terms for common vegetable types (i.e. leafy greens, root vegetables, fruits, legumes, and grains). From each of the five plant food groups, four food types were selected based on commonly consumed foods in the U.S. in each group (19). Tomatoes were included with the fruit group due their designation in plant anatomy. Corn was considered in the grain group even though ground corn products and fresh corn are often categorized into different food groups.

For each food type, ten individual items were purchased from local grocery stores in Atlanta, Georgia (May 2019 to August 2019), totaling 40 samples per group. Food types included corn, brown rice, whole wheat flour, oats, oranges, apples, bananas, tomatoes, carrots, potatoes, onions, sweet potatoes, kale, romaine lettuce, head lettuce, spinach, peanuts, beans, tofu, and peas. Twelve individual items were purchased for apples to increase diversity for subset comparisons as described below. For each food, items were purchased with at least three different brand names or stores for analysis. To address PyC variability due to storage conditions and processing, a selection of fresh, frozen, and canned options was included when available (**Table S3.1**). To explore other factors that could impact PyC concentrations, foods were selected with different varieties, growing locations, and growing conditions (e.g. conventional and organic). Fresh foods were kept refrigerated at 4ºC and frozen foods were kept frozen at -20ºC. Shelf stable items such as dry and canned foods were kept at room temperature until prepared for analysis.

Sample preparation

Samples were prepared from portions of the food item generally considered edible in the United States. For foods usually consumed cooked (i.e. rice, sweet potatoes, potatoes), foods were prepared using typical cooking directions. In brief, sweet potatoes and potatoes were baked in foil at 350ºF for 40-50 minutes until softened. Brown rice was cooked in 2:1 ratio of filtered water to rice by bringing to a boil and then reducing to a simmer for 40 minutes. If necessary, foods were cut before grinding with a mortar and pestle. Liquid nitrogen was used to freeze fruits, leafy greens, and root vegetables prior to grinding to facilitate breakdown of fibrous materials and transfer to extraction vials.

Ground food samples were loaded into 2 mL microcentrifuge tubes, weighed, and stored at -80ºC until extracted for LC-MS/MS analysis. A 2:1 acetonitrile to water solvent mixture was added to ground food samples at a ratio of 2 µL solvent to 1 mg food. Samples were vortexed, sonicated (10 seconds at 30% amp), incubated on ice for 30 min, and centrifuged for 10 min (14,000 rpm) at 4ºC. The supernatant was removed and stored at -80ºC until LC-MS/MS analysis. Sample extraction following reduction with dithiothreitol was tested and not included in the final method as it did not improve PyC recovery.

LC-MS/MS analysis

Food sample supernatants were thawed, vortexed, and loaded into autosampler vials. Pooled samples for each food type were mixed from the supernatant for use as a quality control reference, and seven PyC standards were added at 1 µM for use during quantification. Samples

were randomized within each food type and analyzed on a high resolution LTQ-Velos Orbitrap mass spectrometer (Thermo Fisher). Pooled food samples with and without added PyC standards were analyzed at the end of each food type analysis for estimating PyC concentrations and validating PyCs with chromatographic characteristics and ion dissociation mass spectrometry (MS/MS). Each sample was analyzed in triplicate with a 10 µl injection volume on a HILIC column in positive electrospray ionization mode. The mass spectrometer was operated at 60,000 resolution and mass-to-charge (*m/z*) range of 500-1,250 following optimization for PyC detection and quantification.

Phytochelatin quantification

The intensity of each PyC (M+H adduct) was determined using the area under the curve (Thermo Scientific Xcalibur software, Genesis peak integration algorithm) (**Table S3.2**) (7). For some samples, reduced and oxidized PyC forms of the M+H adduct were detected. Tests with dithiothreitol treatment did not improve reproducible quantification and introduced an additional step in sample preparation, so the approach used for these samples was to sum the intensities for the multiple forms [e.g. PyC_2-Gly: *m/z* 538.1272 (oxidized) and *m/z* 540.1429 (reduced)]. Comparisons of intensities for commercial standards treated in this manner showed this to be reliable, i.e., disulfide bonds form internally within PyC and the signal of the oxidized form is not decreased due to larger aggregate formation. Technical replicates were averaged. Quantification used a method of additions. To account for food specific matrix effects for each food item, intensity values of 1 µM PyC authentic standards were determined by subtracting the intensity of the pooled sample from the intensity of the pooled sample with PyC standards. Estimates of PyC concentrations in each food item were then determined relative to the intensity

of the added 1 µM PyC standard. Samples with MS/MS patterns matching the characteristic fragmentation pattern of the respective PyC standard but within an unreliable range for quantification were designated as "detected but not quantified".

PyC quantities are reported as µg/g edible weight (EW), which refers to the weight of the food item as purchased or prepared. For example, oats and whole wheat flour are reported using the dry weights, potatoes are reported using the prepared weight (i.e. canned or baked), and apples are reported using the fresh weight. PyC quantities per serving size were calculated using the Food Pyramid Equivalent Database (FPED) serving equivalents converted to grams. FPED equivalents for fruits and vegetables are in 1 cup equivalents and for protein and whole grains are in 1 ounce equivalents (20).

Statistical analysis

One-way ANOVA with Tukey's post-hoc test was used to compare PyC_2-Gly concentrations across food groups (e.g. fruits, leafy greens, root vegetables, grains, legumes) and within food groups (e.g. apples, oranges, bananas, tomatoes), PyC_3-Gly and PyC_2-Ala concentrations across food types, and any food types with multiple varieties or processing levels. Unpaired t-tests were used to compare PyC concentrations between two groups (e.g. red kale vs. curly kale; fresh potatoes vs. canned potatoes). Pearson's correlation was used to assess the association between concentrations of multiple PyCs within one food type. All statistical analyses were performed using the software GraphPad Prism version 8.3.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com. Significant differences were defined as p-values < 0.05.

Results and Discussion

Overall distribution of PyCs in commonly consumed foods

Our results establish the novel finding that PyCs are widely distributed in plant foods. Foods from five plant food groups were analyzed with four food types per group and ten food items per type, totaling 202 individual food items (**Table 3.1**). PyCs were found in all five food groups and twenty food types analyzed, demonstrating PyCs are consumed via a variety of foods. Five PyC types were detectable and/or quantifiable in these foods. PyC_2-Gly was the predominant PyC type with quantifiable concentrations in 18 of 20 food types and 182 of 202 food items. In alignment with previous studies in plant seedlings, the four other PyC types measured occurred much less frequently (4). Quantifiable levels of PyC_3-Gly were in six food types and PyC_3-Gly was detectable but not quantifiable in one additional food. PyC_2-Ala was quantifiable in four food types and detectable but not quantifiable in an additional four types (**Figure S3.1**, **Table S3.3**). PyC_4-Gly and PyC_2-Glu were detectable but not quantifiable in five and three food types, respectively (**Table S3.3**). Long-chain PyCs, PyC_5-Gly and PyC_6-Gly, were not detected in any foods. Other studies have not observed long-chain PyCs in plant seedlings grown under control conditions and have observed only trace quantities of long-chain PyCs when plants are grown with exposure to heavy metals (4, 21).

PyC_2-Gly quantification in 18 of 20 food types

As PyC_2-Gly is one of the most common PyCs in the plant kingdom, we focused our analysis on PyC_2-Gly concentrations to determine if the distribution in foods was similar. PyC_2-Gly concentrations were highly variable across the five food groups (one-way ANOVA, $p < 0.05$)

with fruits and root vegetables having the highest PyC_2-Gly concentrations followed by grains, leafy greens, and legumes (**Figure 3.1A**). In root vegetables and fruits, average PyC_2-Gly concentrations were 2.55 µg/g EW (range 0.09-15.26) and 2.63 µg/g EW (range 0.13-8.85), respectively. Root vegetables and fruits had significantly higher PyC_2-Gly concentrations than average legume concentrations (0.30 µg/g EW, range 0.00-2.25, Tukey's, $p < 0.05$). Average PyC_2-Gly concentrations in leafy greens and grains were similar (1.24 µg/g EW, range 0.07-5.44; 1.38 µg/g EW, range 0.00-15.01). The highest PyC_2-Gly concentrations in individual food items were distributed across fruit and vegetable food groups; in descending order of PyC_2-Gly content were carrots, corn, oranges, apples, onions, and romaine lettuce (**Figure 3.1B, C, E, F**). The lowest PyC_2-Gly concentrations were in grains and legumes with no detection of PyC_2-Gly in two legume food types, tofu and beans.

Root vegetables

PyCs are produced by plants in response to soil metals. Previous studies demonstrate PyCs frequently concentrate in or near roots relative to stems and leaves (6, 12). Therefore, we investigated whether edible components of plants which grow closer to the roots would have higher concentrations of PyCs. The four types of root vegetables varied significantly in their PyC_2-Gly concentrations (**Figure 3.1B**). Of the two true roots in our study, carrots had high PyC_2-Gly concentrations with an average of 6.36 µg/g EW. Sweet potatoes, the only other true root of the root vegetables studied, had the lowest average PyC_2-Gly concentrations of the root vegetables at 0.24 µg/g EW with consistently low concentrations across all samples (range 0.09-0.61). Onion and potato samples had similar concentrations to fruits and leafy greens with similar average concentrations and variability across individual samples (2.53 µg/g EW, range

1.22-4.22 and 1.07 µg/g EW, range 0.24-3.35, respectively). These findings suggest that PyC types and concentrations in plant foods are specific to each food and more systematic and complete assessment should be conducted before generalizing to categories of plant food groups.

Grains and legumes

Grains are an important source of nutrition globally, accounting for 50% of calories consumed worldwide and therefore, may serve as an important source of PyCs in the diet. Legumes are a major source of protein in many low- and middle-income countries and have been identified in the U.S. dietary guidelines as an important component within a healthy dietary pattern. Although anatomically distal plant components like grains and legumes are farther from metal exposures at the roots, phytochelatin synthase, the enzyme producing PyCs, is expressed constitutively throughout the plant (22). Among the grains measured, PyC_2-Gly concentrations were very low except in fresh and canned/frozen corn samples (**Figure 3.1C**). Within corn samples, average PyC_2-Gly concentrations were 5.20 µg/g EW and there was high variability in concentrations driven by the very low PyC_2-Gly levels in cornmeal/flour and high levels in fresh corn (range 0.01-15.01). Legumes had low concentrations of PyC_2-Gly (**Figure 3.1D**). Peas had an average PyC_2-Gly concentration of 1.16 µg/g EW (range 0.37-2.25) and peanuts had an average of 0.03 µg/g EW (range 0.00-0.13). The other two legumes studied, tofu and beans, had no detectable PyC_2-Gly. Other studies in beans and soybeans, from which tofu is made, found PyC-Ala but not PyC-Gly (23, 24).

Fruits

Among fruits, oranges had the highest average PyC_2-Gly concentration at 5.09 µg/g EW (Tukey's, $p < 0.05$) (**Figure 3.1E**), followed by apples at 3.23 µg/g EW. Oranges and apples had similar variability in concentrations across samples (range 3.14-8.85 and 0.87-8.17, respectively). Tomatoes had an average PyC_2-Gly concentration of 1.67 µg/g EW (range 0.67-3.52) and bananas had the lowest PyC_2-Gly concentrations of all fruits at 0.40 µg/g EW (range 0.13-0.83). Although no previous studies have examined the PyC concentrations in apple, orange, or banana plants or their fruits, the phytochelatin synthase gene ortholog has been documented in each species (25). Additionally, PyCs have been identified in the roots of 1-year old citrus plants grown with and without Cd exposures (26). Similar studies in tomato plants have identified PyCs in leaves and roots with and without Cd exposures (27, 28). In a study of tomato seedlings, PyC_2-Gly concentrations were about 5 µg/g fresh weight (FW) in leaves and roots with no Cd exposure (27). The slightly higher concentrations observed in the tomato seedlings study are in alignment with previous studies demonstrating roots and young leaves have higher concentrations relative to old leaves and stems (24, 29).

Leafy greens

Compared to fruits and grains, leafy greens are anatomically at an intermediate location between roots and the distal fruiting bodies of plants. Additionally, leafy greens are known to have higher concentrations of Cd than other plant foods. Therefore, we hypothesized leafy greens may have intermediate to high PyC concentrations based on plant anatomy and metal accumulation. Leafy greens actually had lower PyC_2-Gly concentrations relative to root vegetable and fruit groups

(**Figure 3.1A**). Romaine lettuce had the highest average PyC_2-Gly concentration of 2.50 µg/g EW followed by head lettuce at 1.02 µg/g EW. Spinach had the lowest average PyC_2-Gly concentrations of the leafy greens studied (0.54 µg/g EW). The average kale concentration was 0.90 µg/g with one kale sample having a 5-fold higher PyC_2-Gly concentration (5.44 µg/g EW) (**Figure 3.1F**). High exposure to soil heavy metals while growing could explain the high concentrations in this kale plant. In a study of lettuce, increasing Cd exposures led to 2 to 10-fold higher PyC content relative to unexposed plants (29). Other studies of lettuce have found zero to very low levels of PyCs in leaves unexposed to heavy metals with total PyC concentrations around 0.1 µg/g FW (29). When plants are exposed to Cd in hydroponic growing systems, the concentrations increase to about 1.3 µg/g FW in new leaves (29). The low PyC levels may be due to the experimental growing conditions in these studies. In contrast to our findings, a study which grew both spinach and lettuce in a hydroponic system with the same exposures observed the highest PyC concentrations in Cd-exposed shoots and roots of spinach, not lettuce (30). Nutrients in hydroponic systems differ from soil. As other soil metals like copper and zinc induce PyC generation, the difference in our results may reflect the impact of growing leafy greens in soil versus hydroponic conditions.

Difference in PyC types and concentrations across foods

As hypothesized, short chain PyCs (PyC_2-Gly and PyC_3-Gly) are commonly distributed and present at the highest concentrations across all foods studied. Of the subset of six food types with multiple quantifiable PyCs, PyC_2-Gly was at significantly higher concentrations than other PyCs in all foods (**Figure 3.2**). PyC_2-Gly levels were 10 to 34-fold higher than PyC_3-Gly in brown rice, tomatoes, oranges, peas, corn, and potatoes. Of the six foods with quantifiable PyC_3-Gly

concentrations, the highest concentrations were in oranges followed by corn and potatoes (**Figure S3.1A**). In alignment with other studies, the PyC-Ala type predominated in legumes, occurring as the only detectable PyC in tofu and beans (15, 23, 24). In a study including beans and peas, peas contained PyC-Gly and PyC-Ala but beans only contained PyC-Ala (23). Of the four foods with PyC_2-Ala, three of the food types were legumes, with tofu containing the highest concentrations (**Figure S3.1B**). No PyC_2-Ala was detected in peanuts but was detectable in other non-legume foods including corn, brown rice, and apples. PyCs other than PyC-Gly have been previously detected in corn and brown rice, including PyC-Ser and PyC-Glu, but to our knowledge, PyC-Ala has not (6, 21). In foods with multiple quantifiable PyCs, PyC concentrations were positively correlated (**Figure S3.2**). In a study of arsenic-exposed rice varieties, PyC types and chain lengths were also positively correlated (16).

Difference in PyC content by level of food processing

The impact of food processing on PyC content has not been described. However, food processing is known to decrease GSH content (31), the precursor from which PyCs are formed. In alignment with previous research on GSH, foods with increased levels of food processing tended to have lower PyC_2-Gly and PyC_3-Gly content (**Figure 3.3**) (31). This trend was particularly pronounced in corn. Fresh corn had some of the highest PyC_2-Gly concentrations while dried and milled corn had some of the lowest (**Figure 3.3A**). The milling technique could influence PyC concentrations as more or less of the corn kernel will be included in the resulting cornmeal or flour. Also, different corn varieties are used for fresh and canned corn versus milled corn which may contribute to this difference. With higher levels of processing, PyC_2-Gly concentrations were consistently lower in peas and potatoes (**Figure 3.3A**). A significant trend was observed in

peas across step-wise increases in food processing. This may be explained by thermal degradation, a process which frequently reduces phytochemical concentrations (18). Dry grain products in our study had low PyC concentrations which may be due to PyC degradation as these products are stored at room temperature. In our study, fresh peas, corn, and potatoes had the highest concentrations with decreasing concentrations in frozen and canned foods, respectively. Lower PyC concentrations in frozen foods relative to fresh may be due to blanching prior to freezing and packaging (18). Cooking and canning can lead to greater phytochemical reductions due to increased thermal degradation with higher temperatures and longer cooking times (18, 32). In a study of GSH in foods, raw spinach had the highest total GSH concentration (394 nmol/g) followed by cooked (234 nmol/g) and canned (71 nmol/g) (33). This study observed a similar pattern of total GSH concentrations in carrots (raw, 255 nmol/g; cooked 188 nmol/g; canned, 0 nmol/g) (33). Studies of other compounds have found similar patterns, with reductions in phytochemical content with increased food processing (31, 34). Thermal processing could also disrupt the enzymatic activity of phytochelatin synthase, disrupting additional PyC synthesis which could otherwise occur in fresh foods. In some cases, thermal degradation can increase compound extraction through disrupting the food matrix and improving bioaccessibility (18). In our study, tomatoes had slightly higher PyC_2-Gly concentrations in canned as compared to fresh (canned: 2.38 µg/g EW, range 1.62-3.52; fresh: 1.36 µg/g EW, range 0.67-2.50; $p > 0.05$). This may be due to increased extractability of PyCs from the canning process. Additionally, canning may concentrate PyCs in tomatoes more than other canned plant foods like potatoes and peas due to their high water content. Similar patterns were observed with PyC_3-Gly concentrations for these foods (**Figure 3.3B**). Given the known influence and specificity of food processing on

phytochemical concentrations, food processing will likely impact PyC concentrations in other plant foods.

Comparison of PyC concentration by growing conditions and food variety

As PyCs are enzymatically produced by plants in response to their environment, we considered how growing conditions or food variety may impact PyC concentrations. Soil metal concentrations are known to impact PyC concentrations and vary geographically. When we examined how growing locations and organic versus conventional farming may impact PyC concentrations, no differences were found (**Figure S3.3, S3.4A**). Agricultural growing practices do impact soil characteristics and metal exposures but conventional and organic labelling may not capture meaningful differences (35). Additionally, the data available about growing locations were limited to state or country level which may not be granular enough to observe differences. For example, specific geographic regions may have increased heavy metals due to nearby manufacturing processes emitting heavy metals that deposit in nearby soils (35).

As indicated by the trends we observed, plant variety may impact PyC content in some foods. Many plant foods have multiple varieties from the same species widely available in grocery stores. For example, apple varieties include gala, fuji, red delicious, and many others. Our data indicate there may be some difference by food variety in apples, kale, and potatoes but these differences were not significant (**Figure S3.4B**). In a study examining the response of six rice varieties to the same Cd exposures, PyC concentrations varied up to 10-fold (16). More broadly,

phytochemical content is known to differ across common plant food varieties (36). Our data indicate plant variety will be important to consider when studying PyCs in human foods.

PyC_2-Gly content by serving size

To understand the relative proportion of PyC_2-Gly obtained in the diet, we calculated PyC_2-Gly for a typical serving using the mean PyC_2-Gly concentrations in foods and the serving size from USDA Food Pyramid Equivalents Database. Fruits and vegetables had the highest quantity of PyC_2-Gly (**Table 3.2**). The whole grain and protein foods had very low PyC_2-Gly amounts per serving with the highest quantity coming from brown rice (**Table 3.2**). Raw corn and raw carrots had the highest quantity of PyC_2-Gly with over 1000 µg/serving. Oranges, canned tomatoes, frozen or canned corn, and onions followed in descending order with over 400 µg/serving. As previously discussed, we observed higher quantities of PyCs in less processed foods. Raw corn, carrots, and peas had four times the PyC_2-Gly per serving than the canned version, and baked potatoes had two times the quantity relative to canned. In tomatoes, the opposite relationship was observed with canned tomatoes containing two times the PyC_2-Gly per serving than fresh tomatoes.

The values obtained indicate that the average amounts of PyC_2-Gly in the human diet could range from 0.38 mg to 3.75 mg for individuals consuming less than 3 servings of fruits and vegetables per day to 0.88 mg to 8.75 mg for individuals consuming 7 or more servings per day. This may be highly impacted, however, by specific food selections. Of importance for mechanistic studies, the measured amounts suggest that relevant PyC concentration ranges in the

intestinal lumen are 0.35 µM to 8.11 µM if one assumes dilution in 2 L from dietary intake and 0.08 µM to 1.80 µM if one assumes dilution in 9 L from endogenous and exogenous sources (37).

This study establishes PyC types and amounts coming from specific plant foods in the diet. Open questions remain about how the metal-binding properties of PyCs influence metal bioavailability. From previous research on Cd and PyCs, we know PyC exposure can lead to reduced absorption of Cd (10, 11). However, how PyCs impact bioavailability of other toxic metals or essential metals is unknown. Studies of iron bioavailability indicate compounds structurally similar to PyCs, GSH and Cys, can enhance iron absorption (38). Although structurally different, phytic acid is another compound produced by grains, legumes, nuts, and some vegetables with metal-binding properties, which provides context and considerations for studying PyC health effects. Mineral bioavailability studies indicate the impact of phytic acid is highly dependent on dietary context. For example, absorption of minerals may not be negatively impacted in individuals who are well-nourished and eating a diverse diet (39). The affinity of binding will depend on the metal, ratio of phytic acid to metal, and other factors such as pH. Similar factors including gut microbiome activity are likely to influence the availability and function of PyCs in the gastrointestinal tract (7). Other health effects of PyCs could include a role in nutritional immunity wherein the host controls access to micronutrients as a mechanism to protect from bacterial infections. Additionally, the structural similarity of PyCs to GSH, a major antioxidant, suggests PyCs may impact cellular redox mechanisms and availability of redox metals.

Conclusion

We determined PyCs are widely distributed across plant food groups, and high concentrations occur in specific food types and not necessarily specific food groups. PyC_2-Gly was the predominant PyC in 18 of 20 food types, indicating focused study of PyC_2-Gly may be warranted in future studies examining health effects. Although more extensive surveys of foods are required, our data provide a foundation for estimating PyC concentrations consumed in the population and examining these biologically relevant concentrations in model systems for their potential health effects. Additional research is needed to evaluate the extent to which PyC content contributes to health benefits of plant-based diets.

