Development of Enabling Tools for Global Profiling and Quantitative Analysis of Protein Post-translational Modifications

By

Yatao Shi

A dissertation in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Pharmaceutical Sciences)

at the

University of Wisconsin-Madison

2020

Date of final oral examination: 12/20/19

The dissertation is approved by the following members of the Final Oral Committee:

Lingjun Li, Professor, Pharmacy and Chemistry

Tim S. Bugni, Professor, Pharmacy

Paul C. Marker, Professor, Pharmacy

Jiaoyang Jiang, Associate Professor, Pharmacy

Michael R. Sussman, Professor, Biochemistry

Acknowledgements

I would first like to acknowledge my advisor Dr. Lingjun Li, without whom this work would not be possible. During my graduate school days, she gives me constant guidance, support and freedom to pursue my academic and professional goals. I am impressed by her perpetual energy and enthusiasm in research, and feel extremely fortunate and forever grateful for having an opportunity to work for such a brilliant and open-minded mentor.

I would also like to thank my committee members, Dr. Jiaoyang Jiang, Dr. Tim Bugni, Dr. Paul. Marker and Dr. Michael Sussman, for their valuable suggestions and insights in my preliminary exam, annual committee meeting, and this dissertation. I wish to give my thanks to Dr. Jiaoyang Jiang and Dr. Tim Bugni for giving me an opportunity to rotate in their labs.

I have greatly benefited from many previous Li Lab members. When I just joined lab, Dr. Ling Hao mentored me metabolomics and taught me many lab skills. Dr. Bingming Chen introduced me into the field of tissue imaging and spent a lot time train me the instrument and sample preparation. Without her help and support, two imaging projects in this dissertation will be impossible. Dr. Zhengwei Chen taught me proteomics from the sample preparation to data analysis. Dr. Fengfei Ma trained me the maintenance and troubleshooting of multiple instrument platforms. Dr. Jillian Johnson, Dr. Dustin Frost, Dr. Qing Yu, Dr. Matt Glover, Dr. Zhidan Liang, Dr. Shan Jiang and Dr. Chenxi jia have provided me valuable suggestion, support and their expertise to a variety of issues I faced during research.

I also want to say thanks to many current Li lab members. Zihui. Li has been working with me on protein citrullination project for more than 3 years. We worked together to overcome all faced challenges and get the method worked. He makes great contributions to the citrullination

project. It is my honor to get the opportunity to work with such a talented colleague. Yusi Cui gives me great help in the Dileu labeling and data analysis. I am grateful for the opportunity to be in the same lab with them. Dr. Junfeng Huang taught me a lot and is always glad to answer my questions and share his experience. Dr. Xiaofang Zhong, Qingyin Yu, Bin Wang, Mike Li, Chris Saucer, Dylan Tabang and Kellen DeLaney have always been wonderful colleagues and friends.

I want to take this opportunity to thank my research collaborators. Dr. Xudong (Daniel) Shi and Dr. Bowen Wang are knowledgeable collaborators and I am grateful for all their support and suggestion. Dr. Lian-Wang Guo and Dr. K. Craig Kent showed me their great support to my restenosis project. I am also fortunate enough to collaborate with Dr. Manish, Dr.Yajing Peng from Dr. Luigi Puglielli's lab, Dr. Yiping Liu from Dr. Judith Smith's group and Dr. Fen Zhu from Dr. Lixin Rui's lab. They are all talented and experienced people to work with and I learnt a lot from them.

I would like to thank Dr. Zhengqing Ye and Dr. John Feltenberger in Medicinal Chemistry Center for the synthesis of tags for my research projects. Their assistance makes great contribution to my citrullination projects. I also want to thank Dr. Cameron Scarlett and Molly Pellitteri Hahn of the Analytical Instrumentation Center for training me and giving me the access to instruments. In addition, I would like to extend my thanks to Ken Niemeyer, Josh Cutler and Gary Girdaukas for their continuous support and assistance in the past five years.

My final thanks go to my family. I appreciate the unconditional love and support given by my parents to me. I also want to thank my two older sisters to take care of my parents in the past five years. Without the love and support from my family, I could not overcome all the challenges in my life. Love you forever!