Acknowledgements

This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases [T32 DK007734]; National Institute of Environmental Health Sciences [F31 ES030980, R01 ES023485, U2C ES030163]; and National Institutes of Health, Office of the Director [S10 OD018006].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Chapter 3 References

1. Kahleova H, Levin S, Barnard N. Cardio-Metabolic Benefits of Plant-Based Diets. Nutrients. 2017;9(8). doi: 10.3390/nu9080848. PubMed PMID: 28792455; PMCID: PMC5579641.

2. WCRF. Food, Nutrition, and Physical Activity, and the Prevention of Cancer: A Global Perspective. Washington, DC: AICR: World Cancer Research Fund/American Institute for Cancer Research, 2007.

3. Liu RH. Dietary bioactive compounds and their health implications. J Food Sci. 2013;78 Suppl 1:A18-25. Epub 2013/06/29. doi: 10.1111/1750-3841.12101. PubMed PMID: 23789932.

4. Seregin IV, Kozhevnikova AD. Low-molecular-weight ligands in plants: role in metal homeostasis and hyperaccumulation. Photosynth Res. 2020. Epub 2020/07/13. doi: 10.1007/s11120-020-00768-1. PubMed PMID: 32653983.

5. Pal R, Rai JP. Phytochelatins: peptides involved in heavy metal detoxification. Appl Biochem Biotechnol. 2010;160(3):945-63. doi: 10.1007/s12010-009-8565-4. PubMed PMID: 19224399.

6. Mou RX, Cao ZY, Lin XY, Wu L, Cao ZZ, Zhu ZW, Chen MX. Characterization of the phytochelatins and their derivatives in rice exposed to cadmium based on high-performance liquid chromatography coupled with data-dependent hybrid linear ion trap orbitrap mass spectrometry. Rapid Commun Mass Spectrom. 2016;30(16):1891-900. doi: 10.1002/rcm.7669. PubMed PMID: 27426698.

7. Dennis KK, Uppal K, Liu KH, Ma C, Liang B, Go YM, Jones DP. Phytochelatin database: a resource for phytochelatin complexes of nutritional and environmental metals. Database (Oxford). 2019;2019. Epub 2019/07/04. doi: 10.1093/database/baz083. PubMed PMID: 31267134; PMCID: PMC6606759.

8. Mehra RK, Kodati VR, Abdullah R. Chain length-dependent Pb(II)-coordination in phytochelatins. Biochem Biophys Res Commun. 1995;215(2):730-6. doi: 10.1006/bbrc.1995.2524. PubMed PMID: 7488015.

9. Satarug S, Vesey DA, Gobe GC. Health Risk Assessment of Dietary Cadmium Intake: Do Current Guidelines Indicate How Much is Safe? Environ Health Perspect. 2017;125(3):284-8. doi: 10.1289/EHP108. PubMed PMID: 28248635; PMCID: PMC5332171.

10. Fujita Y, el Belbasi HI, Min KS, Onosaka S, Okada Y, Matsumoto Y, Mutoh N, Tanaka K. Fate of cadmium bound to phytochelatin in rats. Res Commun Chem Pathol Pharmacol. 1993;82(3):357-65. PubMed PMID: 8122036.

11. Jumarie C, Fortin C, Houde M, Campbell PG, Denizeau F. Cadmium uptake by Caco-2 cells: effects of Cd complexation by chloride, glutathione, and phytochelatins. Toxicol Appl Pharmacol. 2001;170(1):29-38. doi: 10.1006/taap.2000.9075. PubMed PMID: 11141353.

12. Bardarov K, Naydenov M, Djingova R. HPLC-HRMS method for fast phytochelatins determination in plants. Application to analysis of Clinopodium vulgare L. Talanta. 2015;142:20-7. Epub 2015/05/25. doi: 10.1016/j.talanta.2015.04.014. PubMed PMID: 26003687.

13. Mendoza-Cozatl DG, Butko E, Springer F, Torpey JW, Komives EA, Kehr J, Schroeder JI. Identification of high levels of phytochelatins, glutathione and cadmium in the phloem sap of Brassica napus. A role for thiol-peptides in the long-distance transport of cadmium and the effect of cadmium on iron translocation. Plant J. 2008;54(2):249-59. doi: 10.1111/j.1365-313X.2008.03410.x. PubMed PMID: WOS:000254792700006.

14. Gong JM, Lee DA, Schroeder JI. Long-distance root-to-shoot transport of phytochelatins and cadmium in Arabidopsis. Proc Natl Acad Sci U S A. 2003;100(17):10118-23. Epub 2003/08/12. doi: 10.1073/pnas.1734072100. PubMed PMID: 12909714; PMCID: PMC187785.

15. Oven M, Page JE, Zenk MH, Kutchan TM. Molecular characterization of the homo-phytochelatin synthase of soybean Glycine max: relation to phytochelatin synthase. J Biol Chem. 2002;277(7):4747-54. doi: 10.1074/jbc.M108254200. PubMed PMID: 11706029.

16. Batista BL, Nigar M, Mestrot A, Rocha BA, Barbosa Junior F, Price AH, Raab A, Feldmann J. Identification and quantification of phytochelatins in roots of rice to long-term exposure: evidence of individual role on arsenic accumulation and translocation. J Exp Bot. 2014;65(6):1467-79. Epub 2014/03/07. doi: 10.1093/jxb/eru018. PubMed PMID: 24600019; PMCID: PMC3967088.

17. Dazy M, Beraud E, Cotelle S, Grevilliot F, Ferard JF, Masfaraud JF. Changes in plant communities along soil pollution gradients: responses of leaf antioxidant enzyme activities and phytochelatin contents. Chemosphere. 2009;77(3):376-83. Epub 2009/08/21. doi: 10.1016/j.chemosphere.2009.07.021. PubMed PMID: 19692108.

18. Palermo M, Pellegrini N, Fogliano V. The effect of cooking on the phytochemical content of vegetables. J Sci Food Agric. 2014;94(6):1057-70. Epub 2013/11/15. doi: 10.1002/jsfa.6478. PubMed PMID: 24227349.

19. USDA, ERS. Food Consumption and Nutrient Intakes 2014 [updated October 20, 2016; cited 2016 October]. Available from: https://www.ers.usda.gov/data-products/food-consumption-and-nutrient-intakes/.

20. Bowman SA, Clemens JC, Friday JE, Thoerig RC, Moshfegh AJ. Food Patterns Equivalents Database 2005-06: Methodology and User Guide2014. Available from: http://www.ars.usda.gov/ba/bhnrc/fsrg.

21. Seregin IV, Vooijs R, Kozhevnikova AD, Ivanov VB, Schat H. Effects of cadmium and lead on phytochelatin accumulation in maize shoots and different root parts. Dokl Biol Sci. 2007;415:304-6. Epub 2007/10/13. doi: 10.1134/s0012496607040163. PubMed PMID: 17929673.

22. Clemens S, Ma JF. Toxic Heavy Metal and Metalloid Accumulation in Crop Plants and Foods. Annu Rev Plant Biol. 2016;67:489-512. Epub 2016/04/30. doi: 10.1146/annurev-arplant-043015-112301. PubMed PMID: 27128467.

23. Piechalak A, Tomaszewska B, Baralkiewicz D, Malecka A. Accumulation and detoxification of lead ions in legumes. Phytochemistry. 2002;60(2):153-62. Epub 2002/05/16. doi: 10.1016/s0031-9422(02)00067-5. PubMed PMID: 12009318.

24. Vazquez S, Goldsbrough P, Carpena RO. Comparative analysis of the contribution of phytochelatins to cadmium and arsenic tolerance in soybean and white lupin. Plant Physiol Biochem. 2009;47(1):63-7. Epub 2008/11/14. doi: 10.1016/j.plaphy.2008.09.010. PubMed PMID: 19006673.

25. Yazdi M, Kolahi M, Mohajel Kazemi E, Goldson Barnaby A. Study of the contamination rate and change in growth features of lettuce (Lactuca sativa Linn.) in response to cadmium and a survey of its phytochelatin synthase gene. Ecotoxicol Environ Saf. 2019;180:295-308. Epub 2019/05/18. doi: 10.1016/j.ecoenv.2019.04.071. PubMed PMID: 31100594.

26. López-Climent MF, Arbona V, Pérez-Clemente RM, Zandalinas SI, Gómez-Cadenas A. Effect of cadmium and calcium treatments on phytochelatin and glutathione levels in citrus plants. Plant Biol (Stuttg). 2014;16(1):79-87. Epub 2013/04/12. doi: 10.1111/plb.12006. PubMed PMID: 23574491.

27. Hasan MK, Ahammed GJ, Yin L, Shi K, Xia X, Zhou Y, Yu J, Zhou J. Melatonin mitigates cadmium phytotoxicity through modulation of phytochelatins biosynthesis, vacuolar sequestration, and antioxidant potential in Solanum lycopersicum L. Front Plant Sci. 2015;6:601. Epub 2015/09/01. doi: 10.3389/fpls.2015.00601. PubMed PMID: 26322055; PMCID: PMC4531246.

28. Ben Ammar W, Mediouni C, Tray B, Ghorbel MH, Jemal F. Glutathione and phytochelatin contents in tomato plants exposed to cadmium. Biologia Plantarum. 2008;52(2):314. doi: 10.1007/s10535-008-0065-9.

29. Maier EA, Matthews RD, McDowell JA, Walden RR, Ahner BA. Environmental cadmium levels increase phytochelatin and glutathione in lettuce grown in a chelator-buffered nutrient solution. J Environ Qual. 2003;32(4):1356-64. Epub 2003/08/23. doi: 10.2134/jeq2003.1356. PubMed PMID: 12931891.

30. Meng Y, Zhang L, Wang L, Zhou C, Shangguan Y, Yang Y. Antioxidative enzymes activity and thiol metabolism in three leafy vegetables under Cd stress. Ecotoxicol Environ Saf. 2019;173:214-24. Epub 2019/02/18. doi: 10.1016/j.ecoenv.2019.02.026. PubMed PMID: 30772711.

31. Jones DP, Coates RJ, Flagg EW, Eley JW, Block G, Greenberg RS, Gunter EW, Jackson B. Glutathione in foods listed in the National Cancer Institute's Health Habits and History Food Frequency Questionnaire. Nutr Cancer. 1992;17(1):57-75. doi: 10.1080/01635589209514173. PubMed PMID: 1574445.

32. Ng ZX, Rosman NF. In vitro digestion and domestic cooking improved the total antioxidant activity and carbohydrate-digestive enzymes inhibitory potential of selected edible mushrooms. J Food Sci Technol. 2019;56(2):865-77. Epub 2019/03/25. doi: 10.1007/s13197-018-3547-6. PubMed PMID: 30906044; PMCID: PMC6400758.

33. He M, Openo K, McCullough M, Jones DP. Total equivalent of reactive chemicals in 142 human food items is highly variable within and between major food groups. J Nutr. 2004;134(5):1114-9. doi: 10.1093/jn/134.5.1114. PubMed PMID: 15113955.

34. Nayak B, Liu RH, Tang J. Effect of processing on phenolic antioxidants of fruits, vegetables, and grains--a review. Crit Rev Food Sci Nutr. 2015;55(7):887-919. Epub 2014/06/11. doi: 10.1080/10408398.2011.654142. PubMed PMID: 24915381.

35. ATSDR, Faroon O, Ashizawa A, Wright S, Tucker P, Jenkins K, Ingerman L, Rudisill C. Toxicological Profile for Cadmium. In: ATSDR, editor. Atlanta, GA 2012.

36. Boyer J, Liu RH. Apple phytochemicals and their health benefits. Nutr J. 2004;3:5. Epub 2004/05/14. doi: 10.1186/1475-2891-3-5. PubMed PMID: 15140261; PMCID: PMC442131.

37. Sullivan S, Alpers D, Klein S. Nutritional Physiology of the Alimentary Tract. In: Ross AC, Caballero B, Cousins RJ, editors. Modern Nutrition in Health and Disease. 11th ed. Philadelphia: Lippincott Williams & Wilkins; 2012. p. 540-73.

38. Layrisse M, Martinez-Torres C, Leets I, Taylor P, Ramirez J. Effect of histidine, cysteine, glutathione or beef on iron absorption in humans. J Nutr. 1984;114(1):217-23. Epub 1984/01/01. doi: 10.1093/jn/114.1.217. PubMed PMID: 6693981.

39. Silva EO, Bracarense AP. Phytic Acid: From Antinutritional to Multiple Protection Factor of Organic Systems. J Food Sci. 2016;81(6):R1357-62. Epub 2016/06/09. doi: 10.1111/1750-3841.13320. PubMed PMID: 27272247.

Table 3.1. Plant food groups and food types analyzed per group

Food group	Whole grains	Fruits	Root vegetables	Leafy greens	Legumes
	corn	orange	carrot	kale	peanuts
	brown rice	apple	potato	romaine lettuce	beans
	whole wheat flour	banana	onion	head lettuce	tofu
	oats	tomato	sweet potato	spinach	peas

Table 3.2. PyC_2-Gly content per typical serving.

Food	PyC_2-Gly (µg/g EW)[a]	Serving in grams[b]	PyC_2-Gly (µg/serving)
Fruits & Vegetables			
corn (raw)	13.89	150	2080
carrot (raw)	8.48	125	1060
orange (raw)	5.09	185	942
tomato (canned)	2.38	245	583
corn (frozen)	2.59	165	437
corn (canned)	2.52	165	416
onion (raw)	2.53	160	405
apple (raw)	3.23	110	355
carrot (canned)	1.92	145	278
peas (raw)	1.89	145	274
romaine (raw)	2.50	95	238
tomato (raw)	1.36	170	231
peas (frozen)	1.15	145	167
potato (baked)	1.33	120	160
head lettuce (raw)	1.00	110	110
spinach (canned)	0.60	170	102
peas (canned)	0.46	160	73.6
potato (canned)	0.47	155	72.9
kale (raw)	0.90	70	63.0
banana (raw)	0.40	150	60.0
sweet potato (baked)	0.26	200	52.2
spinach (raw)	0.70	70	49.0
Whole Grains			
brown rice (cooked)	0.29	98	28.6
oats	0.01	28	0.32
corn (cornmeal)	0.01	28	0.24
whole wheat flour	0.00	28	0.07
Protein			
peanuts	0.04	14	0.50
tofu	0.00	31	0.00
beans (canned)	0.00	44	0.00

[a]Values listed are mean edible weights (EW).

[b]Servings are listed as Food Pyramid Equivalents Database (FPED) equivalents based on the USDA dietary guidelines for Americans. Fruits and vegetables are listed as gram weights for 1 cup equivalents. Proteins and whole grains are listed as gram weights for 1 ounce equivalents.

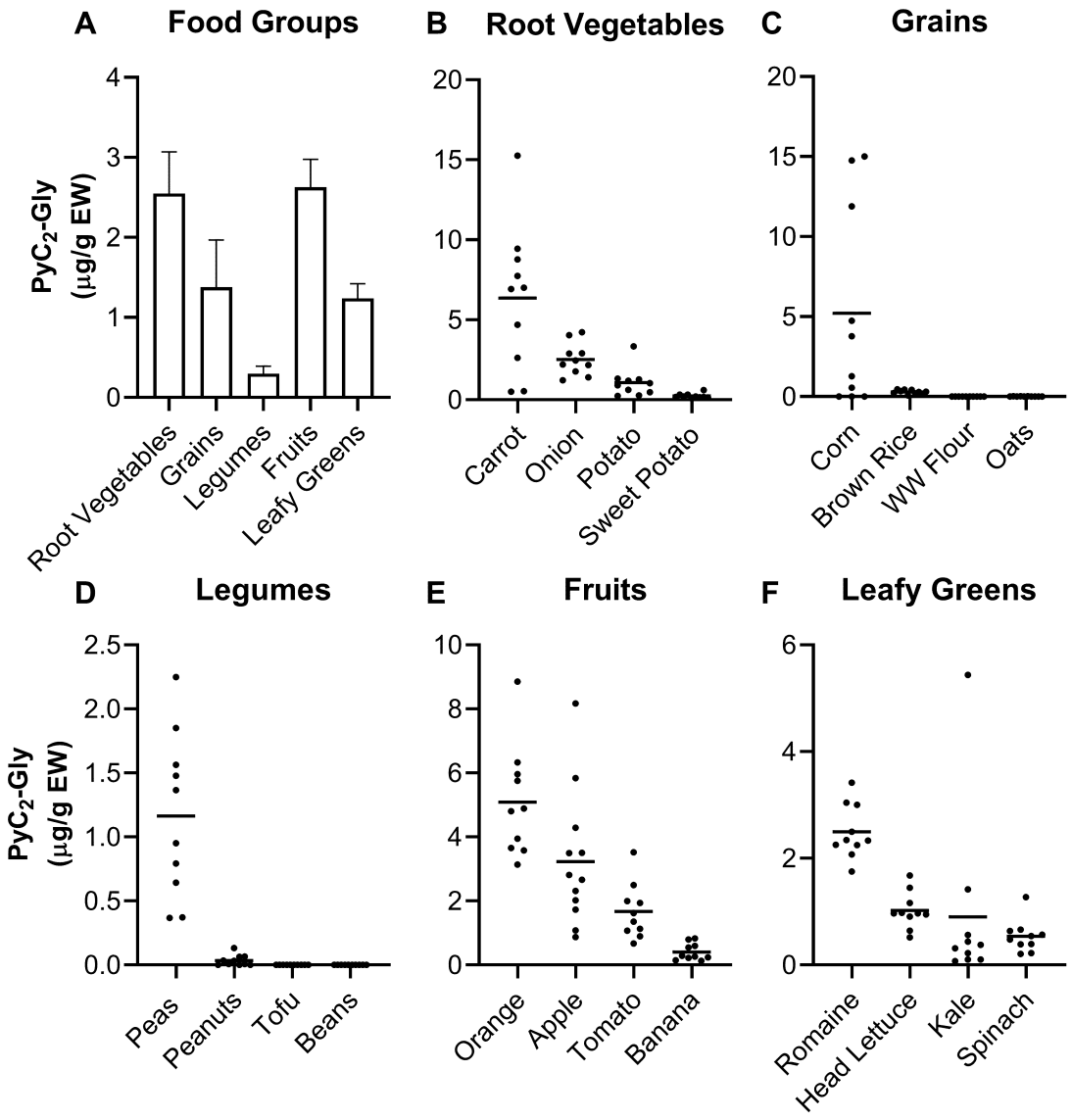

Figure 3.1. PyC_2-Gly concentration of 20 food types across five food groups. A, Bar graphs (mean ± SEM) show differences in PyC_2-Gly concentrations across food groups ($p < 0.05$, one-way ANOVA). B-F, Scatter plots of each food group show variability in PyC_2-Gly concentrations within food groups ($p < 0.05$, one-way ANOVA) and some foods. Horizontal lines indicate mean values. Points represent individual PyC_2-Gly levels in each food sample which are the average of three technical replicates. Carrots (B), corn (C), oranges (E), and apples (E) have the highest concentrations across all 20 food types. Tofu (D) and beans (D) contain no detectable PyC_2-Gly. EW, edible weight.

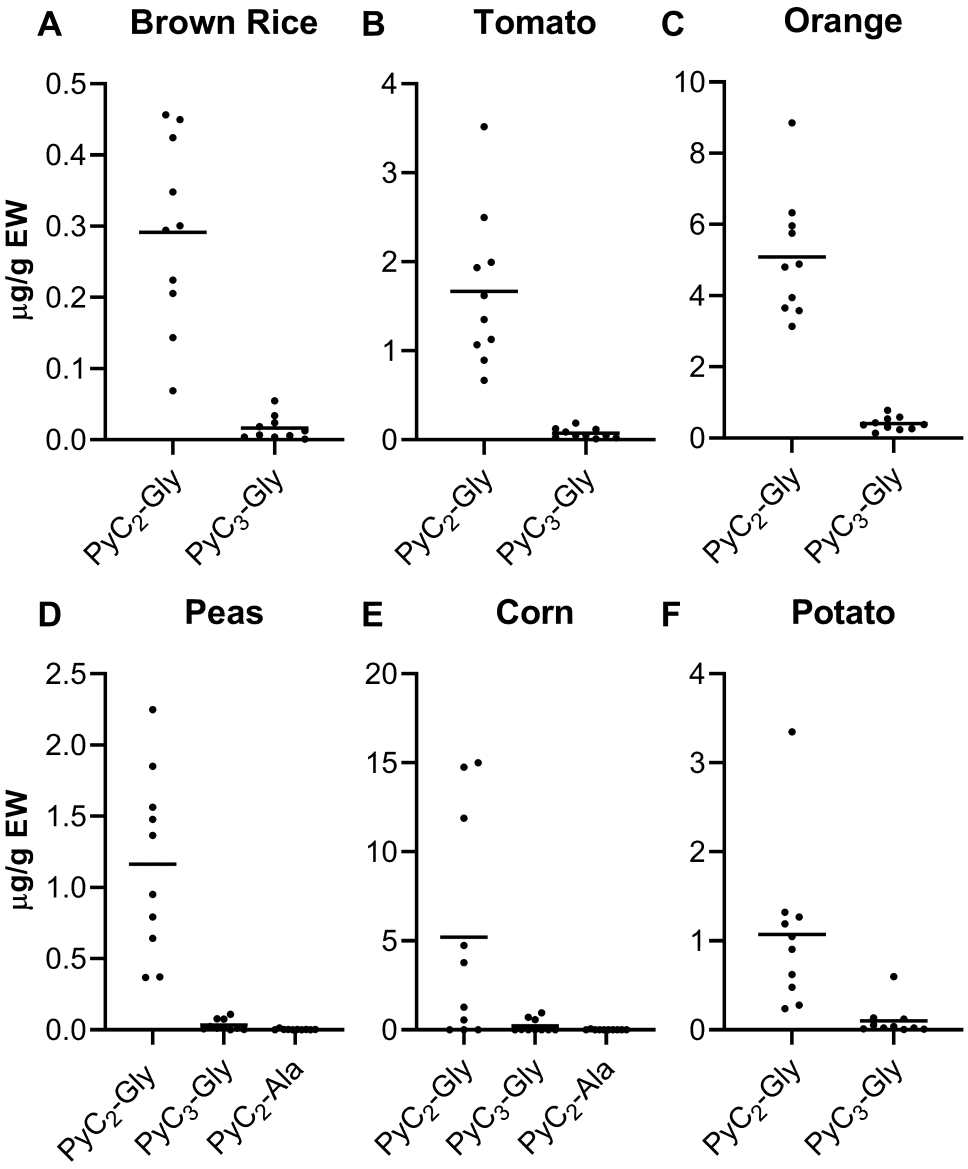

Figure 3.2. PyC_2-Gly concentrations were higher in all foods with multiple quantifiable PyCs (A-F). PyC_3-Gly and PyC_2-Ala concentrations were quantifiable in only this subset of analyzed foods. Points represent individual PyC levels in each food sample and are the average of three technical replicates. Horizontal lines represent the mean value of each PyC in the respective food. EW, edible weight.