Table of Contents

Acknowledgements ..... **i**

List of Abbreviations and Acronyms ..... **v**

Chapter 1. Introduction and Research Summary

Introduction and Research Summary 2
References 8

Chapter 2. Development of a Novel Biotin Tag-assisted Mass Spectrometry-based Method for Large-scale Protein Citrullination Analysis

Abstract 15
Introduction 16
Materials and Methods 18
Results and Discussion 25
Conclusions 30
Acknowledgements 31
References 32
Supplemental Information 49

Chapter 3. MS-based Methods for Quantitative Analysis of Citrullinated Proteins in Biological Samples

Abstract 60
Introduction 61
Materials and Methods 62
Results and Discussion 68
Conclusions 73
References 74
Supplemental Information 85

Chapter 4. Enabling Global Protein Homocitrullination Analysis by Using a Novel Biotin Tag-assisted Mass Spectrometry-based Method

Abstract 93
Introduction 94
Materials and Methods 96
Results and Discussion 101
Conclusions 105
Acknowledgments 106
References 107
Supplemental Information 121

Chapter 5. MS Investigation of Signaling Pathways Involved in Synergistic Effect of LPS and TPG on Enhanced IFN-β Production in Macrophages
Abstract....128
Introduction....129
Materials and Methods....130
Results and Discussion....136
Conclusions....140
Acknowledgements....140
References....141
Supplemental Information....156

Chapter 6. Mass Spectrometry Imaging of N-glycans from Formalin-fixed Paraffin-embedded Tissue Sections Using a Novel Subatmospheric Pressure Ionization Source
Abstract....161
Introduction....162
Materials and Methods....164
Results and Discussion....168
Conclusions....172
Acknowledgments....173
References....174
Supplemental Information....183

Chapter 7. Mass Spectrometric Imaging Reveals Temporal and Spatial Dynamics of Bioactive Lipids in Arteries Undergoing Restenosis
Abstract....191
Introduction....192
Materials and Methods....193
Results and Discussion....197
Conclusions....205
Acknowledgments....206
References....207
Supplemental Information....221

Chapter 8. Conclusions and Future Directions
Conclusions....234
Future Directions....237
References....242

Appendix I. Publications and Presentations
Peer-Reviewed Publications....251
Patent and Conference Presentations....252

Development of Enabling Tools for Global Profiling and Quantitative Analysis of Protein Post-translational Modifications

Yatao Shi

Under the supervision of Professor Lingjun Li

At the University of Wisconsin-Madison

Abstract

Protein post-translational modifications (PTMs) involve the covalent chemical modifications of specific amino acid residues in proteins, which play crucial roles in alteration of protein physiochemical properties, structures and their physiological functions. Alterations in protein PTMs have been implicated in many serious diseases. Comprehensive and systematic study of disease-related protein PTMs is critical to explore their roles in the pathogenesis of diseases, contributing to the diagnosis and clinical treatment of diseases. Recently, given the capability of monitoring and identifying thousands of peptides simultaneously, mass spectrometry (MS) has evolved as a powerful tool in bottom-up proteomics, especially the protein PTM analysis.

This dissertation is devoted to the development and application of novel electrospray ionization mass spectrometry (ESI-MS) based strategies for the in-depth profiling and quantitative analysis of several important protein PTMs as well as matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) technique for the *in-situ* imaging of

biomolecules in complex tissue samples. A portion of this dissertation describes the development of a novel biotin tag-assisted MS-based method, as the first method, for the in-depth profiling of citrullinated and homocitrullinated proteins in biological samples. Its utility is demonstrated through its application in mouse tissue-specific citrullination and homocitrullination analysis. We further demonstrate the combination of the biotin tag-assisted MS-based method with different quantitative techniques, enabling the simultaneously qualitative and quantitative analysis of citrullinated and homocitrullinated proteins from different biological samples for the first time. Meanwhile, quantitative proteomics and phosphoproteomics was performed to reveal the signaling pathways involved in the macrophages treated by lipopolysaccharides (LPS) and thapsigargin (TPG).

This dissertation also highlights another two projects focusing on in-situ imaging of biomolecules from frozen tissue sections or formalin-fixed paraffin-embedded (FFPE) tissue sections using MALDI MS imaging (MSI). MALDI MSI of mouse arteries undergoing restenosis revealed the involvement of many bioactive lipids in the progress of restenosis. In addition, a novel subatmospheric pressure ionization source has been coupled with high resolution accurate mass (HRAM) MS for the high-resolution imaging of N-glycans from FFPE tissue sections. Application of this new SubAP/MALDI MS platform to FFPE mouse ovarian cancer tissue section unraveled the specific distribution of high-mannose N-glycans in the tumor region, suggesting potential association of this type of N-glycans with tumor progression.

List of Abbreviations and Acronyms

ACN	Acetonitrile
MeOH	Methanol
NaCl	Sodium chloride
SDS	Sodium dodecyl sulfate
FA	Formic acid
TFA	Trifluoroacetic acid
EDTA	Ethylenediaminetetracetic acid
BCA	Bicinchoninic acid assay
DTT	Dithiothreitol
IAA	Iodoacetamide
Lys-C/Trypsin	Endoproteinase lysine C/Trypsin mixture
Lys-C	Endoproteinase lysine C
SCX	Strong cation exchange chromatography
NanoUPLC	Nano ultra-performance liquid chromatography
HILIC	Hydrophilic interaction chromatography
RPLC	Reversed phase liquid chromatography
LC	Liquid chromatography
RT	Retention time

MS	Mass spectrometry
HRAM	High-resolution, accurate-mass.
ESI	Electrospray ionization
Da	Dalton, unit of mass (1 Da = 1 g/mol)
z	charge
m/z	mass-to-charge ratio
AGC	Automatic gain control
CID	Collision-induced dissociation
HCD	Higher-energy collisional dissociation
ETD	Electron transfer dissociation
EThcD	Electron-transfer/higher-energy collision dissociation
MS^1	Precursor mass analysis
TIC	Total ion chromatogram
MS^2	Tandem mass spectrometry
NCE	Normalized collision energy
ppm	Parts per million
PSM	Peptide-spectrum match
PTM	Post-translational modification
FDR	False discovery rate

GO	Gene ontology
S/N	Signal-to-noise ratio
DiLeu	Dimethylated leucine
iTRAQ	Isobaric tag for relative and absolute quantification
TMT	Tandem mass tag
WT	Wild type
SMC	Smooth muscle cell
TGF-β	Transforming growth factor beta
MALDI	Matrix-assisted laser desorption ionization
FFPE	Formalin-fixed paraffin-embedded
CHCA	α-Cyano-4-hydroxycinnamic acid
DHB	2,5-Dihydroxybenzoic acid

Chapter 1

Introduction and Research Summary

Introduction and Research Summary

Protein post-translational modifications (PTMs) represent the covalently biochemical modifications of proteins following ribosomal translation to form the mature proteins. More than 200 different types of protein PTMs have currently been identified, for example, protein phosphorylation, ubiquitylation, acetylation and glycosylation. These PTMs play an important role in the regulation of protein folding and translocation, protein-protein interactions, protein degradation, gene expression as well as the signaling and regulatory processes.[1-5] It has been revealed that abnormal alteration of many protein PTMs are associated with the onset and progress of many devastating diseases, such as Alzheimer's disease, cardiovascular diseases and some cancers.[6-8] Hence, comprehensive and systematic profiling of the dynamic changes of disease-related protein PTMs is not only critical for unraveling the pathogenesis process of diseases, but also beneficial for the diagnosis and treatment of diseases in clinic.

In past decades, mass spectrometry (MS) including electrospray ionization mass spectrometry (ESI-MS) and matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS), has evolved as a powerful tool for the high throughput analysis of biological samples.[9-13] The ESI-MS has been widely applied to the field of proteomics, especially the qualitative and quantitative analysis of protein PTMs, while MALDI-MS has an advantage of retaining spatial resolution and directly visualizing biomolecules in tissue sections.

This dissertation is devoted to the development and application of novel liquid chromatography mass spectrometry (LC-MS)-based strategies for the in-depth profiling and quantitative analysis of several important protein PTMs, such as citrullination (Chapters 2-3), homocitrullination (Chapter 4) and phosphorylation (Chapter 5). In addition, MALDI-MS technique is applied to the *in-situ* imaging of biomolecules including N-glycans from FFPE tissue

sections (Chapter 6) and lipids from fresh tissue sections (Chapter 7). Chapter 1 serves as a brief introduction of the dissertation work and the major outcomes of each project. The major findings from each research chapter and future directions for the line of research as well as some preliminary results will be summarized in Chapter 8.

Chapter 2 is focused on the development of a novel biotin tag-assisted, MS-based method for large-scale identification of citrullinated proteins in biological samples. Protein citrullination representing the change of arginine residue to citrulline residue, is an important PTM involved in many important cellular processes and diseases.[14-19] Current protein citrullination analysis mainly relies on conventional antibody-based techniques, but suffers from low throughput capabilities.[20,21].MS-based bottom-up proteomics has been proven to be a useful tool for large-scale protein PTM analysis. However, its application to protein citrullination analysis suffers from several challenges. First, methods for the specific enrichment of citrullinated peptides for MS analysis are not effective; Second, the mass shift of 0.984 Da induced by citrullination is the same as that of another modification called deamidation. Third, ^{13}C isotopic peaks in tandem MS spectra have a mass shift of 1.0033 Da, which is close to that of citrullination. These aspects will interfere with the accurate annotation of peptide fragment ions upon automated MS data processing, leading to misidentification of citrullination sites. Hence, it is critical to develop a novel approach, which is able to simultaneously deal with these challenges, for MS-based in-depth profiling of citrullinated proteins in complex biological samples. In this chapter, I reported a novel biotin thiol tag for the specific chemical derivatization and enrichment of citrullinated peptides from biological samples. Moreover, the biotin tag labeled citrullinated peptides are amenable to MS analysis. Due to the enlarged mass shift post biotin tag labeling, deamidation and ^{13}C isotopic peaks in tandem MS spectra do not cause interference with the accurate identification of citrullinated peptides and

citrullination site annotation. The novel approach has been demonstrated by citrullinated peptide standard and applied to mouse tissue-specific protein citrullination analysis.