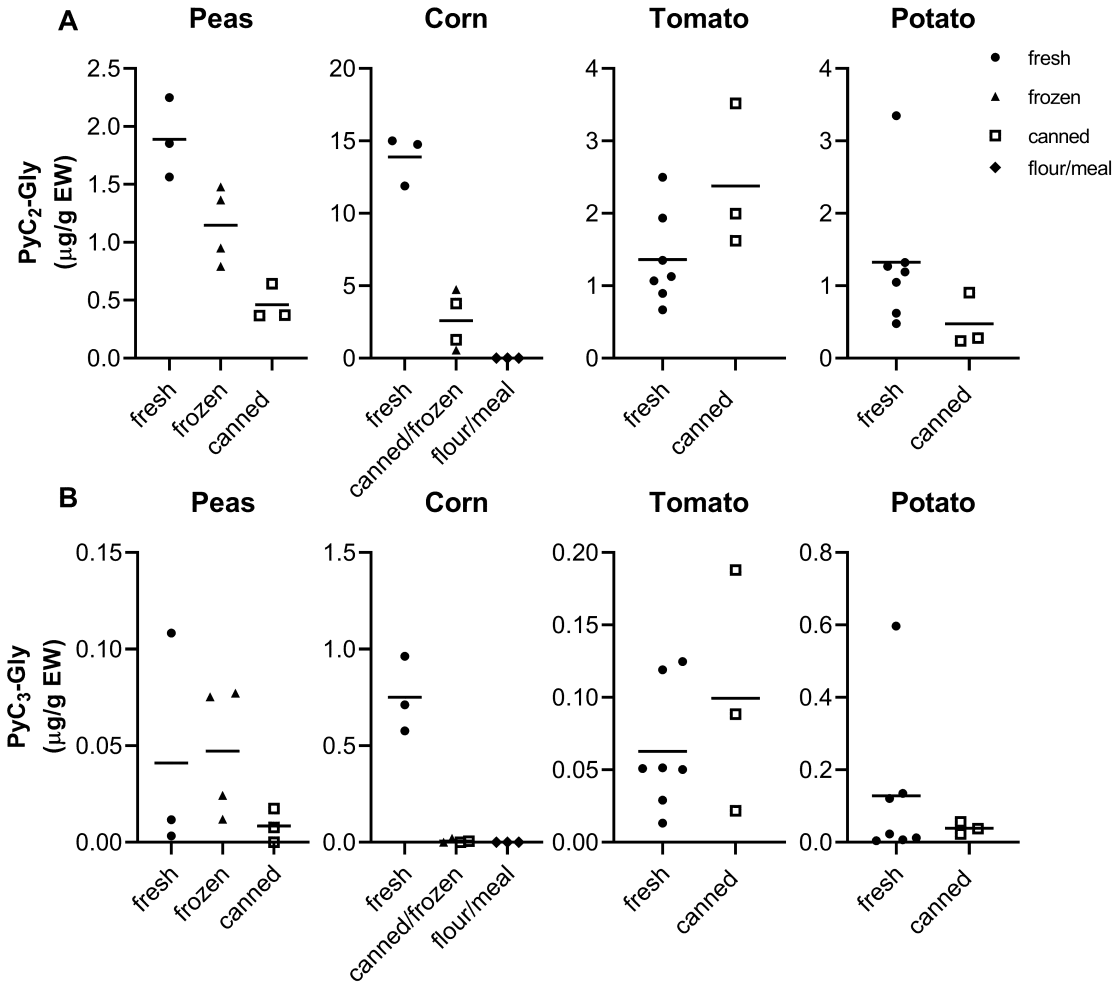

Figure 3.3. Food processing reduces PyC_2-Gly and PyC_3-Gly content in some foods. A, PyC_2-Gly concentrations were higher in corn and pea food samples with less food processing. No difference was detected in concentrations in potato and tomato samples with different processing levels. B, PyC_3-Gly concentrations were higher in corn samples with less food processing and there was with no difference in peas, potatoes, and tomatoes. Points indicate PyC concentrations of individual food samples and are the average of three technical replicates. Horizontal lines indicate mean values. EW, edible weight.

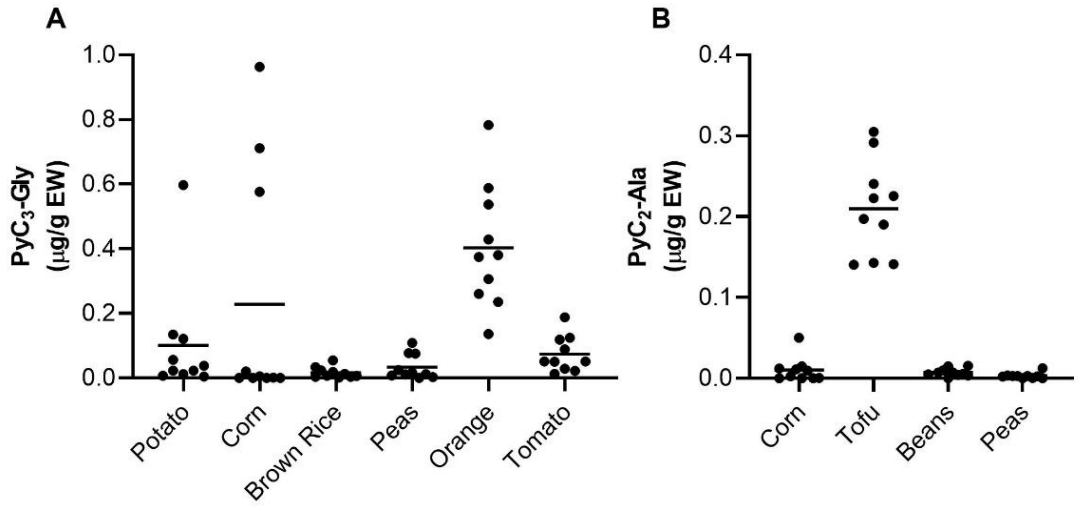

Figure S3.1 (Supplementary). PyC_3-Gly and PyC_2-Ala concentrations detected in a subset of foods. A, Oranges had the highest PyC_3-Gly concentrations of all foods (potatoes, brown rice, peas, tomatoes, Tukey's, $p < 0.05$; corn, Tukey's, NS). B, Tofu had the highest PyC_2-Ala concentrations of all foods (Tukey's, $p < 0.05$). Points indicate PyC concentrations of individual food samples and are the average of three technical replicates. Horizontal lines indicate mean values. EW, edible weight.

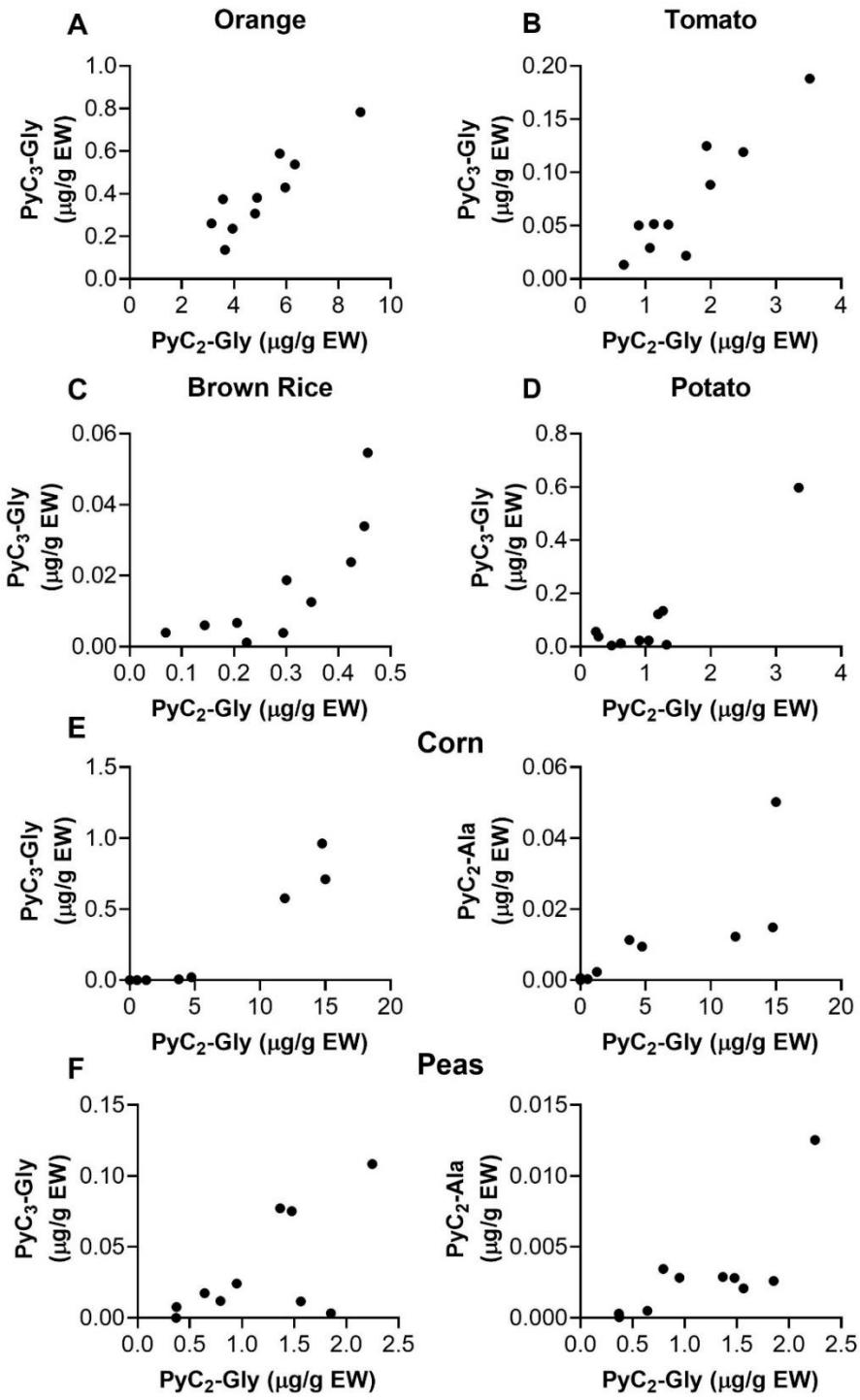

Figure S3.2 (Supplementary). PyC concentrations strongly correlate in individual food items (A-E, $p < 0.05$). Points indicate PyC concentrations of individual food samples and are the average of three technical replicates. EW, edible weight.

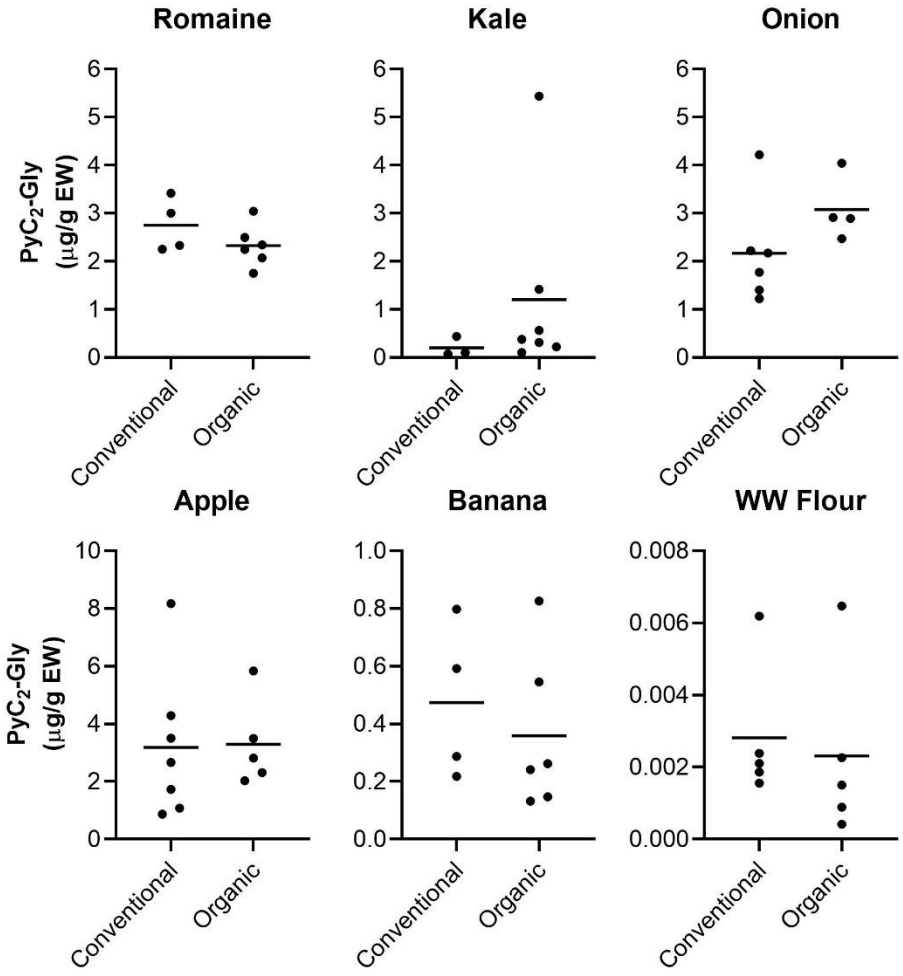

Figure S3.3 (Supplementary). Conventionally and organically grown foods did not differ by PyC_2-Gly content. Points indicate PyC concentrations of individual food samples and are the average of three technical replicates. Horizontal lines indicate mean values. EW, edible weight; WW, whole wheat.

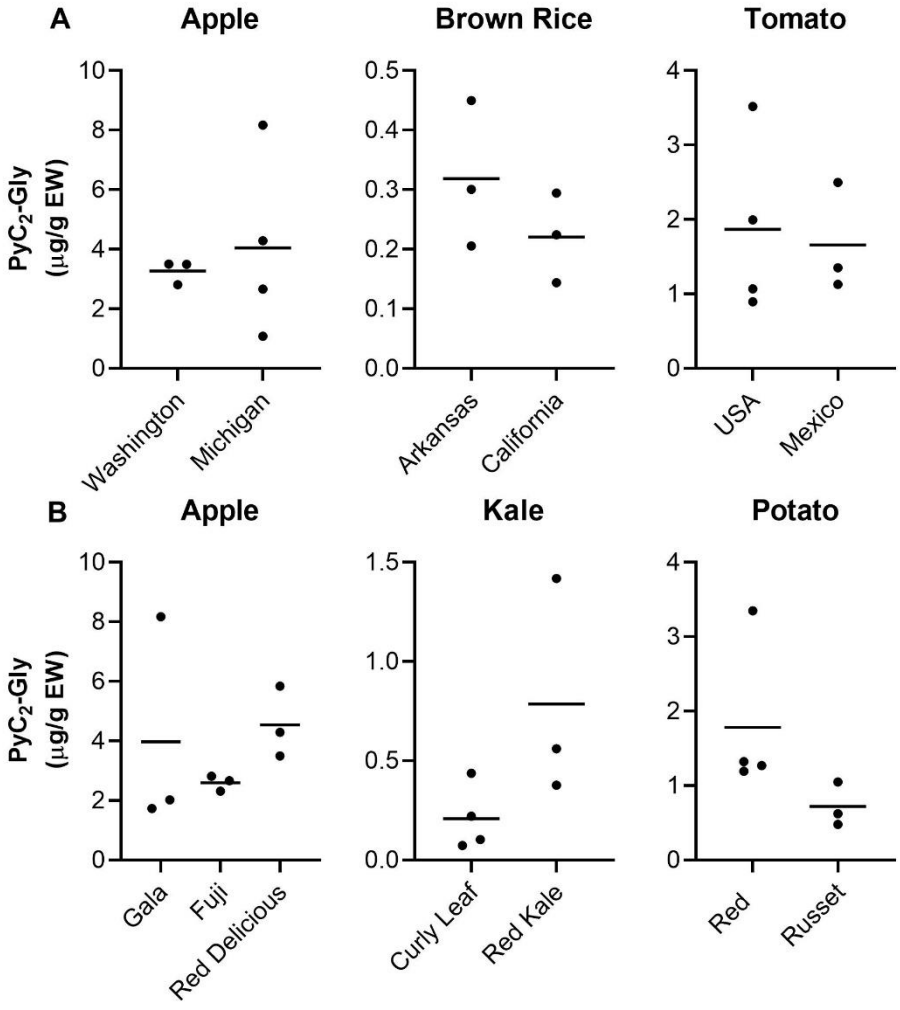

Figure S3.4 (Supplementary). PyC_2-Gly concentration did not differ by growing location (A) or variety (B) for food samples with this information. Horizontal lines indicate mean values. EW, edible weight.

Table S3.1 (Supplementary). Food types with multiple processing levels.

Food type	Processing levels included
corn	fresh, canned, flour/meal
tomato	fresh, canned
pea	fresh, frozen, canned
carrot	fresh, canned
potato	fresh, canned
spinach	fresh, canned

Table S3.2 (Supplementary). Phytochelatin *m/z* targets.

Phytochelatin[a]	*m/z*	Adduct
(S-S)PyC_2-Ala	552.1429	M+H
PyC_2-Ala	554.1585	M+H
(S-S)PyC_2-Glu	610.1483	M+H
PyC_2-Glu	612.1640	M+H
(S-S)PyC_2-Gly	538.1272	M+H
PyC_2-Gly	540.1429	M+H
(S-S)PyC_3-Gly	770.1790	M+H
PyC_3-Gly	772.1946	M+H
$(S\text{-}S)_2PyC_4$-Gly	1000.2151	M+H
(S-S)PyC_4-Gly	1002.2308	M+H
PyC_4-Gly	1004.2464	M+H
$(S\text{-}S)_2PyC_5$-Gly	1232.2669	M+H
(S-S)PyC_5-Gly	1234.2825	M+H
PyC_5-Gly	1236.2982	M+H
$(S\text{-}S)_3PyC_6$-Gly	731.6551	M+2H
$(S\text{-}S)_2PyC_6$-Gly	732.6630	M+2H
(S-S)PyC_6-Gly	733.6708	M+2H
PyC_6-Gly	734.6786	M+2H

[a]Oxidized forms of each phytochelatin represented by the number of disulfide bonds (S-S).

Table S3.3 (Supplementary). MSMS confirmation of detectable but not quantifiable phytochelatins.

Phytochelatin	Food type
PyC_3-Gly	carrot
PyC_4-Gly	potato onion carrot banana orange
PyC_2-Ala	head lettuce brown rice apple orange
PyC_2-Glu	head lettuce kale carrot

CHAPTER 4

Transport and function of a common dietary phytochelatin, PyC_2-Gly, in intestinal epithelial cells

Kristine K. Dennis[a], Michael Orr[b], Karan Uppal[b], Miriam B. Vos[c], Young-Mi Go[b], Dean P. Jones[a, b]

[a]Nutrition and Health Sciences, Laney Graduate School, Emory University, Atlanta, GA, USA

[b]Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, Emory University, Atlanta, GA, USA

[c]Department of Pediatrics, School of Medicine, Emory University, Atlanta, GA, USA

Abstract

Dietary phytochelatins (PyCs) may contribute to the health benefits of plant-based diets due to their metal-binding properties and similarities to glutathione, a central redox molecule. PyC_2-Gly is one of the most common PyCs in the diet but has not been studied in mammalian systems. Here, we examined whether PyC_2-Gly is absorbed into and through intestinal epithelial cells and if it reduces absorption of the toxic metal, cadmium (Cd). Caco-2 cells were grown into confluent monolayers and treated with PyC_2-Gly (5 µM and 10 µM) and Cd (1 µM and 5 µM) in combination and alone. Cells were treated for 2 h, 4 h, and 6 h. Cells and basolateral media were collected and analyzed with liquid chromatography-mass spectrometry and inductively-coupled plasma mass spectrometry for PyC_2-Gly and Cd content, respectively. PyC_2-Gly was absorbed into and through cells in a dose-dependent manner. Mean cellular PyC_2-Gly uptakes at 4 h were 29 µg PyC_2-Gly/µg protein and 53 µg PyC_2-Gly/µg protein for 5 µM and 10 µM, respectively. Basolateral media of PyC_2-Gly treated cells at 4 h contained 0.028 µM PyC_2-Gly and 0.049 µM PyC_2-Gly for 5 µM and 10 µM doses, respectively. Although underpowered to draw definitive conclusions, cells co-treated with PyC_2-Gly and Cd had lower Cd content than cells treated with Cd alone. Our data demonstrate one of the most common dietary PyCs, PyC_2-Gly, is absorbed into and through intestinal epithelial cells. Additional research is required to determine whether PyCs reduce dietary Cd absorption or influence other biological processes such as redox systems or metal homeostasis.

Introduction

Consuming a variety of plant-derived foods is recognized to provide health benefits due to a range of nutritive and non-nutritive components (1-3). A large spectrum of naturally occurring chemicals occurs in plants, and elucidating health benefits is complicated because some chemicals not only influence absorption of other dietary components, but also can participate positively or negatively in complex biological processes such as oxidative stress (4-7). For example, phytates bind to iron, forming insoluble complexes which reduce iron bioavailability (4). Polyphenols such as curcumin and resveratrol modulate antioxidant and anti-inflammatory responses to protect from oxidative stress (8). Because of the large number and diverse nature of phytochemicals, improved understanding of respective impacts of individual chemicals and chemical classes is needed to better understand benefits of plant-derived foods on health (7, 9).