Continuing the work described in the previous chapter, we introduced novel MS-based approaches for quantitative protein citrullination analysis. Some fluorescence probes have been reported for semiquantitative protein citrullination analysis.[22,23] However, these methods cannot identify citrullinated proteins at a large scale, or reveal the dynamic change of individual citrullinated protein in different samples, limiting its application to clinical sample analysis. In this chapter, I integrated the biotin tag-assisted, MS-based method reported in Chapter 2 with the stable-isotope dimethyl labeling technique, for the simultaneously qualitative and quantitative analysis of citrullinated proteins in various biological samples. The dimethyl labeling of citrullinated peptide, including reaction condition and labeling efficiency, was optimized. The quantitative performance of the integrated method was further systematically evaluated using the citrullinated peptide standard and complex biological samples, and reliable quantitation accuracy has been achieved. Furthermore, to increase the multiplexing capacity, I also tested the coupling of the biotin tag-assisted, MS-based method with our own N, N-dimethyl leucine (DiLeu) multiplex isobaric tags to multiplex quantitative protein citrullination analysis. This strategy has been demonstrated using citrullinated peptide standard though its performance on biological samples has not been assessed.

Chapter 4 demonstrated the global protein homocitrullination analysis enabled by the biotin tag-assisted, MS-based method developed in Chapter 2. Protein homocitrullination, also called protein carbamylation, is a type of non-enzymatic protein post-translation modification (PTM) associated with many serious diseases.[24-27] Although Some anti-homocitrullination antibodies have been commercially available and used for ELISA or WB analysis of target homocitrullinated

proteins, method for large-scale protein homocitrullination analysis is still not available. I notice that protein homocitrullination results in the formation of ureido group at the side chain of lysine residue, and thereby could be derivatized by the novel biotin thiol tag reported in Chapter 2. I have tested and validated the applicability of biotin tag-assisted, MS-based method to large-scale protein homocitrullination analysis. Mouse tissue-specific protein homocitrullination analysis has been performed to identify unknown protein substrates containing homocitrullination modification and explore the potential physiological role of this disease-related PTM.

In Chapter 5, quantitative proteomics and phosphoproteomics are performed to reveal the dynamic changes of proteins and phosphoproteins in macrophages treated in different conditions. Lipopolysaccharide (LPS) is a well-known endotoxin, which can induce the production of IFN-β in macrophages. Thapsigargin (TPG) is non-competitive inhibitor of the sarco/endoplasmic reticulum Ca^{2+} ATPase (SERCA). TPG-treated macrophages suffer from endoplasmic reticulum (ER) stress and the subsequent unfolded protein response (UPR).[28] Our collaborators observed that TPG-pretreated macrophages undergoing an intracellular UPR can respond to LPS treatment with greatly enhanced IFN-β production.[29,30] Although they have demonstrated that X-box binding protein 1 (XBP1) splicing and the phosphorylation of interferon regulatory factor (IRF3) contributed to this synergy effect, the global change of cellular proteins and signal pathways in LPS/TPG-treated macrophages has not been explored. To better understand the synergistic effect of LPS and TPG on the enhanced IFN-β production, I conducted quantitative proteomics and phosphoproteomics in mouse macrophages treated with four different conditions (Control, LPS, TPG and LPS/TPG) using our 12-plex DiLeu isobaric labeling tags. A large number of proteins and phosphoproteins were significantly changed in LPS/TPG-treated macrophages compared to

other three groups, providing new explanations and potential mechanisms to the synergy effect of LPS and TPG treatment.

Chapter 6 discusses the development of a novel MS platform, which couples a new subatmospheric pressure (SubAP)/MALDI source with a Q Exactive HF hybrid quadrupole-orbitrap mass spectrometer, for in-situ imaging of N-linked glycans from formalin-fixed paraffin-embedded (FFPE) tissue sections. N-linked glycosylation is a type of common PTM, which plays an important role in many diseases including Alzheimer's disease, cancers and cardiovascular diseases.[6-8] Although ESI-MS-based glycomic and glycoproteomic analyses have been applied to investigate N-glycan expression changes between healthy and disease specimens, [31-34] the homogenization of tissue samples conducted in the vast majority of previous studies inevitably resulted in the loss of spatial information regarding N-glycan localization. MALDI-MS has been introduced for *in-situ* biomolecular analysis due to its capability of retaining spatial information and directly visualizing biomolecules in tissue sections. Its application to *in-situ* N-glycan imaging has been reported.[35-37] In this Chapter, I have successfully established a novel SubAP/MALDI MS platform for the imaging of N-glycans from tissue sections. Compared to traditional vacuum MALDI source, the SubAP/MALDI source is independent of mass spectrometer and can be switched back and forth with an ESI source on an MS instrument platform, providing greater instrumental versatility and flexibility to allow both ESI-based analysis and MALDI-based analysis on a single mass spectrometer. Furthermore, the new SubAP/MALDI MS platform enables high resolution MS imaging acquisition, and therefore can be applied to small tissue section analysis. I have applied the SubAP/MALDI MS to image N-glycans on mouse FFPE brain section and mouse FFPE ovarian tissue section with cancer. High-resolution N-glycan images were collected and different N-glycan distribution patterns were observed.