Phytochelatins (PyCs) are metal-chelating phytochemicals found in the human diet. PyC types and concentrations in plant foods in the human diet were recently defined for the first time (10), and this information enables studies in model systems to better understand potential health benefits of PyCs. PyCs bind cadmium (Cd), a toxic metal regularly consumed in the diet, with high affinity and may reduce Cd bioavailability (11, 12). With the newly defined parameters for concentrations and distributions of dietary PyCs, there is an opportunity to expand our understanding of how specific PyCs impact absorption of low-level dietary Cd.

Cd is widespread in the environment, entering the food chain through absorption into plants and animals grown for food (13). Despite dietary Cd exposures in the U.S. being low relative to some occupational exposures and to cigarette smoking, average dietary Cd exposures are associated with increased risks of numerous diseases, including chronic kidney disease, type 2 diabetes, and cancer (14). Humans have poor ability to eliminate Cd, and the biologic half-life has been estimated to be between 10 and 30 years (15). In principle, reducing dietary Cd exposures will reduce Cd body burdens and related health risks (13). One approach to reduce exposure is decreasing Cd bioavailability through modifying other dietary factors (15). For example, high fiber diets are associated with reduced Cd levels in humans (16). Additionally, vegetarian diets can have higher dietary Cd intake from foods but not result in higher Cd biomarkers (17). It has been postulated that this relates to higher fiber intake in vegetarian diets (18) which positively correlates with metal-binding activity. Our research is based upon the premise that PyCs could contribute to reductions in dietary Cd absorption due to their metal-binding properties and widespread distribution in plant foods.

Evidence from intestinal cell studies have demonstrated PyCs can reduce Cd absorption. For instance, in a study in intestinal cells, PyC_3-Ser (0.3 µM) reduced Cd accumulation by half relative to unbound Cd ($CdCl_2$) (11). Another study of colon-like Caco-2 BBE cells assessed viability in Cd-exposed cells with and without PyC_3-Gly (7 µM) exposure. This study found PyC_3-Gly protected from Cd-induced cell death (19). Although cellular Cd absorption was not measured, protection was likely due to a combination of reduced Cd absorption, PyC_3-Gly sequestration of Cd in a bound complex within the cell, or PyC_3-Gly protection from Cd-induced cellular stress. In the same study, fluorescently-tagged PyC_3-Gly was absorbed into cells and

through the confluent monolayer, suggesting PyCs could have additional functions beyond reducing Cd absorption (19).

PyCs are a diverse class of glutathione-derived molecules, consisting of 2-11 repeating y-Glu-Cys peptide repeats followed by a terminal amino acid (20, 21). The terminal amino acid (Gly, Glu, Ser, β-Ala, Ala, Gln, or no amino acid) determines the PyC type, and the length determines the capacity and affinity for metal-binding (21, 22). In mammalian cells, the type and length will also determine the degree to which the PyC is absorbed (23, 24). In intestinal cells, PyC_3-Gly absorption was partially attributed to receptor-mediated transcytosis with the majority of absorption occurring via an undefined mechanism (19). Transport of oligopeptides containing five or more amino acids can occur via sodium-coupled oligopeptide transporters 1 and 2 (24). In a recent study of commonly consumed foods, PyC_2-Gly was the most common and abundant PyC (10). Of twenty commonly consumed plant-derived foods, PyC_3-Gly was found in one-third of the foods, and PyC_2-Ala was found only at much lower concentrations, much below the concentration ranges tested in the previous studies of Cd absorption (10). Importantly, PyC_2-Gly has not been previously studied in mammalian systems.

In this study, we examine PyC_2-Gly absorption and the impact of PyC_2-Gly on Cd absorption at relevant dietary concentrations in an intestinal epithelial cell model. First, we measure PyC_2-Gly uptake into and through cells by increasing time and dose. Second, we characterize Cd uptake into and through cells at low-level Cd doses experienced in the average U.S. diet. Finally, we examine how cells dosed with combinations of dietary PyC_2-Gly and Cd concentrations absorb

PyC_2-Gly and Cd. This study provides characterization of an intestinal cell model system for studying dietary PyC_2-Gly and its interaction with low-level dietary Cd exposures.

Materials and Methods

Cell culture: Caco-2 cells were obtained from ATCC and maintained under 5% CO_2 conditions at 37ºC in cell culture media (Sigma Aldrich Minimum Essential Media, M4655 formulation) with 20% fetal bovine serum (FBS), 1% sodium pyruvate (Hyclone) and 1% penicillin and streptomycin. For all experimental conditions, serum was reduced to 0.5% FBS. All assays were conducted between cell passages 40 to 52. Cells were maintained in T-75 culture flasks and subcultured at 70-80% confluency. For permeability assessment and uptake experiments, cells were seeded in 24-well plates on polycarbonate filters with 0.33 cm^2 growth area, 0.4 µm pore diameter (Millicell) and 30,000 cells/well and grown for ≥14 days (max 21 days) until transepithelial electrical resistance (TEER) reached ≥300 $\Omega*cm^2$.

Cell dosing and collection: Before dosing, cells were washed twice on apical and basolateral sides with warmed 0.5% FBS media. Cells were treated with $CdCl_2$ at 1 µM and 5 µM and/or PyC_2-Gly at 5 µM and 10 µM concentrations for 2 h, 4 h, and 6 h in combination and alone. These concentrations were selected based upon the expected concentrations in the diet (10, 25). Additionally, previous research with low Cd doses demonstrated no negative effects on cell monolayer permeability in Caco-2 cells grown for 2 to 5 weeks with 1 µM and 5 µM Cd (26). Following treatment, cell culture plates were placed on ice. Treatment media was aspirated and cells were rinsed twice with cold PBS containing calcium and magnesium. Cells were collected

by scraping from polycarbonate filters in cell culture grade water. Basolateral media was collected from all wells. Cells and basolateral media samples were frozen at -80ºC until ICP-MS and LC-MS analyses.

Integrity of cell monolayer: Permeability of the cell monolayer after treatment was assessed using a FITC-Dextran assay. FITC-Dextran (4 kDa, Sigma-Aldrich) at a concentration of 1 mg/ml was added in conjunction with PyC_2-Gly and Cd doses. After 4 h treatment, samples of basolateral media were harvested and fluorescence was measured using a plate reader (Ex 485, Em 538). Results are presented as % of control. Barrier function was assessed with TEER measurements. TEER was measured after cells had been dosed for 4 h using a voltohmmeter (World Precision Instruments) where the measured resistance in Ohms was adjusted for the blank well resistance and multiplied by the area of the transwell filter (0.33 cm^2).

Cd concentration measurements: ^{114}Cd was measured in Caco-2 cells and basolateral media with inductively-coupled plasma mass spectrometry (ICP-MS, iCap Q, ThermoFisher Scientific). Cells were digested in 2% nitric acid to a final volume of 10 mL. A linear standard curve (0.25 to 60 parts per billion) was run at the same time as samples.

PyC_2-Gly concentration measurements: Cell and basolateral media samples were prepared with a 2:1 ratio of acetonitrile to experimental sample (i.e. cells in water or basolateral media). Prepared samples were vortexed, equilibrated on ice for 30 min, and centrifuged for 10 min (14,000 rpm) at 4ºC. Supernatant from cell and basolateral media samples were analyzed with LC-MS (10-min

method, *m/z* 500-1250) on a Thermo Fisher LTQ-Velos Orbitrap mass spectrometer at 60,000 resolution. Control samples with and without added PyC_2-Gly authentic standard (10 nM and 1 µM) were analyzed at the end of the analysis for estimating PyC_2-Gly concentrations. Each sample was analyzed in triplicate with a 10 µl injection on a HILIC column in positive electrospray ionization mode. Cell pellets were saved for later analysis with the Bradford assay for protein content. The intensity of PyC_2-Gly (*m/z* 538.1272; M+H adduct) was determined using the area under the curve (Thermo Scientific Xcalibur software, Genesis peak integration algorithm). Technical replicates were averaged. PyC_2-Gly quantification used a method of additions and estimates of PyC_2-Gly concentrations determined relative to the intensity of the added PyC_2-Gly standard. PyC_2-Gly concentrations were reported as µg PyC_2-Gly/µg protein for cell samples and µM PyC_2-Gly for basolateral media samples.

Statistical Analysis

Statistics were performed using the software GraphPad Prism version 8.3.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com. Data were analyzed using Student's t-test or one-way analysis of variance (ANOVA) with post-hoc Dunnett's multiple comparisons test or linear trend test. Data are summarized by box and whisker plots including median, minimum, and maximum. Summary statistics are reported as mean values in text. Statistically significant differences were defined as p-values < 0.05.

Results

Integrity of Caco-2 cell monolayer

To assess the integrity of the cell monolayer, we measured TEER and apical-to-basolateral movement of FITC-dextran for each dose and combination dose of PyC_2-Gly (0, 5, and 10 µM) and Cd (0, 1, and 5 µM). FITC-Dextran fluorescence did not differ from control for all experimental doses (**Figure 4.1A**). The positive control dosed with 25 µM of Cd for 4 h had 50% higher FITC-Dextran in the basolateral media than all other doses. Mean TEER was greater than 400 $\Omega*cm^2$ for all experimental doses at 4 h (**Figure 4.1B**). The positive control (25 µM Cd) had an average TEER lower than the monolayer integrity threshold of 300 $\Omega*cm^2$.

PyC_2-Gly uptake into and through Caco-2 cell monolayers

Analysis of cells after 4 h of treatment with two doses (5 and 10 µM) of PyC_2-Gly showed a dose-dependent increase of cellular PyC_2-Gly content (**Figure 4.2**). Cells dosed with 5 and 10 µM PyC_2-Gly had mean uptakes of 29 and 53 µg PyC_2-Gly/µg protein, respectively (Student's t-test, $p < 0.05$). Timecourse (2, 4, and 6 h) experiments of cells treated with 10 µM PyC_2-Gly showed the highest cellular PyC_2-Gly content at 2 h with decreasing content at 4 h and 6 h (**Figure 4.3A**). Basolateral PyC_2-Gly content from cell monolayers dosed with 10 µM PyC_2-Gly showed a time-dependent increase in PyC_2-Gly (mean PyC_2-Gly: 2 h, 0.047 µM; 4 h, 0.049 µM; 6 h, 0.074 µM) which was not significant by one-way ANOVA ($p > 0.05$). Cells dosed with 5 µM PyC_2-Gly over time had no difference in PyC_2-Gly cellular content (**Figure 4.3A**) or basolateral media content (**Figure 4.3B**) (one-way ANOVA; $p > 0.05$).

Cadmium uptake into and through Caco-2 cell monolayers

We examined Cd uptake into cells by Cd treatment (0, 1, and 5 µM) at 4 h. Cd content was increased in cells treated with 1 and 5 µM $CdCl_2$ compared to controls (**Figure 4.4A**; one-way ANOVA, $p < 0.05$). Mean cellular content at 4 h was 22 and 79 ng/mg protein for 1 and 5 µM $CdCl_2$ treatments, respectively. Cells treated with 5 µM $CdCl_2$ for 2, 4, and 6 h showed an increase in cellular Cd content by linear trend test (**Figure 4.4B**; $p < 0.05$). When assessing Cd transport through Caco-2 cell monolayers, basolateral media of the 5 µM $CdCl_2$ treated cells had a mean Cd level of 0.31 ppb and Cd was not detectable in the 1 µM $CdCl_2$ treated group (**Figure S4.1A**). Time point measurements of basolateral media showed no differences in Cd levels (**Figure S4.1B**; one-way ANOVA, $p > 0.05$).

Impact of PyC_2-Gly on Cd uptake in Caco-2 cell monolayers

Cd content was measured in cells and basolateral media of Caco-2 cell monolayers dosed for 4 h with combinations of PyC_2-Gly (0, 5, and 10 µM) and $CdCl_2$ (0, 1, and 5 µM). Although the data showed no significant differences, some trends were observed. In cells treated with 5 µM $CdCl_2$, there was a dose-dependent decrease in mean Cd content with increasing PyC_2-Gly dose (**Figure 4.5**; 74 and 59 ng ^{114}Cd /mg protein compared to control 79 ng ^{114}Cd /mg protein). In basolateral media of 5 µM $CdCl_2$ treated cells, average Cd content was lower in cells treated with 5 and 10 µM PyC_2-Gly (0.30 and 0.27 ppb) compared to no PyC_2-Gly (0.31 ppb), but this was not significant (Student's t-test, $p > 0.05$) (**Figure S4.2**). In cells treated with 1 µM $CdCl_2$, there was also a dose-dependent decrease in mean Cd content with increasing PyC_2-Gly dose (**Figure 4.5**; 19 and 14 ng ^{114}Cd/mg protein compared to control 22 ng ^{114}Cd/mg protein). In the basolateral media of 1 µM $CdCl_2$ treated cells, no Cd was detected (data not shown).

Impact of Cd on PyC_2-Gly uptake in Caco-2 cell monolayers

PyC_2-Gly content was also measured in cells co-treated with the same combinations of $CdCl_2$ and PyC_2-Gly listed above. In cells treated with 5 µM PyC_2-Gly, mean PyC_2-Gly cellular content was similar across doses of Cd (0, 1, and 5 µM) with a slight increase with higher Cd doses (29, 30, and 33 µg PyC_2-Gly/µg protein, respectively) (**Figure S4.3A**). Cells treated with 10 µM PyC_2-Gly had no difference in PyC_2-Gly content across Cd doses (53, 55, and 53 µg PyC_2-Gly/µg protein, respectively) (**Figure S4.3B**). No significant differences in cellular or basolateral media PyC_2-Gly content were observed in co-treated cells (one-way ANOVA, $p > 0.05$; **Figure S4.4**).

Discussion

The present study was to characterize the uptake of PyC_2-Gly and the impact of PyC_2-Gly on Cd absorption in an intestinal cell model system using PyC_2-Gly and Cd concentrations similar to those in the human diet. Previous research on PyCs identified PyC_3-Gly absorption into and through colon-like Caco-2 BBE cells and reduced Cd absorption in the presence of PyC_3-Ser (11, 19). The present study shows PyC_2-Gly, the most abundant PyC in the diet, can be taken up into and through Caco-2 cell monolayers which may have implications for metal homeostasis and cellular redox systems (27). We also observed dose-dependent decreases in Cd uptake with increasing PyC_2-Gly dose, supporting other literature suggesting consumption of dietary PyCs via plant foods may contribute to reduced Cd body burdens.

Low-level dietary Cd exposures are of human health concern due to the long half-life of Cd, leading to increased Cd body burden and associated increases in chronic disease risk (13, 17). In our study, we found a 25% reduction in cellular Cd content in cells co-treated with 10 µM PyC_2-Gly. However, our study was underpowered to test for protection against Cd absorption and will require repetition to confirm protection under usual dietary intakes. Another study examining the difference in Cd accumulation of inorganic Cd versus Cd-PyC_3-Ser complexes found a 50% reduction in Cd uptake in cells dosed with Cd-PyC_3-Ser complexes (11). Similarly, rats treated with Cd-PyC_3-Gly directly in the stomach had 50% lower Cd distribution to the liver and kidney relative to $CdCl_2$. In the same study, those treated with Cd-PyC_3-Gly directly in the duodenum had even lower Cd distribution to the liver and kidney relative to $CdCl_2$ (12). Their research suggests the protective effect of PyCs will depend on where the compounds become available during digestion (28).

As PyCs are distributed within plant matrices of foods, they may be liberated much later in digestion as the microbiome breaks down the indigestible plant matrices in the colon (29). In addition to research on PyCs, research has demonstrated that metal complexation with other organic ligands reduce intestinal Cd absorption (16). PyC length may also impact Cd absorption as longer PyCs have more thiol groups, the primary Cd binding site, due to their higher cysteine content. However, even though PyC_2-Gly is shorter in length from other PyCs studied regarding PyC and Cd absorption, the reduction in Cd absorption was similar. PyC_2 and PyC_3 may be similar enough for Cd binding and related absorption not to be impacted. Reduced absorption of dietary Cd at the levels observed in our study and others via consumption of PyC-containing

plant foods could have implications for population-level Cd body burdens and contribute to reduced Cd-related disease risks.

In our study, PyC_2-Gly was taken up into intestinal cells and through the cell monolayer into the basolateral media. Another study of fluorescently-tagged PyC_3-Gly demonstrated similar results with transfer into and through colon-like Caco-2 BBE cells (19). Similar to our study, Langelueddecke et al. also observed increased absorption with increasing PyC_3-Gly doses from 0 to 5.6 µM. However, whether PyC_2-Gly and PyC_3-Gly absorption occurs via similar mechanisms remains to be elucidated. Although the transport mechanisms of PyC_2-Gly could not be determined by our study design, previous research has demonstrated that peptide length and amino acid composition impacts absorption (23, 24). As PyC transport has not been quantified in other studies, we could not directly compare PyC_2-Gly and PyC_3-Gly uptake and transport efficiencies. Our data demonstrate about 0.5% of the apical PyC_2-Gly dose is transferred through the cell monolayer at both 5 µM and 10 µM treatments at 4 h. In the previous study of PyC_3-Gly, about 50% of PyC_3-Gly transferred out of the apical media reached the basolateral media after 8 h, suggesting about 50% remained in the cells (19). However, quantitative comparisons were not possible as the authors did not report the quantity of PyC_3-Gly initially transferred out of the apical media. Mechanisms for PyC_2-Gly uptake may differ from PyC_3-Gly and other PyCs as peptide transfer is impacted by length and terminal amino acids (23, 24). PyC_2-Gly is a pentapeptide which could be transported by oligopeptide transporters, endocytosis, or other mechanisms (30). Longer PyCs such as PyC_3-Gly may be more dependent on endocytosis mechanisms for uptake. In a study examining the impact of N-terminus and C-terminus amino acids on alanine-based pentapeptide uptake, EAAAA had intermediate permeability while

AAAAG had low permeability (23). Intestinal metals could impact absorption of PyC as PyC-metal complexes form, potentially preventing uptake of PyC into the cell. However, in cells co-treated with PyC_2-Gly and Cd, we did not observe a decrease. We even observed a slight increase in PyC_2-Gly uptake with increasing Cd doses in cells treated with 5 µM but not 10 µM PyC_2-Gly. Additional research to elucidate the factors influencing transport of PyC_2-Gly, the most abundant PyC in the human diet, will inform how other dietary factors influence PyC uptake.

With their high thiol content and similarity to GSH, PyCs may interact with redox systems through high affinity binding of metal cations to thiol groups (27, 31). The interplay with redox systems could occur within intestinal cells and systemically. At the cellular level, Cd can disrupt redox mechanisms despite being a redox-stable metal (32). Some of Cd's toxic effects are linked to its disruption of redox systems through depletion of GSH and aberrant binding to critical proteins via thiol groups (27, 32). PyCs bind to Cd with higher affinity than GSH which could protect the GSH pool from Cd (33). As GSH is a central endogenous antioxidant, depletion of the GSH pool by Cd can induce oxidative stress (34). Cd-induced oxidative stress can occur when Cd replaces iron in cellular proteins, increasing the quantity of free iron, a redox-active metal, within the cell (35). Other mechanisms of Cd-induced oxidative stress are linked to inhibition of selenoenzymes, disruptions in mitochondrial function, and overactivation of signaling pathways such as induction of NADPH oxidases (35). Systemically, the role of PyCs is unclear. Redox signaling is essential for normal physiological functioning and intricate control mechanisms exist to ensure reductive and oxidative metabolism can occur within cells at the same time (27). The redox network systemically allows the body to respond to the surrounding

environment (36, 37). Due to their structure, properties, and absorption, PyCs will interface with redox systems (27). How PyCs interact with redox processes intracellularly remains an intriguing open question for redox biology and nutrition.

To our knowledge, this is the first study to examine uptake of the common dietary PyC, PyC_2-Gly, and its impact on Cd absorption. Our results support previous research demonstrating PyCs can be absorbed intact into and through intestinal epithelial cells. Additionally, at concentrations of Cd and PyC_2-Gly common in the U.S. diet, PyC_2-Gly reduces Cd absorption by 25%. Thus, our data support that PyCs found in plant foods could contribute to reduced Cd absorption and over the lifespan, reduced overall Cd body burden. Due to the potential implications of absorbed PyCs, future investigations into the biological impact of PyCs on metal homeostasis and cellular redox systems are warranted.

Acknowledgements

This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases [T32 DK007734]; National Institute of Environmental Health Sciences [F31 ES030980, R01 ES023485, U2C ES030163]; and National Institutes of Health, Office of the Director [S10 OD018006]. We thank Raven Peterson and Jason Matthews for their technical assistance.

Chapter 4 References

1. Kahleova H, Levin S, Barnard N. Cardio-Metabolic Benefits of Plant-Based Diets. Nutrients. 2017;9(8). doi: 10.3390/nu9080848. PubMed PMID: 28792455; PMCID: PMC5579641.

2. Le LT, Sabate J. Beyond meatless, the health effects of vegan diets: findings from the Adventist cohorts. Nutrients. 2014;6(6):2131-47. doi: 10.3390/nu6062131. PubMed PMID: 24871675; PMCID: PMC4073139.

3. Slavin JL, Lloyd B. Health benefits of fruits and vegetables. Adv Nutr. 2012;3(4):506-16. Epub 2012/07/17. doi: 10.3945/an.112.002154. PubMed PMID: 22797986; PMCID: PMC3649719.

4. Schlemmer U, Frolich W, Prieto RM, Grases F. Phytate in foods and significance for humans: food sources, intake, processing, bioavailability, protective role and analysis. Mol Nutr Food Res. 2009;53 Suppl 2:S330-75. Epub 2009/09/24. doi: 10.1002/mnfr.200900099. PubMed PMID: 19774556.

5. Anjum NA, Hasanuzzaman M, Hossain MA, Thangavel P, Roychoudhury A, Gill SS, Rodrigo MA, Adam V, Fujita M, Kizek R, Duarte AC, Pereira E, Ahmad I. Jacks of metal/metalloid chelation trade in plants-an overview. Front Plant Sci. 2015;6:192. Epub 2015/04/18. doi: 10.3389/fpls.2015.00192. PubMed PMID: 25883598; PMCID: PMC4382971.

6. Olszowy M. What is responsible for antioxidant properties of polyphenolic compounds from plants? Plant Physiol Biochem. 2019;144:135-43. Epub 2019/09/30. doi: 10.1016/j.plaphy.2019.09.039. PubMed PMID: 31563754.

7. Liu RH. Dietary bioactive compounds and their health implications. J Food Sci. 2013;78 Suppl 1:A18-25. Epub 2013/06/29. doi: 10.1111/1750-3841.12101. PubMed PMID: 23789932.

8. Cardozo LF, Pedruzzi LM, Stenvinkel P, Stockler-Pinto MB, Daleprane JB, Leite M, Jr., Mafra D. Nutritional strategies to modulate inflammation and oxidative stress pathways via activation of the master antioxidant switch Nrf2. Biochimie. 2013;95(8):1525-33. Epub 2013/05/07. doi: 10.1016/j.biochi.2013.04.012. PubMed PMID: 23643732.

9. Liu RH. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. Am J Clin Nutr. 2003;78(3 Suppl):517S-20S. Epub 2003/08/26. doi: 10.1093/ajcn/78.3.517S. PubMed PMID: 12936943.

10. Dennis KK, Liu KH, Uppal K, Go YM, Jones DP. Distribution of phytochelatins, metal-binding compounds, in plant foods: A survey of commonly consumed fruits, vegetables, grains and legumes. Food Chem. 2020;339:128051. Epub 2020/09/21. doi: 10.1016/j.foodchem.2020.128051. PubMed PMID: 32950899.

11. Jumarie C, Fortin C, Houde M, Campbell PG, Denizeau F. Cadmium uptake by Caco-2 cells: effects of Cd complexation by chloride, glutathione, and phytochelatins. Toxicol Appl Pharmacol. 2001;170(1):29-38. doi: 10.1006/taap.2000.9075. PubMed PMID: 11141353.

12. Fujita Y, el Belbasi HI, Min KS, Onosaka S, Okada Y, Matsumoto Y, Mutoh N, Tanaka K. Fate of cadmium bound to phytochelatin in rats. Res Commun Chem Pathol Pharmacol. 1993;82(3):357-65. PubMed PMID: 8122036.

13. ATSDR, Faroon O, Ashizawa A, Wright S, Tucker P, Jenkins K, Ingerman L, Rudisill C. Toxicological Profile for Cadmium. In: ATSDR, editor. Atlanta, GA 2012.

14. Satarug S, Vesey DA, Gobe GC. Health Risk Assessment of Dietary Cadmium Intake: Do Current Guidelines Indicate How Much is Safe? Environ Health Perspect. 2017;125(3):284-8. doi: 10.1289/EHP108. PubMed PMID: 28248635; PMCID: PMC5332171.

15. Scientific Opinion of the Panel on Contaminants in the Food Chain on a request from the European Commission on cadmium in food. The EFSA Journal. 2009;980:1-139.

16. Berglund M, Akesson A, Nermell B, Vahter M. Intestinal absorption of dietary cadmium in women depends on body iron stores and fiber intake. Environ Health Perspect. 1994;102(12):1058-66. Epub 1994/12/01. doi: 10.1289/ehp.941021058. PubMed PMID: 7713018; PMCID: PMC1567470.

17. Satarug S, Vesey DA, Gobe GC. Current health risk assessment practice for dietary cadmium: Data from different countries. Food Chem Toxicol. 2017;106(Pt A):430-45. doi: 10.1016/j.fct.2017.06.013. PubMed PMID: 28602857.

18. Jarup L, Berglund M, Elinder CG, Nordberg G, Vahter M. Health effects of cadmium exposure--a review of the literature and a risk estimate. Scand J Work Environ Health. 1998;24 Suppl 1:1-51. Epub 1998/05/07. PubMed PMID: 9569444.

19. Langelueddecke C, Lee WK, Thevenod F. Differential transcytosis and toxicity of the hNGAL receptor ligands cadmium-metallothionein and cadmium-phytochelatin in colon-like Caco-2 cells: implications for in vivo cadmium toxicity. Toxicol Lett. 2014;226(2):228-35. doi: 10.1016/j.toxlet.2014.01.049. PubMed PMID: 24518829.

20. Seregin IV, Kozhevnikova AD. Low-molecular-weight ligands in plants: role in metal homeostasis and hyperaccumulation. Photosynth Res. 2020. Epub 2020/07/13. doi: 10.1007/s11120-020-00768-1. PubMed PMID: 32653983.

21. Pal R, Rai JP. Phytochelatins: peptides involved in heavy metal detoxification. Appl Biochem Biotechnol. 2010;160(3):945-63. doi: 10.1007/s12010-009-8565-4. PubMed PMID: 19224399.

22. Dennis KK, Uppal K, Liu KH, Ma C, Liang B, Go YM, Jones DP. Phytochelatin database: a resource for phytochelatin complexes of nutritional and environmental metals. Database (Oxford). 2019;2019. Epub 2019/07/04. doi: 10.1093/database/baz083. PubMed PMID: 31267134; PMCID: PMC6606759.

23. Ding L, Wang L, Yu Z, Ma S, Du Z, Zhang T, Liu J. Importance of Terminal Amino Acid Residues to the Transport of Oligopeptides across the Caco-2 Cell Monolayer. J Agric Food Chem. 2017;65(35):7705-12. Epub 2017/08/16. doi: 10.1021/acs.jafc.7b03450. PubMed PMID: 28812357.

24. Chothe P, Singh N, Ganapathy V. Evidence for two different broad-specificity oligopeptide transporters in intestinal cell line Caco-2 and colonic cell line CCD841. Am J

Physiol Cell Physiol. 2011;300(6):C1260-9. Epub 2011/02/11. doi: 10.1152/ajpcell.00299.2010. PubMed PMID: 21307350.

25. Go YM, Orr M, Jones DP. Actin cytoskeleton redox proteome oxidation by cadmium. Am J Physiol Lung Cell Mol Physiol. 2013;305(11):L831-43. doi: 10.1152/ajplung.00203.2013. PubMed PMID: 24077948; PMCID: PMC3882527.

26. Blais A, Lecoeur S, Milhaud G, Tome D, Kolf-Clauw M. Cadmium uptake and transepithelial transport in control and long-term exposed Caco-2 cells: the role of metallothionein. Toxicol Appl Pharmacol. 1999;160(1):76-85. doi: 10.1006/taap.1999.8735. PubMed PMID: 10502504.

27. Dennis KK, Go YM, Jones DP. Redox Systems Biology of Nutrition and Oxidative Stress. J Nutr. 2019;149(4):553-65. Epub 2019/04/06. doi: 10.1093/jn/nxy306. PubMed PMID: 30949678; PMCID: PMC6461723.

28. Capuano E, Pellegrini N. An integrated look at the effect of structure on nutrient bioavailability in plant foods. J Sci Food Agric. 2019;99(2):493-8. Epub 2018/08/02. doi: 10.1002/jsfa.9298. PubMed PMID: 30066376.

29. Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. Nat Rev Microbiol. 2008;6(2):121-31. Epub 2008/01/09. doi: 10.1038/nrmicro1817. PubMed PMID: 18180751.

30. Artursson P, Palm K, Luthman K. Caco-2 monolayers in experimental and theoretical predictions of drug transport. Adv Drug Deliv Rev. 2001;46(1-3):27-43. Epub 2001/03/22. doi: 10.1016/s0169-409x(00)00128-9. PubMed PMID: 11259831.

31. Cobbett C, Goldsbrough P. Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. Annu Rev Plant Biol. 2002;53:159-82. doi: 10.1146/annurev.arplant.53.100301.135154. PubMed PMID: 12221971.

32. Liu J, Qu W, Kadiiska MB. Role of oxidative stress in cadmium toxicity and carcinogenesis. Toxicol Appl Pharmacol. 2009;238(3):209-14. Epub 2009/02/25. doi: 10.1016/j.taap.2009.01.029. PubMed PMID: 19236887; PMCID: PMC4287357.

33. Chekmeneva E, Gusmao R, Diaz-Cruz JM, Arino C, Esteban M. From cysteine to longer chain thiols: thermodynamic analysis of cadmium binding by phytochelatins and their fragments. Metallomics. 2011;3(8):838-46. doi: 10.1039/c1mt00028d. PubMed PMID: 21687859.

34. Lopez E, Arce C, Oset-Gasque MJ, Canadas S, Gonzalez MP. Cadmium induces reactive oxygen species generation and lipid peroxidation in cortical neurons in culture. Free Radic Biol Med. 2006;40(6):940-51. Epub 2006/03/17. doi: 10.1016/j.freeradbiomed.2005.10.062. PubMed PMID: 16540389.

35. Cuypers A, Plusquin M, Remans T, Jozefczak M, Keunen E, Gielen H, Opdenakker K, Nair AR, Munters E, Artois TJ, Nawrot T, Vangronsveld J, Smeets K. Cadmium stress: an oxidative challenge. Biometals. 2010;23(5):927-40. Epub 2010/04/03. doi: 10.1007/s10534-010-9329-x. PubMed PMID: 20361350.

36. Jones DP, Sies H. The Redox Code. Antioxid Redox Signal. 2015;23(9):734-46. doi: 10.1089/ars.2015.6247. PubMed PMID: 25891126; PMCID: PMC4580308.

37. Go YM, Jones DP. Redox theory of aging: implications for health and disease. Clin Sci (Lond). 2017;131(14):1669-88. doi: 10.1042/CS20160897. PubMed PMID: 28667066; PMCID: PMC5773128.

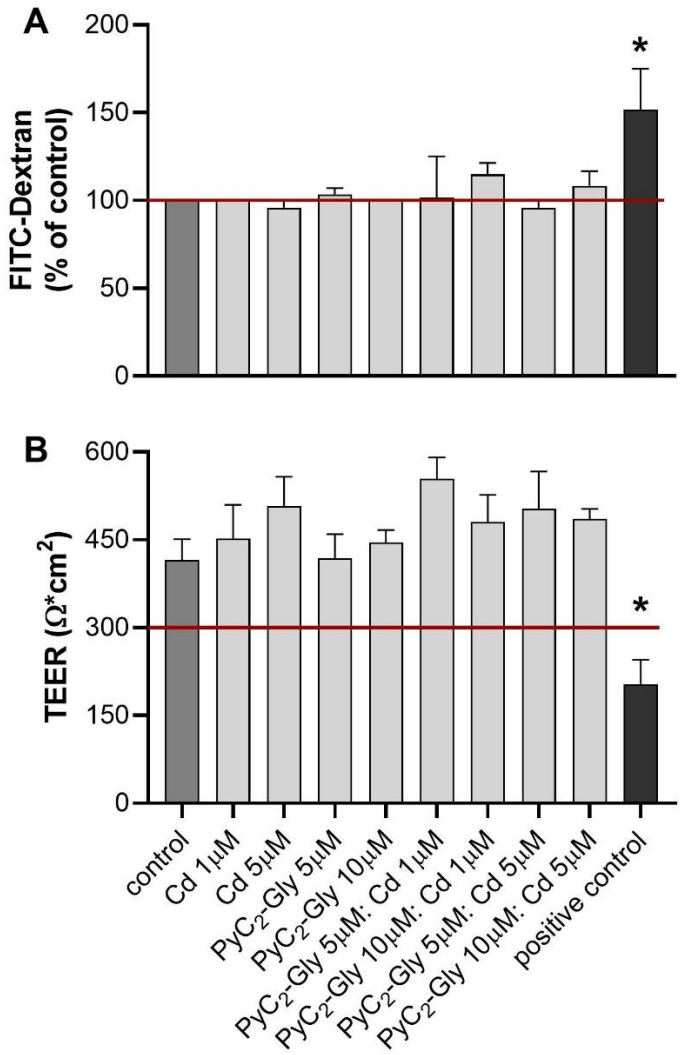

Figure 4.1. Caco-2 cell monolayer integrity as measured by FITC-Dextran permeability and transepithelial electrical resistance (TEER) after dosed with PyC_2-Gly and Cd for 4 h. A, FITC-Dextran level for control cells indicated by red line. B, TEER threshold for adequate monolayer integrity indicated by red line. Values presented are mean ± SEM; FITC-Dextran, n = 2; TEER, n = 3. Asterisk indicates difference from control tested by Student's t-test; * $p < 0.05$.

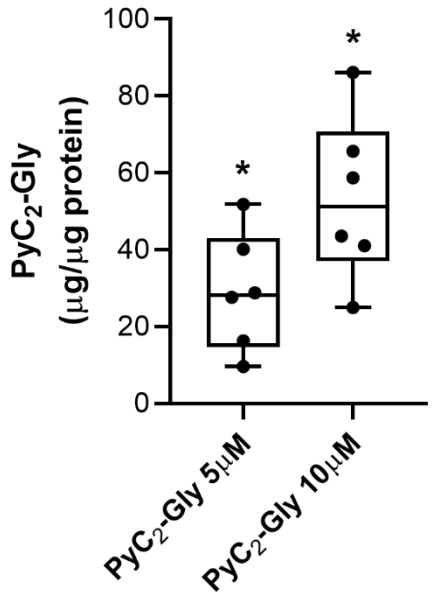

Figure 4.2. PyC_2-Gly uptake into cells at 4 hours by dose. Caco-2 cell monolayers were treated with 0, 5, and 10 µM PyC_2-Gly concentrations. Individual experiments are represented by filled circles. Data are summarized by box and whisker plots including the median, minimum, and maximum values. Asterisks indicate significant differences from control and between 5 and 10 µM by Student's t-test, *$p < 0.05$.

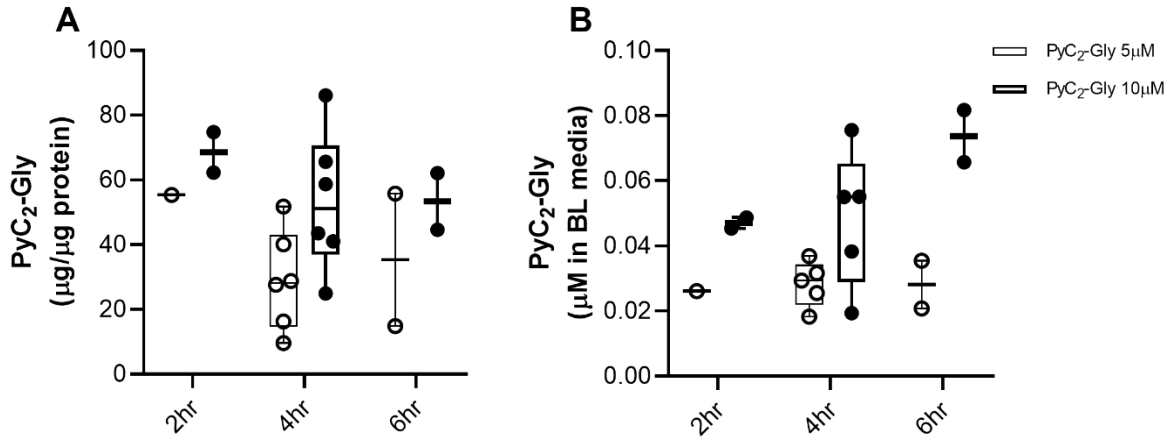

Figure 4.3. PyC_2-Gly uptake into and through cells over time. Individual experiments are represented by open (5 µM PyC_2-Gly) and filled (10 µM PyC_2-Gly) circles. Cellular PyC_2-Gly content increases initially at 2 h followed by a decrease and stabilization at 4 and 6 h for cells treated with 10 µM PyC_2-Gly (A). Concurrently, PyC_2-Gly content of basolateral media has a time-dependent increase (B). Cells treated with 5 µM PyC_2-Gly did not differ by treatment time in cellular or basolateral media PyC_2-Gly content. Data are summarized by box and whisker plots including the median, minimum, and maximum values. No significant differences were observed within treatment groups by time (one-way ANOVA; $p > 0.05$).

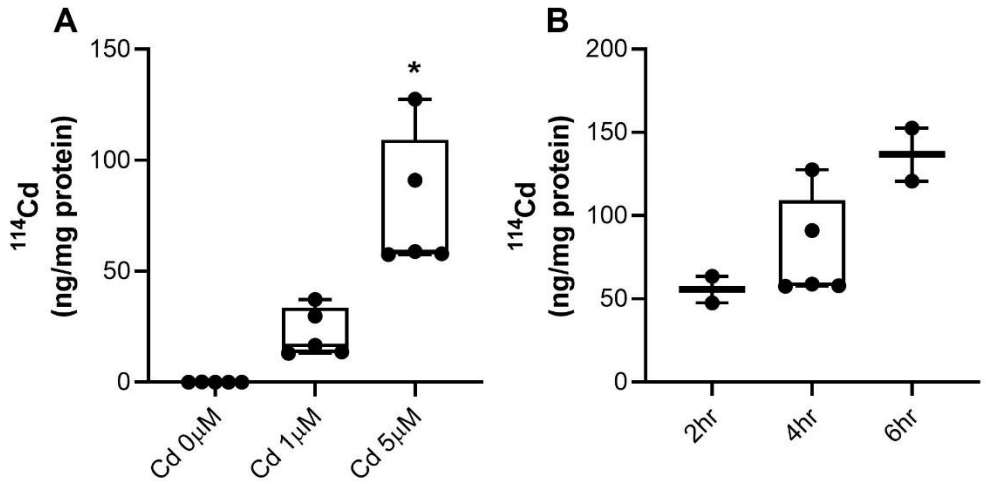

Figure 4.4. Cd uptake with increasing dose and time. Cells were treated with 0, 1, and 5 μM $CdCl_2$ for 4 h (A). Cells treated with 5 μM $CdCl_2$ were harvested at 2, 4, and 6 h (B). Caco-2 cellular content is increased with increasing dose at 4 h. In Caco-2 cells dosed with 5 μM $CdCl_2$, cellular content of Cd increases with increasing treatment time. Individual experiments are represented by filled circles. Data are summarized by box and whisker plots including the median, minimum, and maximum values. Significant differences were observed by dose (one-way ANOVA, $p < 0.05$). Asterisks indicate significant differences from control with 5 μM $CdCl_2$ treatment by post-hoc Dunnett's multiple comparisons test, * $p < 0.05$. By time, differences trended towards significance (one-way ANOVA, $p = 0.055$) with a significant linear trend ($p < 0.05$).

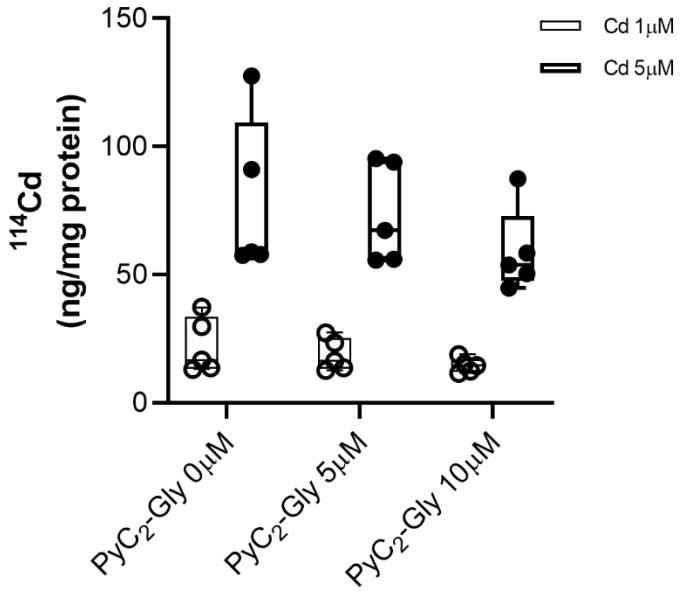

Figure 4.5. Cd uptake into cells co-treated with Cd and PyC_2-Gly. Cells were analyzed after treatment with 1 or 5 μM $CdCl_2$ in combination with 0, 5, or 10 μM PyC_2-Gly for 4 h. Individual experiments are represented by open (1 μM $CdCl_2$) and filled (5 μM $CdCl_2$) circles. Data are summarized by box and whisker plots including the median, minimum, and maximum values. No significant differences were observed by individual Cd doses with PyC_2-Gly doses (one-way ANOVA, $p > 0.05$).

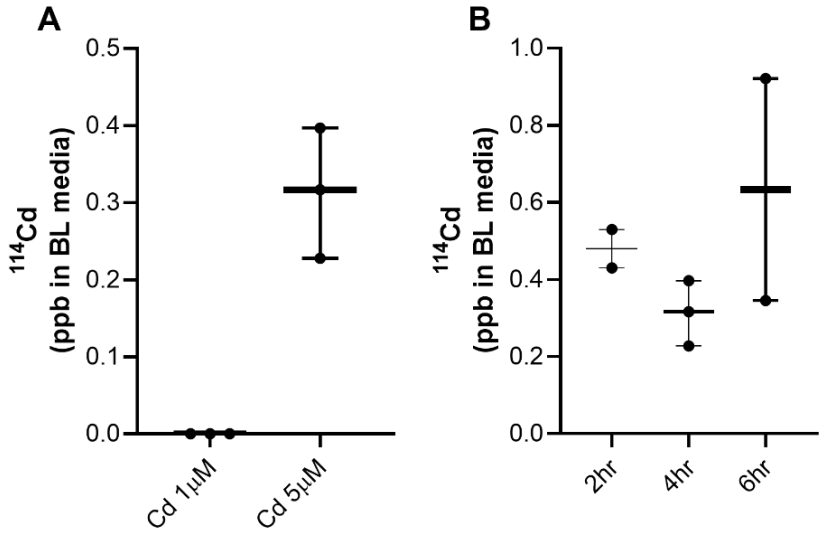

Figure S4.1 (Supplementary). Cd uptake through cells by dose and time. Individual experiments are represented by filled circles. Cells were treated with 1 and 5 µM $CdCl_2$ for 4 h (A). Basolateral media from cells treated with 5 µM $CdCl_2$ were harvested at 2, 4, and 6 h (B). Data are summarized by box and whisker plots including the median, minimum, and maximum values. No significant differences were observed by dose or time (one-way ANOVA, $p > 0.05$).