Chapter 7 describes the *in situ* lipidomic imaging in a rat model mimicking angioplasty-elicited clinical restenosis using MALDI imaging MS. The rat model features well-characterized stages of disease progression. Rat carotid arteries over the course of restenosis progression at three time points post angioplasty (i.e. day 3, 7 and 14) were subjected to MALDI imaging analysis. After MALDI-MSI acquisition, the sections were further stained using Verhoeff-Van Gieson (VVG) staining. The overlap of the histology and MALDI images enables the discovery of bioactive lipids involved in the progression of restenosis. These results demonstrate the utility of MALDI image MS in unraveling the pathological mechanisms of diseases.

Chapter 8 offers conclusions drawn from the results obtained from each previous chapter and introduces future directions of my work. Despite that a novel biotin tag-assisted, MS-based method has been developed, which is the first method enabling large-scale identification of citrullinated and homocitrullinated proteins in biological samples (Chapters 2 and 4), I plan to develop a second-generation, more improved method, called filter-aided, thiol resin-assisted chemoenzymatic (FATRAC) method, for protein citrullination and homocitrullination analysis. The new FATRAC method will use customized multi-functional agarose beads as an alternative of biotin tag and streptavidin beads for the derivatization and enrichment of citrullinated peptides for MS analysis, overcoming several challenges of current biotin tag-assisted, MS-based method. I hope my methodology development work could provide the scientific community a powerful tool to unravel the important roles of protein citrullination and homocitrullination in the onset and progression of many devastating diseases.

References

(1) Butterfield, S.; Hejjaoui, M.; Fauvet, B.; Awad, L.; Lashuel, H. A., Chemical strategies for controlling protein folding and elucidating the molecular mechanisms of amyloid formation and toxicity. *J Mol Biol* 2012, 421, (2-3), 204-36.

(2) Yang, J. M.; Chin, K. V.; Hait, W. N., Interaction of P-glycoprotein with protein kinase C in human multidrug resistant carcinoma cells. *Cancer Res* 1996, 56, (15), 3490-4.

(3) Hochstrasser, M., Ubiquitin-dependent protein degradation. *Annu Rev Genet* 1996, 30, 405-39.

(4) Newell-Price, J.; Clark, A. J.; King, P., DNA methylation and silencing of gene expression. *Trends Endocrinol Metab* 2000, 11, (4), 142-8.

(5) Graves, J. D.; Krebs, E. G., Protein phosphorylation and signal transduction. *Pharmacol Ther* 1999, 82, (2-3), 111-21. (18) Liang, H. C.; Russell, C.; Mitra, V.; Chung, R.; Hye, A.; Bazenet, C.; Lovestone, S.; Pike, I.; Ward, M. Glycosylation of Human Plasma Clusterin Yields a Novel Candidate Biomarker of Alzheimer's Disease. *J. Proteome. Res.* 2015, *14*, 5063-5076.

(6) Lau, K. S.; Dennis, J. W. N-Glycans in cancer progression. *Glycobiology* 2008, *18*, 750-760.

(7) de Leoz, M. L.; Young, L. J.; An, H. J.; Kronewitter, S. R.; Kim, J.; Miyamoto, S.; Borowsky, A. D.; Chew, H. K.; Lebrilla, C. B. High-mannose glycans are elevated during breast cancer progression. *Mol. Cell. Proteomics* 2011, *10*, M110 002717.

(8) Yang, S.; Mishra, S.; Chen, L.; Zhou, J. Y.; Chan, D. W.; Chatterjee, S.; Zhang, H. Integrated Glycoprotein Immobilization Method for Glycopeptide and Glycan Analysis of Cardiac Hypertrophy. *Anal. Chem.* 2015, *87*, 9671-9678.

(9) Liu, X.; McNally, D. J.; Nothaft, H.; Szymanski, C. M.; Brisson, J. R.; Li, J., Mass spectrometry-based glycomics strategy for exploring N-linked glycosylation in eukaryotes and bacteria. *Anal Chem* 2006, 78, (17), 6081-7.

(10) Hao, L.; Shi, Y.; Thomas, S.; Vezina, C. M.; Bajpai, S.; Ashok, A.; Bieberich, C. J.; Ricke, W. A.; Li, L.; Comprehensive urinary metabolomic characterization of a genetically induced mouse model of prostatic inflammation. *Int J Mass Spectrom* 2018, 434, 185-192.

(11) Hao, L.; Greer, T.; Page, D.; Shi, Y.; Vezina, C. M.; Macoska, J. A.; Marker, P. C.; Bjorling, D. E.; Bushman, W.; Ricke, W. A.; Li, L.; In-Depth Characterization and Validation of Human Urine Metabolomes Reveal Novel Metabolic Signatures of Lower Urinary Tract Symptoms. *Sci Rep* 2016, 6, 30869.

(12) Shi, Y.; Li, Z.; Felder, M.; Yu, Q.; Shi, X.; Peng, Y.; Cao Q.; Wang, B.; Puglielli, L.; Patankar, M.; Li, L.; High-resolution mass spectrometry imaging of N-glycans from formalin-fixed paraffin-embedded tissue sections using a novel sub-atmospheric pressure ionization source. *Anal Chem*. 2019, 91, 20, 12942-12947.

(13) Shi, Y.; Johnson, J.; Wang, B.; Chen, B.; Fisher, G. L.; Urabe, G.; Shi, X.; Kent, K. C.; Guo, L. W.; Li, L.; Mass Spectrometric Imaging Reveals Temporal and Spatial Dynamics of Bioactive Lipids in Arteries Undergoing Restenosis. *J. Proteome Res* 2019, 18, (4), 1669-1678.

(14) Pruijn, G. J., Citrullination and carbamylation in the pathophysiology of rheumatoid arthritis. *Front Immunol* 2015, 6, 192.

(15) Wang, Y.; Li, M.; Stadler, S.; Correll, S.; Li, P.; Wang, D.; Hayama, R.; Leonelli, L.; Han, H.; Grigoryev, S. A.; Allis, C. D.; Coonrod, S. A., Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J Cell Biol* 2009, 184, (2), 205-13.

(16) Ishigami, A.; Ohsawa, T.; Hiratsuka, M.; Taguchi, H.; Kobayashi, S.; Saito, Y.; Murayama, S.; Asaga, H.; Toda, T.; Kimura, N.; Maruyama, N., Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease. *J Neurosci Res* 2005, 80, (1), 120-8.

(17) Ishigami, A.; Masutomi, H.; Handa, S.; Nakamura, M.; Nakaya, S.; Uchida, Y.; Saito, Y.; Murayama, S.; Jang, B.; Jeon, Y. C.; Choi, E. K.; Kim, Y. S.; Kasahara, Y.; Maruyama, N.; Toda, T., Mass spectrometric identification of citrullination sites and immunohistochemical detection of citrullinated glial fibrillary acidic protein in Alzheimer's disease brains. *J Neurosci Res* 2015, 93, (11), 1664-74.

(18) Martinod, K.; Demers, M.; Fuchs, T. A.; Wong, S. L.; Brill, A.; Gallant, M.; Hu, J.; Wang, Y.; Wagner, D. D., Neutrophil histone modification by peptidylarginine deiminase 4 is critical for deep vein thrombosis in mice. *Proc Natl Acad Sci* U S A 2013, 110, (21), 8674-9.

(19) Mastronardi, F. G.; Wood, D. D.; Mei, J.; Raijmakers, R.; Tseveleki, V.; Dosch, H. M.; Probert, L.; Casaccia-Bonnefil, P.; Moscarello, M. A., Increased citrullination of histone H3 in multiple sclerosis brain and animal models of demyelination: a role for tumor necrosis factor-induced peptidylarginine deiminase 4 translocation. *J Neurosci* 2006, 26, (44), 11387-96.

(20) Nicholas, A. P.; King, J. L.; Sambandam, T.; Echols, J. D.; Gupta, K. B.; McInnis, C.; Whitaker, J. N., Immunohistochemical localization of citrullinated proteins in adult rat brain. *J Comp Neurol* 2003, 459, (3), 251-66.

(21) Nicholas, A. P.; Whitaker, J. N., Preparation of a monoclonal antibody to citrullinated epitopes: its characterization and some applications to immunohistochemistry in human brain. *Glia* 2002, 37, (4), 328-36.

(22) Bicker, K. L.; Subramanian, V.; Chumanevich, A. A.; Hofseth, L. J.; Thompson, P. R., Seeing citrulline: development of a phenylglyoxal-based probe to visualize protein citrullination. *J Am Chem Soc* 2012, 134, (41), 17015-8.

(23) Kunieda, K.; Yamauchi, H.; Kawaguchi, M.; Ieda, N.; Nakagawa, H., Development of a fluorescent probe for detection of citrulline based on photo-induced electron transfer. *Bioorg Med Chem Lett* 2018, 28, (5), 969-973.

(24) Shi, J.; Knevel, R.; Suwannalai, P.; van der Linden, M. P.; Janssen, G. M.; van Veelen, P. A.; Levarht, N. E.; van der Helm-van Mil, A. H.; Cerami, A.; Huizinga, T. W.; Toes, R. E.; Trouw, L. A., Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. *Proc Natl Acad Sci* U S A 2011, 108, (42), 17372-7.