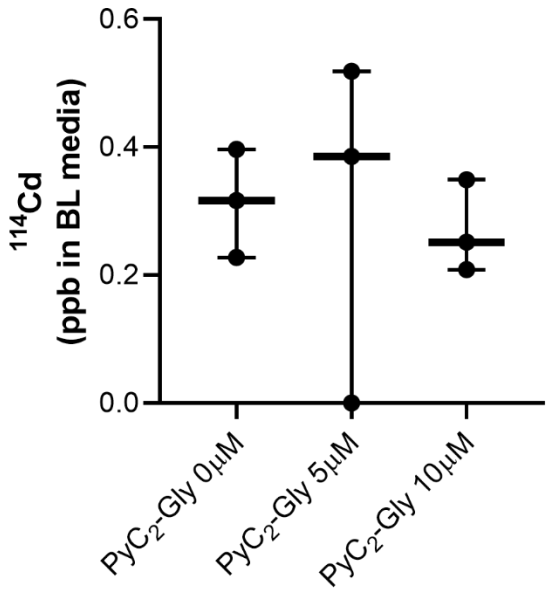

Figure S4.2 (Supplementary). Cd uptake through cells co-treated with $CdCl_2$ and PyC_2-Gly. Individual experiments are represented by filled circles. Basolateral media was collected after cells were treated with 5 μM $CdCl_2$ in combination with 0, 5, or 10 μM PyC_2-Gly for 4 h. Data are summarized by box and whisker plots including the median, minimum, and maximum values. No significant differences were observed by dose or time (one-way ANOVA, $p > 0.05$).

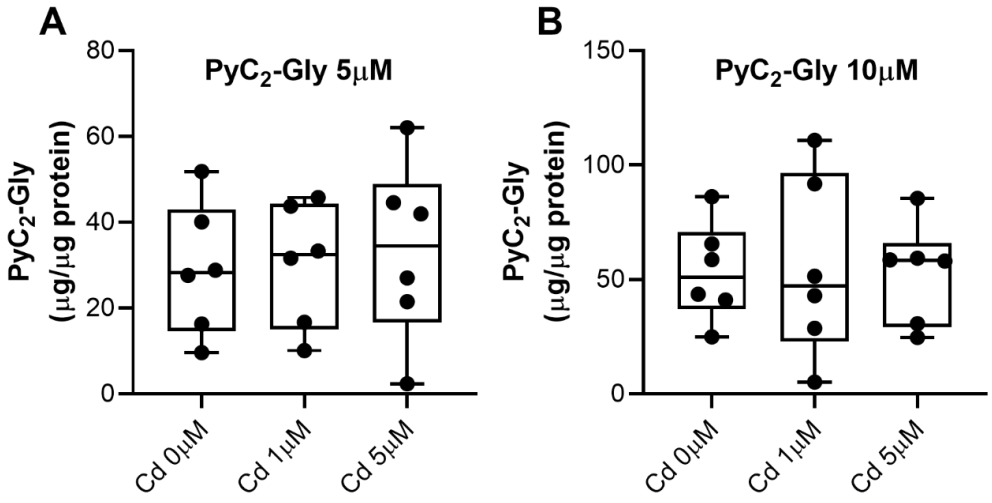

Figure S4.3 (Supplementary). PyC uptake into cells is similar in cells co-treated with Cd and PyC_2-Gly. Cells were analyzed for PyC_2-Gly content after co-treatment with combinations of Cd (0, 1, and 5 μM $CdCl_2$) and PyC_2-Gly (5 and 10 μM PyC_2-Gly) for 4 h. Individual experiments are represented by filled circles. Data are summarized by box and whisker plots including the median, minimum, and maximum values. No significant differences were observed by dose combination (one-way ANOVA, $p > 0.05$).

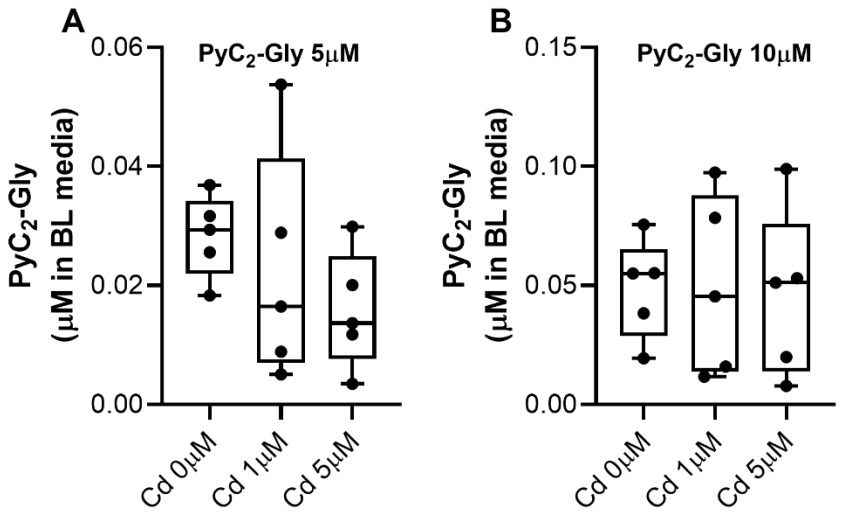

Figure S4.4 (Supplementary). PyC_2-Gly uptake through cells co-treated with Cd and PyC_2-Gly. Basolateral media was analyzed for PyC_2-Gly content after co-treatment with combinations of Cd (0, 1, and 5 µM $CdCl_2$) and PyC_2-Gly (5 and 10 µM PyC_2-Gly) for 4 h. Individual experiments are represented by filled circles. Data are summarized by box and whisker plots including the median, minimum, and maximum values. No significant differences were observed by dose combination (one-way ANOVA, $p > 0.05$).

CHAPTER 5

Plant food intake is associated with lower cadmium body burden in middle-aged adults

Kristine K. Dennis[a], Suzanne E. Judd[b], Jessica A. Alvarez[a,c], Ka Kahe[d], Dean P. Jones[a,e], Terryl J. Hartman[a,f]

[a]Nutrition and Health Sciences, Laney Graduate School, Emory University, Atlanta, GA, USA

[b]School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA

[c]Division of Endocrinology, Metabolism, and Lipids, Department of Medicine, Emory University, Atlanta, GA, USA

[d]Epidemiology and Obstetrics and Gynecology, Columbia University Irving Medical Center, New York City, NY, USA

[e]Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, Emory University, Atlanta, GA, USA

[f]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA

Submitted to *European Journal of Nutrition*, August 2020

Abstract

Purpose: Dietary intake is a primary source of cadmium (Cd) exposure in the non-smoking population. Plant foods containing metal-binding plant compounds such as polyphenols, phytates, and phytochelatins may reduce Cd bioavailability and result in lower Cd body burden. In this study, we investigated the association between plant food intake and urinary creatinine-adjusted Cd (uCd), a well-established marker of Cd body burden.

Methods: Participants were from a cross-sectional sample of 1,904 adults in the REasons for Geographic and Racial Differences in Stroke (REGARDS) cohort. Dietary intake was assessed with a food frequency questionnaire. We created a 12-point plant food score (PFS) based on reported intake across seven categories (fruits, vegetables, legumes, nuts/seeds, whole grains, tea, and wine). Higher scores indicated higher consumption and diversity of plant food intake. Multivariable linear regression models were used to estimate the association between PFS and uCd. Due to the influence of age and smoking on Cd status, stratified analyses were conducted.

Results: Mean PFS was 5.4 (SD: 2.5) and mean uCd was 0.53 µg/g creatinine (SD: 0.39). In adjusted models, PFS was not associated with uCd ($p > 0.05$). In stratified analyses, PFS was inversely associated with uCd ($p = 0.01$) with a 1-point higher PFS associated with 0.022 µg/g lower uCd among middle-aged (45-60) adults. No significant association was observed between PFS and uCd in older (>60) adults. The association of PFS and uCd did not differ by smoking status.

Conclusion: Our findings suggest higher plant food intake is associated with lower Cd body burden in middle-aged adults.

Key words: diet, plant food score, metal-binding, epidemiology, toxic metals

Introduction

The diet is the primary source of Cd exposure in the non-smoking U.S. population. Cd is a toxic metal which accumulates in the body over time, resulting in Cd body burdens with known health risks (1). Average levels of dietary Cd intake in the U.S. leads to Cd levels associated with increased risks of kidney dysfunction, bone demineralization, type 2 diabetes, and cancer, among others (2). Reductions in Cd exposure can be achieved by decreasing consumption of high Cd foods such as shellfish, organ meats, spinach, and potatoes (3). However, foods with low Cd levels can contribute proportionally more to overall Cd exposure due to high consumption (4). Minimizing dietary Cd exposure is an important component of reducing Cd risk but factors impacting bioavailability must also be considered as they are key determinants of Cd body burden.

In population studies, dietary Cd is not always associated with Cd body burden due to differences in bioavailability from factors such as age, fiber intake, and nutritional status (5). Cd bioavailability ranges from 5-10% but can be as high as 30% (1, 6-8). This high variation is not fully explained by known factors. Other dietary components may meaningfully impact Cd absorption and provide modifiable targets to reduce Cd body burden. Some compounds in the diet can bind to Cd ions and alter their absorption. Metal-binding plant compounds such as metallothioneins, phytochelatins, phytates and polyphenols are produced in the edible components of plants regularly consumed in the human diet (9-11). These compounds can impair absorption of specific metals. For example, phytates can decrease Cd absorption, while ascorbic acid can decrease lead absorption (12). The combined influence of metal-binding plant compounds may reduce toxic metal absorption and consequently, total body burden. If overall

plant food consumption in the diet can reduce Cd absorption, this would provide an important strategy for mitigating Cd-related disease risks.

In order to study the impact of interacting dietary components, dietary scores can be utilized for population health studies with dietary recall information. Different dietary scores have been successfully used to study health outcomes and intermediate biomarkers of health or disease including oxidative balance scores, the Mediterranean Diet Score, the Alternative Healthy Eating Index, and others (13-16). Dietary scores facilitate the investigation of the cumulative effects of multiple nutrients or compounds, as small effects of single nutrients can be challenging to observe in free-living populations (17). As comprehensive measurements of metal-binding plant compounds cannot be directly assessed from dietary recall information, a plant food score could capture the cumulative contributions of metal-binding phytochemicals consumed across the diet.

The impact of diet on Cd status is a product of many interacting factors including Cd directly consumed in the diet, mineral status of the individual, and influence of other compounds within the food matrix (18). Metal-binding plant compounds may be an important dietary component influencing Cd absorption, and may help explain the differential associations between dietary Cd consumed and overall Cd body burden. The purpose of this study was to investigate the association between a plant food intake score and Cd body burden, and whether this association differs in individuals by smoking status and age.

Methods

Study Design

The REGARDS (REasons for Geographic and Racial Differences in Stroke) study is a national population-based prospective cohort in the United States. The details of study design, recruitment, and data collection have been described previously (19). Briefly, 30,239 individuals ages 45 and older were randomly selected and recruited into the study between January 2003 and October 2007. Black individuals and individuals residing in the Stroke Belt region (i.e. eight southeastern states – Alabama, Mississippi, Louisiana, Georgia, Arkansas, Tennessee, South Carolina and North Carolina – with higher stroke mortality than the rest of the U.S.) were oversampled.

Trained personnel collected demographic characteristics, socioeconomic factors, and medical history of recruited participants through a telephone survey. Following the survey, a health professional visited each participant's home to conduct an in-person assessment to collect anthropometric measurements as well as blood and urine samples. The REGARDS study was approved by the institutional review boards at participating institutions, and all participants provided written informed consent.

Diet was assessed using a self-administered Block 98 Food Frequency Questionnaire (FFQ), which participants were asked to complete at the end of their baseline in-person assessment and return within 3 months. The FFQ assessed dietary intake over the past year and has been

described in detail elsewhere (20, 21). Completed FFQs were sent to Nutrition Quest (Berkeley, CA, USA) for the analysis of the nutrition content.

Population

The data for the current analysis are from a case-cohort study designed to investigate associations between trace elements and stroke risk. The case-cohort sample includes a random cohort sample from the REGARDS cohort (n = 2,666) and all incident ischemic stroke cases identified through September 2012 and not in the random cohort sample (n = 650). Details of the original case-cohort study are provided elsewhere (22). From the original case-cohort study (n = 3,316), 2,262 participants had urinary Cd measures and FFQ data within the normal energy intake range (M: 800-4,500 kcal; F: 500-3,500 kcal). We excluded an additional 358 participants with missing data on ACR, eGFR, smoking status, income, or education (**Figure 5.1**). Excluded participants were more likely to be Black, female, and have a lower level of education. A total of 1,904 participants were included in the final analysis.

Plant Food Score

Plant food score (PFS) includes five categories of plant foods and two plant-derived beverages, tea and wine (**Table 5.1**). These seven plant food score components incorporate major dietary factors contributing to consumption of metal-binding plant compounds. Plant food intake for fruit (including 100% fruit juice), vegetable, legumes, nuts/seeds, and whole grains were categorized into low, medium, and high intake categories by sex-specific tertiles. Low through high categories were assigned 0-2 points. For plant-derived beverage intake, those who

consumed at least 1 cup of tea/day or 1 drink of wine/week received 1 point and everyone else received 0 points. The points for each component were summed to calculate the overall PFS where higher total scores indicate greater consumption of plant foods.

Mixed dishes in the Block 98 FFQ were disaggregated using the My Pyramids Equivalent Database (MPED 2.0) using a modified version of a previously published method (23, 24). Briefly, each recipe variation was identified in MPED 2.0 for the mixed dishes in the FFQ. The frequency of recipe consumption for each mixed dish was determined using 2-day food records in Black and white individuals ≥ 45 years old from NHANES 2003-2004 and used to weight the contribution of each recipe to overall consumption of each mixed dish (25). The equivalents of each plant food group contained in the mixed dish recipes were recorded and the mean food group equivalents determined for each mixed dish using the weights for recipe consumption. The MPED equivalents were calculated based on the mean food group equivalents for that mixed dish and the quantity of the mixed dish consumed by each study participant. MPED equivalents were converted to grams for each food group. For each food group, the grams of consumption for mixed dishes and single food items were summed to determine grams consumed per food group for all foods.

Urinary cadmium

Urine samples were collected during the baseline visit and stored at -80ºC for future analyses. Urinary Cd was measured using the National Health and Nutrition Examination Survey (NHANES) method with a NEXION 300X quadrupole inductively coupled plasma mass

spectrometer (ICP-MS; Perkin Elmer, Waltham, MA) operated with a dynamic reaction cell to prevent interference from urinary molybdenum (26, 27). Urinary Cd was adjusted for urinary creatinine to account for dilution effects. Urinary creatinine was assessed with the Modular-P chemistry analyzer from Roche/Hitachi (Tokyo, Japan) (28).

Covariates

Age, sex, race (Black or white), smoking (never, former, or current), education (< high school, high school, some college, or ≥ college), and income (< 20k, 20-34k, 35-74k, ≥ 75k) were self-reported. Region was classified as Stroke Buckle (coastal plain regions of Georgia, North Carolina, and South Carolina), rest of Stroke Belt (the remaining area of Georgia, North Carolina, and South Carolina, plus Mississippi, Alabama, Tennessee, Louisiana, and Arkansas), or Non-Stroke Belt region. Total energy, fiber, and iron intake were obtained from the FFQ and analyzed as continuous variables. Weight (kilograms) and height (meters) were measured during the in-home visit using a standardized protocol. Body mass index (BMI) was calculated as weight (kg)/height (m^2). Albumin-to-creatinine ratio (ACR), a marker of kidney damage, was calculated using urinary albumin and creatinine measures. The estimated glomerular filtration rate (eGFR), an indicator of kidney function, was calculated using the Chronic Kidney Disease Epidemiology Collaboration equation (29).

Statistical Analyses

Characteristics of the study population were described for all participants and for participants in low (0-4 points), medium (5-8 points), and high (9-12 points) PFS categories. Mean values with

standard deviation (SD) or median with interquartile range (IQR) values were reported for continuous variables and frequencies for categorical variables. Differences in characteristics were compared across PFS categories with analysis of variance used for normally distributed and continuous variables, Kruskal-Wallis test for skewed and continuous variables, and X^2 test for categorical variables.

PFS was the independent variable of interest and urinary creatinine-adjusted Cd (uCd) was the dependent variable of interest. PFS was treated as a continuous variable in all regression models and divided into three categories for analyses of descriptive statistics. Confounding variables were determined a priori based on previous literature and biological plausibility. Multivariate linear regression models were used to estimate the association between PFS and uCd. Model 1 was adjusted for age, sex, total energy intake, and smoking. Model 2 was additionally adjusted for iron and fiber intake and Model 3 was additionally adjusted for eGFR, education, income, region, and race. We considered age (45-60 years vs. >60 years) and smoking status (current/former vs. never) as effect modifiers in stratified analyses. Interaction terms between PFS and age, as well as between PFS and smoking status (current/former/never) were included in models to determine if the associations between PFS and uCd were comparable across different ages and smoking status.

To explore the robustness of our findings, sensitivity analyses were conducted. As fiber intake is related to plant food consumption, we excluded fiber as a covariate to determine if it altered our findings. Additionally, kidney function can impact Cd excretion (1). Therefore, we excluded

individuals with estimated glomerular filtration rate (eGFR) below 30 and below 60 mL/min/1.73m^2, established thresholds indicative of mild to moderate kidney function loss (30). We analyzed all regression models with natural log-transformed uCd as the dependent variable due to the right skew of the biomarker data. All data analyses were performed in R (31).

Results

Of the 1,904 participants included in this analysis, 50.8% of participants were female and 66.3% were white. On average, participants were 65.6 years of age (SD: 9.2) and overweight (BMI: 29.0). About 60% of participants were either current or former smokers and over half had at least some college education. Median eGFR was 86.4 mL/min/1.73m^2 (IQR: 71.7-97.1) and median ACR was 7.4 mg/g (4.6-15.4) (**Table 5.2**). Mean uCd was 0.53 µg/g creatinine (SD: 0.39).

PFS was normally distributed and ranged from 0-12 points with a mean score of 5.4 (SD: 2.5). A greater proportion of individuals in the high PFS category were white and never smokers relative to those in the low PFS category. Participants in the highest PFS intake category tended to be older, were more likely to have completed at least college, and had higher incomes. Individuals in the low PFS category had about 3-fold lower fiber intake (9.6 ± 3.5 grams vs. 27.3 ± 8.2 grams; mean ± SD) and about 2-fold lower iron intake (9.1 ± 4.1 grams vs. 18.3 ± 6.4 grams; mean ± SD) on average. Energy intakes increased across categories of low, medium, and high PFS (1,324 kcal vs. 1,831 kcal vs. 2,338 kcal) (**Table 5.2**).

The unadjusted model of the association between PFS and uCd was significant, with uCd -0.011 µg/g creatinine lower for every 1-point higher in PFS. In the fully-adjusted model, the associations between PFS and uCd were attenuated and null (**Table 5.3**); some covariates were significantly associated with uCd including age, sex, smoking, eGFR, and race (**Table S5.1**).

When we explored effect modification by smoking status, a significant interaction was observed between current smokers and PFS (p-interaction = 0.04) but not between former smokers and PFS (p-interaction = 0.9). In stratified analyses by smoking status, PFS was not associated with uCd in current/former or never smokers (**Table 5.4**). Age, sex, eGFR, and region were associated with uCd in never smokers; sex, energy intake, iron intake, eGFR, income ≥ 75k and race were associated with uCd in current/former smokers (**Table S5.2**).

In analyses assessing effect modification by age, no significant interaction was observed between age and PFS (p-interaction = 0.2). In stratified analyses by age group, individuals ages 45-60 years had a significant association between PFS and uCd, such that for a 1-point increase in PFS there was a 0.022 µg/g decrease in uCd. In participants over age 60, there was no association between PFS and uCd (**Table 5.4**) but there were significant associations between uCd and sex, smoking, eGFR, and race (**Table S5.3**). In sensitivity analyses which excluded fiber intake as a covariate and individuals with either eGFR <30 mL/min/1.73m^2 or <60 mL/min/1.73m^2, the associations between PFS and uCd did not appreciably change. Results were similar when regression models were analyzed with natural log transformed uCd (data not shown).

Discussion

In this cross-sectional analysis, differential associations between PFS and uCd were observed by age. Among middle-aged (45-60) REGARDS participants, PFS was inversely associated with uCd independent of total energy intake and other potential confounding factors. No association was observed between PFS and uCd in older adults. The differential results observed across these two age groups may be due to multiple factors known to change with age. Cd accumulation and distribution starts to change around age 50 with greater accumulation in the liver relative to the kidney resulting in peak kidney Cd levels around age 60 (1, 3). Age-related increases in uCd start to plateau after age 50 and can decrease in older age (18). Additionally, differences in diet, nutritional status, and kidney function occur with increasing age (30, 32, 33). These biological changes along with other social factors can lead to reduced diet quality including inadequate consumption of legumes, whole grains, and vegetables (34, 35). Similarly, nutrient absorption and utilization can decrease over time, contributing to declines in nutritional status which could alter Cd bioavailability (36). Kidney function is an additional consideration as it declines with age. In our population, overall eGFR was decreased in older adults and was significantly associated with uCd in older but not middle-aged adults. Alterations in kidney function can impact excretion of Cd and has implications for the utility of using uCd as a marker of Cd body burden in those individuals (3).

In our study, a 1-point higher PFS was associated with 0.022 µg/g creatinine lower uCd in middle-aged adults. Considering this in a broader dietary context, an increase in PFS from a low to high level (e.g. 2 points to 10 points) would be associated with a 0.176 µg/g lower uCd level. Although we cannot assess causality in this study, this degree of decrease in uCd could result in

meaningful reductions in Cd-related disease risks across the population. In a study examining bone mineral density in women, women with uCd levels > 0.50 µg/g creatinine were at greater risk of osteoporosis (37). Another study in the general U.S. population found increased risk of all-cause and cardiovascular disease-related mortality for individuals with uCd levels of 0.57 µg/g creatinine (38). Recent NHANES biomonitoring data from 2015-2016 indicates a substantial portion of the U.S. population over age 50 has uCd levels within these ranges (2, 39). Our results suggest increased plant food intake could contribute to meaningfully reduced Cd body burden in middle-aged adults; however, the specific impact of consumption of metal-binding compounds through plant foods could not be directly assessed in our study. With increased characterization of metal-binding compounds in foods, we may be able to specifically assess the role of these compounds in Cd bioavailability. Given mineral intake and overall nutritional status can influence Cd absorption, future research investigating the association of plant-based dietary patterns with Cd body burden may capture both the influence of metal-binding plant compounds and other nutritional factors impacting Cd absorption.