(25) Ok, E.; Basnakian, A. G.; Apostolov, E. O.; Barri, Y. M.; Shah, S. V., Carbamylated low-density lipoprotein induces death of endothelial cells: a link to atherosclerosis in patients with kidney disease. *Kidney Int* 2005, 68, (1), 173-8.

(26) Asselbergs, F. W.; Reynolds, W. F.; Cohen-Tervaert, J. W.; Jessurun, G. A.; Tio, R. A., Myeloperoxidase polymorphism related to cardiovascular events in coronary artery disease. *Am J Med* 2004, 116, (6), 429-30.

(27) Baldus, S.; Heitzer, T.; Eiserich, J. P.; Lau, D.; Mollnau, H.; Ortak, M.; Petri, S.; Goldmann, B.; Duchstein, H. J.; Berger, J.; Helmchen, U.; Freeman, B. A.; Meinertz, T.; Munzel, T., Myeloperoxidase enhances nitric oxide catabolism during myocardial ischemia and reperfusion. *Free Radic Biol Med* 2004, 37, (6), 902-11.

(28). Oslowski, C. M.; Urano, F., Measuring ER stress and the unfolded protein response using mammalian tissue culture system. *Methods Enzymol* 2011, 490, 71-92.

(29) Zeng, L.; Liu, Y. P.; Sha, H.; Chen, H.; Qi, L.; Smith, J. A., XBP-1 couples endoplasmic reticulum stress to augmented IFN-beta induction via a cis-acting enhancer in macrophages. *J Immunol* 2010, 185, (4), 2324-30.

(30) Liu, Y. P.; Zeng, L.; Tian, A.; Bomkamp, A.; Rivera, D.; Gutman, D.; Barber, G. N.; Olson, J. K.; Smith, J. A., Endoplasmic reticulum stress regulates the innate immunity critical transcription factor IRF3. *J Immunol* 2012, 189, (9), 4630-9.

(31) Park, D.; Brune, K. A.; Mitra, A.; Marusina, A. I.; Maverakis, E.; Lebrilla, C. B. Characteristic Changes in Cell Surface Glycosylation Accompany Intestinal Epithelial Cell (IEC) Differentiation: High Mannose Structures Dominate the Cell Surface Glycome of Undifferentiated Enterocytes. *Mol. Cell. Proteomics* 2015, *14*, 2910-2921.

(32) Yu, Q.; Wang, B.; Chen, Z.; Urabe, G.; Glover, M. S.; Shi, X.; Guo, L. W.; Kent, K. C.; Li, L. Electron-Transfer/Higher-Energy Collision Dissociation (EThcD)-Enabled Intact Glycopeptide/Glycoproteome Characterization. *J. Am. Soc. Mass Spectrom.* 2017, *28*, 1751-1764.

(33) Sun, S.; Shah, P.; Eshghi, S. T.; Yang, W.; Trikannad, N.; Yang, S.; Chen, L.; Aiyetan, P.; Hoti, N.; Zhang, Z.; Chan, D. W.; Zhang, H. Comprehensive analysis of protein glycosylation by solid-phase extraction of N-linked glycans and glycosite-containing peptides. *Nat. Biotechnol.* 2016, *34*, 84-88.

(34) Chen, Z.; Yu, Q.; Hao, L.; Liu, F.; Johnson, J.; Tian, Z.; Kao, W. J.; Xu, W.; Li, L. Site-specific characterization and quantitation of N-glycopeptides in PKM2 knockout breast cancer cells using DiLeu isobaric tags enabled by electron-transfer/higher-energy collision dissociation (EThcD). *Analyst* 2018, *143*, 2508-2519.

(35) Scott, D. A.; Casadonte, R.; Cardinali, B.; Spruill, L.; Mehta, A. S.; Carli, F.; Simone, N.; Kriegsmann, M.; Del Mastro, L.; Kriegsmann, J.; Drake, R. R. Increases in Tumor N-Glycan Polylactosamines Associated with Advanced HER2-Positive and Triple-Negative Breast Cancer Tissues. *Proteomics Clin Appl* **2019**, *13*, e1800014.

(36) Everest-Dass, A. V.; Briggs, M. T.; Kaur, G.; Oehler, M. K.; Hoffmann, P.; Packer, N. H. N-glycan MALDI Imaging Mass Spectrometry on Formalin-Fixed Paraffin-Embedded Tissue Enables the Delineation of Ovarian Cancer Tissues. *Mol. Cell. Proteomics* **2016**, *15*, 3003-3016.