Although biologically plausible, smoking status did not materially modify the association between PFS and uCd. In analyses assessing effect modification by smoking status, we found a significant interaction between current smoking and PFS. In stratified analyses, we found no association between PFS and uCd in never or current/former smokers. Among current smokers, their greatest Cd exposure is through inhaled tobacco smoke and therefore, changes to dietary Cd absorption will have less influence on overall Cd body burden. In never smokers, we expected plant food intake could influence Cd absorption and therefore, overall Cd body burden. However, we did not see this association and found significant associations with other factors

known to be related to increased Cd body burden such as age and sex (3). Similar to our study, many studies have observed that age and smoking status are key determinants of uCd (40, 41). Overall declines in uCd have been observed with declining smoking rates (42). In a nationally representative biomonitoring survey (NHANES 2011-2012), smokers age 50 and older had mean uCd levels of 0.629 µg/g creatinine while nonsmokers of the same age had mean uCd of 0.296 µg/g creatinine (39). We observed similar differences in our study with smokers (current/former) over age 45 having a mean uCd of 0.648 µg/g creatinine and never smokers having a mean uCd of 0.383 µg/g creatinine. Despite the influence of other factors such as smoking status, we did observe lower uCd levels in middle-aged adults with higher consumption and diversity of plant food intakes as assessed by PFS.

Previous studies investigating the relationship between diet and Cd body burden have identified associations with specific foods and overall dietary patterns. Individual foods have been associated with higher Cd levels including processed meat, potatoes, bread products, and seafood (43-46). However, these findings tend to be population specific and indicate the importance of considering overall dietary pattern and other individual factors in determining Cd exposure and absorption. Bioavailability studies with controlled diets demonstrate Cd absorption is influenced by the overall dietary context (47). Additionally, iron status is a major determinant of Cd absorption, particularly in women (48). In a study of pregnant women in the United Kingdom (UK), a “health conscious” dietary pattern including foods like salad, fruit, rice, pasta, legumes, and fish was associated with reduced blood Cd levels (49). In analyses of individual food groups in the UK study, greater consumption of leafy greens and green vegetables was associated with lower blood Cd levels. Although some foods such as leafy greens, shellfish, and potatoes are

known to have higher Cd contents, the correlation of estimated dietary Cd with Cd biomarkers is often low due to other major factors influencing Cd bioavailability in low exposure populations (5).

Strengths of this study include the dietary data, urinary creatinine-adjusted Cd measurements, and wide range of covariate data available with established and potential relevance to Cd body burden in almost 2000 people from the REGARDS cohort. We were able to adjust for eGFR, a sensitive marker of kidney function, in our analyses. Additionally, we created a novel plant food score using dietary recall data to assess the cumulative impact of consumption of metal-binding dietary compounds in a human population. Dietary intake was assessed with FFQ data which provides both strengths and limitations in our study. The FFQ provided information on food and beverage intake over the previous year which provides a long-term assessment of diet relevant to our research question. Our plant food score leveraged the FFQ data to investigate our research question but the plant food score has not been formally validated for this purpose. A limitation of the FFQ is the potential for misclassification due to self-report. Additionally, our study is cross-sectional which limits our ability to assess causality and creates challenges for evaluating Cd body burden which results from long-term exposures across the lifespan. Although we were able to account for iron intake, we did not have hemoglobin measures for our whole study population which would have allowed us to consider iron status. Finally, selection bias may be a concern as individuals excluded due to missing data were more likely be to Black, female, and have a lower level of education.

In conclusion, our study suggests that higher plant food intake, as assessed by PFS, is associated with lower levels of uCd in middle-aged but not older adults. This finding has important implications for long-term health risks from Cd because of the very long biologic half-life of Cd. The results suggest that in addition to other known health benefits from increased plant food intake, a decrease in Cd burden could improve population health through decreased risk of Cd-associated diseases, including cancer at multiple sites, osteoporosis, and kidney disease.

Acknowledgements

The authors thank the other investigators, the staff, and the participants of the REGARDS study for their valuable contributions. A full list of participating REGARDS investigators and institutions can be found at http://www.uab.edu/soph/regardsstudy/.

Funding

The REGARDS research project is supported by a cooperative agreement U01 NS041588 co-funded by the National Institute of Neurological Disorders and Stroke (NINDS) and the National Institute on Aging (NIA), National Institutes of Health, Department of Health and Human Services. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NINDS or the NIA. Representatives of the NINDS were involved in the review of the manuscript but were not directly involved in the collection, management, analysis or interpretation of the data. Additional funding was provided by the National Institute of Environmental Health Sciences (NIEHS, R01 ES021735). Representatives from NIEHS did not have any role in the design and conduct of the study, the collection,

management, analysis, and interpretation of the data, or the preparation or approval of the manuscript. The work of KD was supported by research grants from the NIH (T32 DK007734, F31 ES030980).

Author contributions All authors were responsible for the study conceptualization and design. KD was responsible for the data analysis. All authors contributed to the interpretation of the results. The manuscript was drafted by KD with input and critical revision from TH and DJ. All authors contributed to the final manuscript.

Ethics declarations

Conflict of interest The authors declare no conflict of interest.

Chapter 5 References

1. ATSDR, Faroon O, Ashizawa A, Wright S, Tucker P, Jenkins K, Ingerman L, Rudisill C. Toxicological Profile for Cadmium. In: ATSDR, editor. Atlanta, GA 2012.

2. Satarug S, Vesey DA, Gobe GC. Health Risk Assessment of Dietary Cadmium Intake: Do Current Guidelines Indicate How Much is Safe? Environ Health Perspect. 2017;125(3):284-8. doi: 10.1289/EHP108. PubMed PMID: 28248635; PMCID: PMC5332171.

3. Satarug S. Dietary Cadmium Intake and Its Effects on Kidneys. Toxics. 2018;6(1). doi: 10.3390/toxics6010015. PubMed PMID: 29534455; PMCID: PMC5874788.

4. Kim K, Melough MM, Vance TM, Noh H, Koo SI, Chun OK. Dietary Cadmium Intake and Sources in the US. Nutrients. 2018;11(1). Epub 2018/12/24. doi: 10.3390/nu11010002. PubMed PMID: 30577418; PMCID: PMC6356330.

5. Satarug S, Vesey DA, Gobe GC. Current health risk assessment practice for dietary cadmium: Data from different countries. Food Chem Toxicol. 2017;106(Pt A):430-45. doi: 10.1016/j.fct.2017.06.013. PubMed PMID: 28602857.

6. Flanagan PR, McLellan JS, Haist J, Cherian G, Chamberlain MJ, Valberg LS. Increased dietary cadmium absorption in mice and human subjects with iron deficiency. Gastroenterology. 1978;74(5 Pt 1):841-6. PubMed PMID: 640339.

7. Kikuchi Y, Nomiyama T, Kumagai N, Dekio F, Uemura T, Takebayashi T, Nishiwaki Y, Matsumoto Y, Sano Y, Hosoda K, Watanabe S, Sakurai H, Omae K. Uptake of cadmium in meals from the digestive tract of young non-smoking Japanese female volunteers. J Occup Health. 2003;45(1):43-52. PubMed PMID: 14605428.

8. Horiguchi H, Oguma E, Sasaki S, Miyamoto K, Ikeda Y, Machida M, Kayama F. Comprehensive study of the effects of age, iron deficiency, diabetes mellitus, and cadmium burden on dietary cadmium absorption in cadmium-exposed female Japanese farmers. Toxicol Appl Pharmacol. 2004;196(1):114-23. doi: 10.1016/j.taap.2003.11.024. PubMed PMID: 15050413.

9. Anjum NA, Hasanuzzaman M, Hossain MA, Thangavel P, Roychoudhury A, Gill SS, Rodrigo MA, Adam V, Fujita M, Kizek R, Duarte AC, Pereira E, Ahmad I. Jacks of metal/metalloid chelation trade in plants-an overview. Front Plant Sci. 2015;6:192. Epub 2015/04/18. doi: 10.3389/fpls.2015.00192. PubMed PMID: 25883598; PMCID: PMC4382971.

10. Schlemmer U, Frolich W, Prieto RM, Grases F. Phytate in foods and significance for humans: food sources, intake, processing, bioavailability, protective role and analysis. Mol Nutr Food Res. 2009;53 Suppl 2:S330-75. Epub 2009/09/24. doi: 10.1002/mnfr.200900099. PubMed PMID: 19774556.

11. Olszowy M. What is responsible for antioxidant properties of polyphenolic compounds from plants? Plant Physiol Biochem. 2019;144:135-43. Epub 2019/09/30. doi: 10.1016/j.plaphy.2019.09.039. PubMed PMID: 31563754.

12. Jadan-Piedra C, Chiocchetti GM, Clemente MJ, Velez D, Devesa V. Dietary compounds as modulators of metals and metalloids toxicity. Crit Rev Food Sci Nutr. 2018;58(12):2055-67. Epub 2017/07/08. doi: 10.1080/10408398.2017.1302407. PubMed PMID: 28686469.

13. Van Hoydonck PG, Temme EH, Schouten EG. A dietary oxidative balance score of vitamin C, beta-carotene and iron intakes and mortality risk in male smoking Belgians. J Nutr. 2002;132(4):756-61. Epub 2002/04/02. doi: 10.1093/jn/132.4.756. PubMed PMID: 11925473.

14. Goodman M, Bostick RM, Dash C, Flanders WD, Mandel JS. Hypothesis: oxidative stress score as a combined measure of pro-oxidant and antioxidant exposures. Ann Epidemiol. 2007;17(5):394-9. Epub 2007/04/28. doi: 10.1016/j.annepidem.2007.01.034. PubMed PMID: 17462547.

15. McCullough ML, Willett WC. Evaluating adherence to recommended diets in adults: the Alternate Healthy Eating Index. Public Health Nutr. 2006;9(1A):152-7. Epub 2006/03/04. doi: 10.1079/phn2005938. PubMed PMID: 16512963.

16. Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. N Engl J Med. 2003;348(26):2599-608. Epub 2003/06/27. doi: 10.1056/NEJMoa025039. PubMed PMID: 12826634.

17. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol. 2002;13(1):3-9. Epub 2002/01/16. doi: 10.1097/00041433-200202000-00002. PubMed PMID: 11790957.

18. Scientific Opinion of the Panel on Contaminants in the Food Chain on a request from the European Commission on cadmium in food. The EFSA Journal. 2009;980:1-139.

19. Howard VJ, Cushman M, Pulley L, Gomez CR, Go RC, Prineas RJ, Graham A, Moy CS, Howard G. The reasons for geographic and racial differences in stroke study: objectives and design. Neuroepidemiology. 2005;25(3):135-43. Epub 2005/07/02. doi: 10.1159/000086678. PubMed PMID: 15990444.

20. Gutierrez OM, Muntner P, Rizk DV, McClellan WM, Warnock DG, Newby PK, Judd SE. Dietary patterns and risk of death and progression to ESRD in individuals with CKD: a cohort study. Am J Kidney Dis. 2014;64(2):204-13. Epub 2014/04/01. doi: 10.1053/j.ajkd.2014.02.013. PubMed PMID: 24679894; PMCID: PMC4111976.

21. Newby PK, Noel SE, Grant R, Judd S, Shikany JM, Ard J. Race and region have independent and synergistic effects on dietary intakes in black and white women. Nutr J. 2012;11:25. Epub 2012/04/17. doi: 10.1186/1475-2891-11-25. PubMed PMID: 22500645; PMCID: PMC3386860.

22. Chen C, Xun P, Tsinovoi C, McClure LA, Brockman J, MacDonald L, Cushman M, Cai J, Kamendulis L, Mackey J, He K. Urinary cadmium concentration and the risk of ischemic stroke. Neurology. 2018;91(4):e382-e91. Epub 2018/06/24. doi: 10.1212/WNL.0000000000005856. PubMed PMID: 29934425; PMCID: PMC6070385.

23. Bowman SA, Friday JE, Moshfegh AJ. MyPyramid Equivalents Database, 2.0 for USDA Survey Foods, 2003-2004: Documentation and User Guide2008.

24. Byrd DA, Judd SE, Flanders WD, Hartman TJ, Fedirko V, Bostick RM. Development and Validation of Novel Dietary and Lifestyle Inflammation Scores. J Nutr. 2019;149(12):2206-18. Epub 2019/08/03. doi: 10.1093/jn/nxz165. PubMed PMID: 31373368; PMCID: PMC6887697.

25. CDC. NHANES 2003-2004: Data Documentation, Codebook, and Frequencies [Internet]2007.

26. Jarrett JM, Xiao G, Caldwell KL, Henahan D, Shakirova G, Jones RL. Eliminating molybdenum oxide interference in urine cadmium biomonitoring using ICP-DRC-MS. J Anal At Spectrom. 2008;23:962-7.

27. McElroy JA, Kruse RL, Guthrie J, Gangnon RE, Robertson JD. Cadmium exposure and endometrial cancer risk: A large midwestern U.S. population-based case-control study. PLoS One. 2017;12(7):e0179360. Epub 2017/07/26. doi: 10.1371/journal.pone.0179360. PubMed PMID: 28742092; PMCID: PMC5524364.

28. Aaron KJ, Campbell RC, Judd SE, Sanders PW, Muntner P. Association of dietary sodium and potassium intakes with albuminuria in normal-weight, overweight, and obese participants in the Reasons for Geographic and Racial Differences in Stroke (REGARDS) Study. Am J Clin Nutr. 2011;94(4):1071-8. Epub 2011/09/02. doi: 10.3945/ajcn.111.013094. PubMed PMID: 21880845; PMCID: PMC3173025.

29. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, Coresh J, Ckd EPI. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12. Epub 2009/05/06. doi: 10.7326/0003-4819-150-9-200905050-00006. PubMed PMID: 19414839; PMCID: PMC2763564.

30. Romagnani P, Remuzzi G, Glassock R, Levin A, Jager KJ, Tonelli M, Massy Z, Wanner C, Anders HJ. Chronic kidney disease. Nat Rev Dis Primers. 2017;3:17088. Epub 2017/11/24. doi: 10.1038/nrdp.2017.88. PubMed PMID: 29168475.

31. R Core Team. R: A Language and Environment for Statistical Computing. 2019.

32. Gekle M. Kidney and aging - A narrative review. Exp Gerontol. 2017;87(Pt B):153-5. Epub 2016/04/02. doi: 10.1016/j.exger.2016.03.013. PubMed PMID: 27032877.

33. Wakimoto P, Block G. Dietary intake, dietary patterns, and changes with age: an epidemiological perspective. J Gerontol A Biol Sci Med Sci. 2001;56 Spec No 2:65-80. Epub 2001/12/04. doi: 10.1093/gerona/56.suppl_2.65. PubMed PMID: 11730239.

34. Shlisky J, Bloom DE, Beaudreault AR, Tucker KL, Keller HH, Freund-Levi Y, Fielding RA, Cheng FW, Jensen GL, Wu D, Meydani SN. Nutritional Considerations for Healthy Aging and Reduction in Age-Related Chronic Disease. Adv Nutr. 2017;8(1):17-26. Epub 2017/01/18. doi: 10.3945/an.116.013474. PubMed PMID: 28096124; PMCID: PMC5227979.

35. Ervin RB. Healthy Eating Index scores among adults, 60 years of age and over, by sociodemographic and health characteristics: United States, 1999-2002. Adv Data. 2008(395):1-16. Epub 2008/07/24. PubMed PMID: 18646580.

36. Remond D, Shahar DR, Gille D, Pinto P, Kachal J, Peyron MA, Dos Santos CN, Walther B, Bordoni A, Dupont D, Tomas-Cobos L, Vergeres G. Understanding the gastrointestinal tract

of the elderly to develop dietary solutions that prevent malnutrition. Oncotarget. 2015;6(16):13858-98. Epub 2015/06/20. doi: 10.18632/oncotarget.4030. PubMed PMID: 26091351; PMCID: PMC4546438.

37. Gallagher CM, Kovach JS, Meliker JR. Urinary cadmium and osteoporosis in U.S. Women >or= 50 years of age: NHANES 1988-1994 and 1999-2004. Environ Health Perspect. 2008;116(10):1338-43. doi: 10.1289/ehp.11452. PubMed PMID: 18941575; PMCID: PMC2569092.

38. Tellez-Plaza M, Navas-Acien A, Menke A, Crainiceanu CM, Pastor-Barriuso R, Guallar E. Cadmium exposure and all-cause and cardiovascular mortality in the U.S. general population. Environ Health Perspect. 2012;120(7):1017-22. doi: 10.1289/ehp.1104352. PubMed PMID: 22472185; PMCID: PMC3404657.

39. CDC. Fourth Report on Human Exposure to Environmental Chemicals, Updated Tables. U.S. Department of Health and Human Services; 2019.

40. Gunier RB, Horn-Ross PL, Canchola AJ, Duffy CN, Reynolds P, Hertz A, Garcia E, Rull RP. Determinants and within-person variability of urinary cadmium concentrations among women in northern California. Environ Health Perspect. 2013;121(6):643-9. Epub 2013/04/05. doi: 10.1289/ehp.1205524. PubMed PMID: 23552363; PMCID: PMC3672909.

41. Riederer AM, Belova A, George BJ, Anastas PT. Urinary cadmium in the 1999-2008 U.S. National Health and Nutrition Examination Survey (NHANES). Environ Sci Technol. 2013;47(2):1137-47. Epub 2012/12/21. doi: 10.1021/es303556n. PubMed PMID: 23253114.

42. Tellez-Plaza M, Navas-Acien A, Caldwell KL, Menke A, Muntner P, Guallar E. Reduction in cadmium exposure in the United States population, 1988-2008: the contribution of declining smoking rates. Environ Health Perspect. 2012;120(2):204-9. Epub 2011/11/09. doi: 10.1289/ehp.1104020. PubMed PMID: 22062584; PMCID: PMC3279452.

43. Vacchi-Suzzi C, Eriksen KT, Levine K, McElroy J, Tjonneland A, Raaschou-Nielsen O, Harrington JM, Meliker JR. Dietary Intake Estimates and Urinary Cadmium Levels in Danish Postmenopausal Women. PLoS One. 2015;10(9):e0138784. Epub 2015/09/22. doi: 10.1371/journal.pone.0138784. PubMed PMID: 26390122; PMCID: PMC4577120.

44. Copes R, Clark NA, Rideout K, Palaty J, Teschke K. Uptake of cadmium from Pacific oysters (Crassostrea gigas) in British Columbia oyster growers. Environ Res. 2008;107(2):160-9. Epub 2008/03/25. doi: 10.1016/j.envres.2008.01.014. PubMed PMID: 18358470.

45. Sand S, Becker W. Assessment of dietary cadmium exposure in Sweden and population health concern including scenario analysis. Food Chem Toxicol. 2012;50(3-4):536-44. Epub 2012/01/14. doi: 10.1016/j.fct.2011.12.034. PubMed PMID: 22240415.

46. Olmedo P, Grau-Perez M, Fretts A, Tellez-Plaza M, Gil F, Yeh F, Umans JG, Francesconi KA, Goessler W, Franceschini N, Lee ET, Best LG, Cole SA, Howard BV, Navas-Acien A. Dietary determinants of cadmium exposure in the Strong Heart Family Study. Food Chem Toxicol. 2017;100:239-46. Epub 2016/12/26. doi: 10.1016/j.fct.2016.12.015. PubMed PMID: 28012896; PMCID: PMC5373690.

47. Vahter M, Berglund M, Nermell B, Akesson A. Bioavailability of cadmium from shellfish and mixed diet in women. Toxicol Appl Pharmacol. 1996;136(2):332-41. Epub 1996/02/01. doi: 10.1006/taap.1996.0040. PubMed PMID: 8619241.

48. Berglund M, Akesson A, Nermell B, Vahter M. Intestinal absorption of dietary cadmium in women depends on body iron stores and fiber intake. Environ Health Perspect. 1994;102(12):1058-66. Epub 1994/12/01. doi: 10.1289/ehp.941021058. PubMed PMID: 7713018; PMCID: PMC1567470.

49. Taylor CM, Doerner R, Northstone K, Kordas K. Maternal Diet During Pregnancy and Blood Cadmium Concentrations in an Observational Cohort of British Women. Nutrients. 2020;12(4). Epub 2020/04/01. doi: 10.3390/nu12040904. PubMed PMID: 32224907; PMCID: PMC7230298.

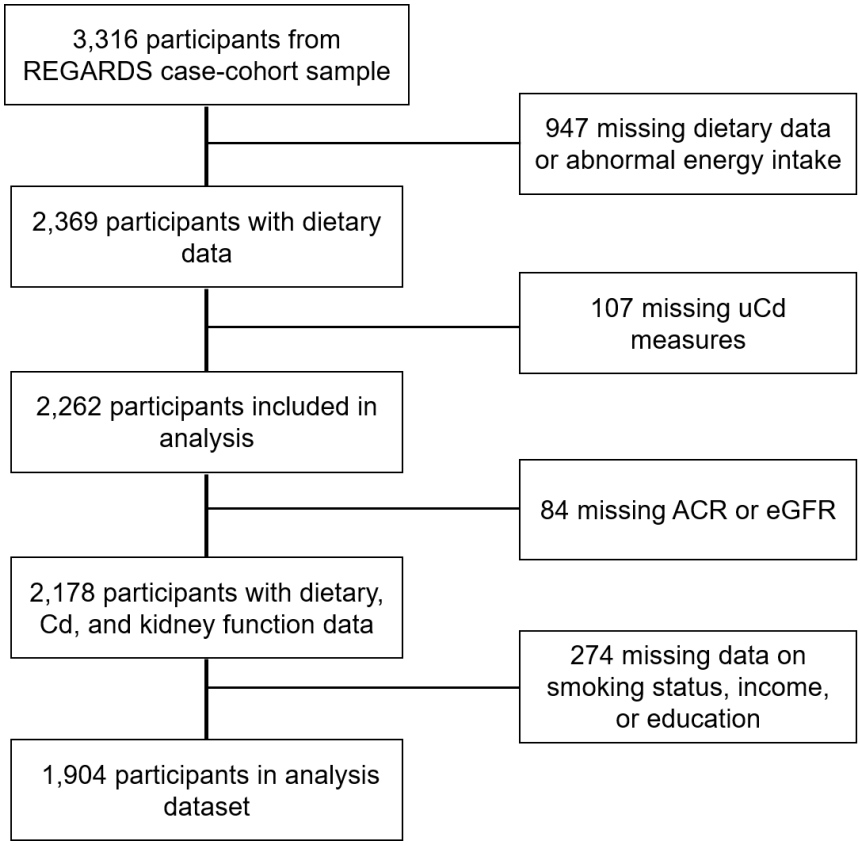

Figure 5.1. Flowchart of participants in REGARDS study who were included in the analysis. uCd, urinary cadmium; ACR, albumin/creatinine ratio; eGFR, estimated glomerular filtration rate

Table 5.1. Plant food score (PFS) components[a]

1. Fruit	0 = low (1st tertile), 1 = medium (2nd tertile), 2 = high (3rd tertile)
2. Vegetable	0 = low (1st tertile), 1 = medium (2nd tertile), 2 = high (3rd tertile)
3. Legumes	0 = low (1st tertile), 1 = medium (2nd tertile), 2 = high (3rd tertile)
4. Nuts/seeds	0 = low (1st tertile), 1 = medium (2nd tertile), 2 = high (3rd tertile)
5. Whole grains	0 = low (1st tertile), 1 = medium (2nd tertile), 2 = high (3rd tertile)
6. Tea	0 = less than 1 cup/day, 1 = greater than or equal to 1 cup/day
7. Wine	0 = less than 1 drink/week, 1 = greater than or equal to 1 drink/week

[a]Low, medium, and high categories correspond to sex-specific 1st, 2nd, and 3rd tertiles

Table 5.2. Characteristics of participants in REGARDS cohort sample by plant food score (PFS) interval (n = 1,904)

Characteristic[a]	All	PFS		
		low 0-4	medium 5-8	high 9-12
n	1,904	708	974	222
Age, years	65.6 (9.2)	64.9 (9.5)	65.9 (9.1)	66.6 (8.7)[b]
Female, %	50.8	50.6	52.5	44.1
White, %	66.3	61.7	68.1	73.4[b]
Total energy intake, kcal	1,702 (671)	1,324 (489)	1,831 (631)	2,338 (645)[b]
Smoking, %				
current	15.2	20.5	12.9	8.1[b]
former	44.9	38.9	39.5	46.8
never	39.9	40.7	47.5	45
BMI (missing = 14)	29.0 (5.8)	29.2 (5.8)	28.9 (5.7)	28.7 (6.0)
Region, %				
stroke buckle	20.8	20.1	20.9	22.5
rest of stroke belt	34.7	34.6	34.5	36
non-stroke belt	44.5	45.3	44.6	41.4
eGFR, mL/min/1.73m^2	86.4 (71.7-97.1)	86.2 (69.7-97.4)	86.5 (72.2-96.8)	86.0 (75.3-97.1)
ACR, mg/g	7.4 (4.6-15.4)	7.2 (4.4-15.0)	7.4 (4.6-15.4)	7.7 (5.0-16.1)
Education, %				
≥ college	37.7	32.6	39.6	45.0[b]
some college	28	27.1	28.6	28.4
high school	25	27.7	24	20.7
< high school	9.3	12.6	7.7	5.9
Income USD, %				
75k and above	18.6	16	19.3	23.9[b]
35-74k	36.7	35.3	36.7	41.4
20-34k	27.7	29.8	27.1	23.4
less than 20k	17	18.9	16.9	11.3
Iron intake, mg	12.6 (6.1)	9.1 (4.1)	13.8 (5.8)	18.3 (6.4)[b]
Fiber intake, g	15.9 (8.0)	9.6 (3.5)	17.8 (6.3)	27.3 (8.2)[b]

BMI, body mass index; eGFR, estimated glomerular filtration rate; ACR, albumin/creatinine ratio; uCd, urinary cadmium. [a] Values for age, BMI, total energy intake, fiber and iron intake are reported as mean (± SD). Values for eGFR and ACR are reported as median (IQR). Sex, race, smoking status, region, income, and education are reported as percent. [b] $p < 0.05$ based on the ANOVA, Kruskal-Wallis, or X^2 tests, as appropriate.

Table 5.3. Multivariate linear regression models of PFS with urinary cadmium[a].

	β	95% CI	*p* value
Unadjusted	-0.0110	-0.0181, -0.0040	0.002
Model 1[b]	-0.0009	-0.0082, 0.0064	0.8
Model 2[c]	-0.0002	-0.0095, 0.0091	0.06
Model 3[d]	-0.0014	-0.0106, 0.0079	0.8

[a] Urinary cadmium is creatinine-adjusted.

[b] Adjusts for age, sex, total energy intake, and smoking.

[c] Additionally adjusts for iron intake and fiber intake.

[d] Additionally adjusts for eGFR, education, income, region, and race.

Table 5.4. Fully adjusted models showing the association between PFS and uCd, stratified by age and smoking status.

Age[a]	45-60 years	>60 years
n	524	1380
PFS	-0.022	0.006
95% CI	-0.040, -0.005	-0.005, 0.017
p value	0.01	0.3
Smoking[b]	Never	Current/Former
n	855	1049
PFS	-0.001	-0.003
95% CI	-0.011, 0.009	-0.019, 0.013
p value	0.8	0.7

PFS, plant food score; uCd, urinary creatinine-adjusted cadmium; eGFR, estimated glomerular filtration rate; CI, confidence interval. [a] Adjusted for sex, total energy intake, smoking status, iron intake, fiber intake, eGFR, education, income, region, and race. [b] Adjusted for age, sex, iron intake, fiber intake, eGFR, education, income, region, race.

Table S5.1 (Supplementary). Fully-adjusted multivariate linear regression model showing the association between all independent variables and uCd.

	β	95% CI	*p* value
Intercept	0.024	-0.162, 0.210	0.8
PFS	-0.001	-0.011, 0.008	0.8
Age, years	0.006	0.004, 0.008	<0.001
Sex (M, 1; F, 0)	-0.245	-0.278, -0.212	<0.001
Total energy intake, kcal	0.000	0.000, 0.000	0.6
Former smoker	0.220	0.186, 0.252	<0.001
Current smoker	0.559	0.513, 0.605	<0.001
Iron intake, mg	-0.003	-0.007, 0.001	0.1
Fiber intake, g	0.000	-0.003, 0.004	0.9
eGFR, mL/min/1.73m^2	0.002	0.001, 0.003	<0.001
≥ college	0.001	-0.033, 0.035	1
income 20-34k	0.044	-0.003, 0.090	0.7
income 35-74k	0.010	-0.038, 0.035	0.7
income ≥ 75k	-0.051	-0.108, 0.006	0.1
Rest of stroke belt	-0.020	-0.054, 0.013	0.2
Stroke buckle	-0.028	-0.068, 0.012	0.2
Race (B, 1; W, 0)	-0.092	-0.126, -0.058	<0.001

PFS, plant food score; uCd, urinary creatinine-adjusted cadmium; M, male; F, female; eGFR, estimated glomerular filtration rate; B, black; W, white.

Table S5.2 (Supplementary). Smoking status stratified, fully-adjusted multivariate linear regression model showing the association between all independent variables and uCd.

Smoking status	Never (n = 855)			Current/Former (n = 1049)		
	β	95% CI	*p* value	β	95% CI	*p* value
Intercept	0.0380	-0.154, 0.230	0.7	0.5882	0.272, 0.904	0.0003
PFS	-0.0010	-0.011, 0.009	0.8	-0.0032	-0.019, 0.013	0.7
Age, years	0.0042	0.002, 0.006	<0.001	0.0033	-0.000, 0.007	0.053
Sex (M, 1; F, 0)	-0.1711	-0.207, -0.135	<0.001	-0.3262	-0.380, -2.72	<0.001
Total energy intake, kcal	0.0000	-0.000, 0.000	0.09	0.0001	0.000, 0.000	0.003
Iron intake, mg	0.0000	-0.004, 0.004	1.0	-0.0065	-0.013, -0.0003	0.04
Fiber intake, g	0.0024	-0.001, 0.006	0.2	-0.0056	-0.011, 0.0002	0.06
eGFR, mL/min/1.73m^2	0.0021	0.001, 0.003	<0.001	0.0019	0.0004, 0.003	0.01
≥ college	0.0343	-0.003, 0.071	0.07	-0.0216	-0.078, 0.035	0.4
income 20-34k	-0.0093	-0.060, 0.041	0.7	0.0436	-0.033, 0.120	0.3
income 35-74k	0.0280	-0.024, 0.080	0.3	-0.0677	-0.145, 0.009	0.09
income ≥ 75k	-0.0533	-0.115, 0.009	0.1	-0.1579	-0.250, -0.066	<0.001
Rest of stroke belt	-0.0479	-0.084, -0.012	0.009	0.0106	-0.046, 0.067	0.7
Stroke buckle	-0.0578	-0.101, -0.015	0.008	0.0056	-0.060, 0.072	0.9
Race (B, 1; W, 0)	-0.0422	-0.078, -0.006	0.2	-0.1315	-0.188, -0.075	<0.001

PFS, plant food score; uCd, urinary creatinine-adjusted cadmium; M, male; F, female; eGFR, estimated glomerular filtration rate; B, black; W, white.

Table S5.3 (Supplementary). Age stratified, fully-adjusted multivariate linear regression model showing the association between all independent variables and uCd.

Age	45-60 years (n = 524)			>60 years (n = 1380)		
	β	95% CI	*p* value	β	95% CI	*p* value
Intercept	0.432	0.248, 0.616	<0.001	0.422	0.322, 0.523	<0.001
PFS	-0.022	-0.040, -0.005	0.01	0.006	-0.005, 0.017	0.3
Sex (M, 1; F, 0)	-0.222	-0.282, -0.162	<0.001	-0.249	-0.288, -0.210	<0.001
Total energy intake, kcal	0.000	0.000, 0.000	0.5	0.000	0.000, 0.000	0.9
Former smoker	0.188	0.121, 0.254	<0.001	0.236	0.198, 0.275	<0.001
Current smoker	0.418	0.344, 0.493	<0.001	0.625	0.567, 0.683	<0.001
Iron intake, mg	-0.004	-0.013, 0.004	0.3	-0.003	-0.007, 0.001	0.1
Fiber intake, g	0.005	-0.001, 0.011	0.1	-0.001	-0.004, 0.003	0.8
eGFR, mL/min/1.73m^2	0.001	-0.001, 0.003	0.3	0.002	0.001, 0.003	0.002
≥ college	0.034	-0.029, 0.100	0.3	-0.004	-0.044, 0.036	0.8
income 20-34k	0.039	-0.061, 0.140	0.4	0.046	-0.007, 0.099	0.09
income 35-74k	0.019	-0.074, 0.111	0.7	0.003	-0.053, 0.060	0.9
income ≥ 75k	-0.060	-0.161, 0.040	0.2	-0.066	-0.136, 0.004	0.07
Rest of stroke belt	-0.069	-0.131, -0.006	0.03	-0.010	-0.050, 0.030	0.6
Stroke buckle	-0.088	-0.162, -0.015	0.02	-0.015	-0.062, 0.032	0.5
Race (B, 1; W, 0)	-0.019	-0.079, 0.041	0.5	-0.115	-0.156, -0.075	<0.001

PFS, plant food score; uCd, urinary creatinine-adjusted cadmium; M, male; F, female; eGFR, estimated glomerular filtration rate; B, black; W, white.

CHAPTER 6

A perspective on phytochelatins as beneficial components of plant-based diets

The mechanisms through which plant-based dietary patterns are beneficial extend to the interactions of nutritional and environmental exposures occurring simultaneously in the diet. However, a full understanding of these mechanisms remains incomplete. With the thousands of compounds we are exposed to in the diet, many interactions remain unknown or understudied. With advancements in high throughput technologies such as ultra-high resolution liquid chromatography mass spectrometry, there is increased opportunity to investigate the complexity of dietary factors. In this dissertation, we contribute to the knowledge base regarding the mechanisms through which plant-based dietary patterns contribute to health by establishing the plausibility that PyCs, metal-binding plant compounds, are beneficial through functions in metal homeostasis and protection from toxic metal exposures. Studies with molecular and population-based approaches provide critical new information on PyCs found in foods that we eat, showing their distributions and abundances as well as their absorption and functional impacts.

Key Findings

We created a PyC database (PyCDB) as a tool for investigations of PyC and PyC-metal complexes of nutritional and environmental metals (Chapter 2). With over 46,000 unique PyC masses, the database provides a tool for use with high-resolution mass spectrometry data to facilitate investigations of PyCs. Additionally, we enhanced the utility of the database by creating a web-based tool for direct mass queries with experimental data and providing an example workflow for using the PyCDB. Previous research has focused on only a few PyCs at a

time. With advancements in high throughput analyses, our database and related tools provide an opportunity to significantly advance our understanding of PyCs in biological systems.

To understand the contributions of individual foods to dietary PyC consumption, we characterized the types and concentrations of PyCs in 20 commonly consumed plant foods for the first time (Chapter 3). We determined PyC_2-Gly is the most abundant and widely found PyC. Additionally, PyC_3-Gly, PyC_4-Gly, PyC_2-Ala, and PyC_2-Glu were found at low concentrations in some foods. More processed foods tended to have lower PyC concentrations. Previous research has focused on PyCs in agricultural plants in response to soil metal exposures but the contents in edible plant components consumed in the human diet was unknown (1). Our research demonstrates that PyCs are widely found in the diet with concentrations specific to individual food types.

Using the most common PyC observed in foods, PyC_2-Gly, we investigated the transport and function of PyC_2-Gly in a human intestinal epithelial cell model system (Chapter 4). We found that PyC_2-Gly was transported into and through epithelial cells when treated with PyC_2-Gly levels similar to those in the human diet. One other study has demonstrated PyC_3-Gly uptake into and through intestinal cells but no previous studies have quantified PyC transport or investigated PyC_2-Gly (2). In functional assessments investigating if PyC_2-Gly reduces Cd absorption, we found a small but non-significant reduction in Cd absorption with PyC_2-Gly co-treatment. Other studies with different PyC types (i.e. PyC_3-Ser, PyC_3-Gly, and PyC_5-Gly) have found reduced Cd uptake in intestinal cells and reduced Cd distribution to the liver and kidneys in rats (3, 4). Based

upon these studies, there is a likelihood that our study was underpowered and needs to be repeated with a larger number of replicates. Nonetheless, our research establishes that a common dietary PyC, PyC_2-Gly, is absorbed into and through intestinal epithelial cells at concentrations expected in the diet. Additionally, when considered in the context of other functional PyC studies, our research also suggests PyC_2-Gly may reduce Cd absorption.

To explore how overall consumption of metal-binding plant compounds may impact Cd accumulation in human populations, we created a plant food score (PFS) to assess the association between higher plant food intake and Cd body burden in a cross-sectional study within the REGARDS cohort (Chapter 5). Among middle-aged (45-60) but not older (>60) adults, we found PFS was inversely associated with urinary Cd (uCd), a well-established marker of Cd body burden. Every 1-point higher PFS was associated with 0.022 µg/g uCd lower, suggesting higher intake of plant-based foods over the long-term could reduce Cd body burden, resulting in meaningful decreases in disease risks. Prior research has found some dietary patterns and individual foods associated with lower Cd levels but to our knowledge, no prior research has attempted to capture the role of metal-binding plant compounds on Cd body burden via a dietary score in a human population (5, 6). Our novel approach allowed us to assess the contributions of overall plant food intake to Cd status. Our study suggests higher plant food intake is associated with lower Cd body burdens and establishes an approach for future investigations of the interaction of metal-binding plant compounds and toxic metal body burdens.

Future directions

Our research establishes novel findings regarding PyCs in the human diet and their potential roles in health. First, this dissertation contributes foundational knowledge of dietary PyCs and tools for advancing PyC research. In addition, the findings of this research and its limitations also establish the needs and open opportunities for greater understanding of the biological functions and impacts of dietary PyCs on human health.

Through our initial survey of plant foods, we determined PyC quantities are highly variable within and across plant food groups, and other factors such as food processing and variety likely impact PyC contents. An extensive survey of food types by different factors will allow more precise assessment of the contributions of different plant foods to overall PyC consumption. With more comprehensive characterization of PyCs in foods, we can strengthen our investigations in human populations. While our novel PFS served as a useful proxy for assessing intake of metal-binding plant compounds, direct measures of PyCs in all plant foods captured in food frequency questionnaires would facilitate more exact examinations of the relationship between PyCs and toxic metals in population studies. Additionally, future investigations in longitudinal cohorts which include other age groups would overcome the limitations of cross-sectional analyses and improve the generalizability of our research. Kidney function declines with age and can impact uCd measures, highlighting the importance of longitudinal data for assessing outcomes related to long-term exposures (7-9). As smoking is a major determinant of Cd body burden, larger sample sizes would allow more nuanced investigations by smoking status, particularly with detailed measures of lifelong cigarette dose.

As PyCs are produced in response to soil metals, understanding if metal contents of plant foods relate to PyC contents is an important question. In preliminary metal analyses of the 20 foods in our study, there were significant correlations between Cd and PyC_2-Gly contents in only two foods (data not shown). Similarly, nutritional metals, including selenium and manganese, had a few significant correlations with PyC_2-Gly within individual foods but were also limited. PyC synthesis in plants is induced by a variety of metals which likely contributes to this complexity (1). Whether there are common correlations between specific metals and PyCs in plant foods remains an open area for exploration.

Finally, many questions remain regarding the biological mechanisms of PyC transport and functions within the body. There is evidence of PyC_3-Gly absorption via receptor-mediated endocytosis in intestinal cell models and more generally, transport of peptides with five amino acids via oligopeptide transporters (2, 10). Understanding the transport mechanisms of dietary PyCs, particularly in regards to peptide length and terminal amino acid, will (i) inform if transport occurs competitively with other compounds, (ii) inform if intact PyC-metal complexes could be transported, and (iii) inform how absorbed PyC may interact with other cell types once within the body. PyCs bind with other metals, including essential metals such as copper and zinc, and the binding affinity is dependent on the specific metal and biological context, such as pH. Whether PyCs functionally influence other metals is an important open question for nutrition science and toxicology. With this knowledge, we could extend our research into other human health questions. For example, could diets high in PyCs protect from toxic metal exposures in drinking water or even decrease toxic metal body burdens? Additionally, PyCs are redox active

molecules with capacity to bind metals and impact redox systems. Even low PyC concentrations in the body could have a role in metal homeostasis and overall redox control, potentially protecting from oxidative stress.

Conclusions

This dissertation establishes that PyCs are widely found in plant-based diets and could have a primary role in protecting from toxic metal exposures. With expanded databases of PyC types and abundances across foods in the human diet, future investigations can inform the broader role of dietary PyCs, their interactions with toxic metals, and their role in human health. Through building upon the knowledge created in this dissertation with expanded assessments of PyCs, we may be able to improve dietary recommendations to mitigate risks of toxic metal exposures and contribute to lifelong health.

Chapter 6 References

1. Seregin IV, Kozhevnikova AD. Low-molecular-weight ligands in plants: role in metal homeostasis and hyperaccumulation. Photosynth Res. 2020. Epub 2020/07/13. doi: 10.1007/s11120-020-00768-1. PubMed PMID: 32653983.

2. Langelueddecke C, Lee WK, Thevenod F. Differential transcytosis and toxicity of the hNGAL receptor ligands cadmium-metallothionein and cadmium-phytochelatin in colon-like Caco-2 cells: implications for in vivo cadmium toxicity. Toxicol Lett. 2014;226(2):228-35. doi: 10.1016/j.toxlet.2014.01.049. PubMed PMID: 24518829.

3. Jumarie C, Fortin C, Houde M, Campbell PG, Denizeau F. Cadmium uptake by Caco-2 cells: effects of Cd complexation by chloride, glutathione, and phytochelatins. Toxicol Appl Pharmacol. 2001;170(1):29-38. doi: 10.1006/taap.2000.9075. PubMed PMID: 11141353.

4. Fujita Y, el Belbasi HI, Min KS, Onosaka S, Okada Y, Matsumoto Y, Mutoh N, Tanaka K. Fate of cadmium bound to phytochelatin in rats. Res Commun Chem Pathol Pharmacol. 1993;82(3):357-65. PubMed PMID: 8122036.

5. Taylor CM, Doerner R, Northstone K, Kordas K. Maternal Diet During Pregnancy and Blood Cadmium Concentrations in an Observational Cohort of British Women. Nutrients. 2020;12(4). Epub 2020/04/01. doi: 10.3390/nu12040904. PubMed PMID: 32224907; PMCID: PMC7230298.

6. Vahter M, Berglund M, Nermell B, Akesson A. Bioavailability of cadmium from shellfish and mixed diet in women. Toxicol Appl Pharmacol. 1996;136(2):332-41. Epub 1996/02/01. doi: 10.1006/taap.1996.0040. PubMed PMID: 8619241.

7. ATSDR, Faroon O, Ashizawa A, Wright S, Tucker P, Jenkins K, Ingerman L, Rudisill C. Toxicological Profile for Cadmium. In: ATSDR, editor. Atlanta, GA 2012.

8. Gekle M. Kidney and aging - A narrative review. Exp Gerontol. 2017;87(Pt B):153-5. Epub 2016/04/02. doi: 10.1016/j.exger.2016.03.013. PubMed PMID: 27032877.

9. Romagnani P, Remuzzi G, Glassock R, Levin A, Jager KJ, Tonelli M, Massy Z, Wanner C, Anders HJ. Chronic kidney disease. Nat Rev Dis Primers. 2017;3:17088. Epub 2017/11/24. doi: 10.1038/nrdp.2017.88. PubMed PMID: 29168475.

10. Chothe P, Singh N, Ganapathy V. Evidence for two different broad-specificity oligopeptide transporters in intestinal cell line Caco-2 and colonic cell line CCD841. Am J Physiol Cell Physiol. 2011;300(6):C1260-9. Epub 2011/02/11. doi: 10.1152/ajpcell.00299.2010. PubMed PMID: 21307350.