(37) Holst, S.; Heijs, B.; de Haan, N.; van Zeijl, R. J.; Briaire-de Bruijn, I. H.; van Pelt, G. W.; Mehta, A. S.; Angel, P. M.; Mesker, W. E.; Tollenaar, R. A.; Drake, R. R.; Bovee, J. V.; McDonnell, L. A.; Wuhrer, M. Linkage-Specific in Situ Sialic Acid Derivatization for N-Glycan Mass Spectrometry Imaging of Formalin-Fixed Paraffin-Embedded Tissues. *Anal. Chem.* **2016**, *88*, 5904-5913.

Chapter 2

Development of a Novel Biotin Tag-assisted Mass Spectrometry-based Method for Large-scale Protein Citrullination Analysis

This project is performed in Prof. Lingjun Li lab. Yatao Shi leads the project and is responsible for the biotin tag design and experiment design as well as MS-based analysis. The biotin tag is synthesized by Dr. Zhengqing Ye from Medicinal Chemistry Center (MCC), School of Pharmacy, University of Wisconsin–Madison.

Adapted from: Yatao Shi, Zihui Li, Xudong Shi, Graham Delafield, Zhengqing Ye, Bin Wang, Hui Ye, Zhengwei Chen, Fengfei Ma, Lingjun Li; Development of a novel biotin tag-assisted mass spectrometry-based method for large-scale protein citrullination analysis. (To be submitted)

Abstract

Citrullination is a type of protein post-translational modification (PTM) involved in many serious diseases. Owing to the lack of effective method for citrullination analysis, comprehensive study of this modification is an as-yet-unresolved challenge. Here, we designed a novel biotin thiol tag for the specific modification of ureido group at the side chain of citrulline residue. A novel multifaceted method, which combines chemical derivatization, biotin tag-assisted enrichment strategy with mass spectrometry (MS)-based technology, has been developed for the large-scale identification of citrullinated proteins from complex biological samples. This method has been successfully applied to mouse tissue-specific protein citrullination analysis. In total, 1551 unique citrullinated proteins and 2972 citrullination sites were identified with high confidence. Most of citrullinated proteins and citrullination sites were identified for the first time. Our results indicate that citrullinated proteins are widely distributed in cellular compartments, and are involved in many critical cellular and physiological processes, providing important insights into the role of protein citrullination in diseases.

Introduction

Protein citrullination, catalyzed by a small family of enzymes called peptidylarginine deiminases (PADs), is an important PTM crucial for maintaining protein structures, functions and physiological homeostasis.[1-3] Proteins with aberrant citrullination could serve as autoantigens to stimulate immune system and initiate pathogenic immune responses, contributing to numerous autoimmune diseases, such as rheumatoid arthritis (RA) and multiple sclerosis.[4-6] One instance is the presence of anti-citrullinated protein antibodies (ACPA) in the plasma of majority patients with RA, which has been used in clinic as a biomarker for RA diagnosis.[7] Recently, accumulating evidence revealed the essential roles of protein citrullination in many important cellular processes and diseases.[8-13] Although functionally important, our current understanding about citrullination in terms of its cellular distribution and function is rather limited and primarily impeded by the lack of effective analytical tools.

Current citrullination studies mainly rely on conventional antibody-based techniques, such as immunohistochemistry (IHC) and Western blotting (WB).[14-17] Although effective, these methods suffer from low-throughput and need to know targeted proteins or require specific antibodies, and therefore are not available for global citrullination analysis. Recently, some fluorescence probes are reported to visualize and relatively quantify citrullinated proteins from different biological samples.[18, 19] However, the fluorescence probe-based detection method cannot either identify citrullinated proteins in a large scale, or reveal the dynamic change of individual citrullinated protein in different samples, limiting its application to clinical sample analysis. Moreover, critical information in terms of the changes of citrullination site and citrullinated form of protein was missing. Hence, there is an urgent need to develop alternative strategies, enabling the large-scale protein citrullination analysis of complex biological samples.

In recent years, mass spectrometry (MS)-based bottom-up proteomics has been proven to be a useful tool for large-scale protein PTM analysis. However, its application to protein citrullination analysis suffers from several challenges including: i) Methods enabling specific capture of citrullinated peptides for MS analysis are not effective; ii) The mass shift of 0.984 Da induced by citrullination is the same as that of another modification called deamidation. Furthermore, ^{13}C isotopic peaks in tandem MS spectra have a mass shift of 1.0033 Da, which is close to that of citrullination. These factors will interfere with the accurate annotation of peptide fragment ions upon automated MS data processing, leading to misidentification of citrullination sites.

To deal with these challenges, various strategies have been employed for MS-based citrullination analysis. Chemical derivatization of citrullinated peptides enlarges the mass shift induced by citrullination, avoiding the interference of deamidation and ^{13}C isotopic peaks.[20-22] Furthermore, citrullinated peptides can also be distinguished from their corresponding unmodified peptides according to either their retention time difference on C_{18} column,[23-24] or the characteristic neutral loss of isocyanic acid upon MS fragmentation.[25] However, the application of these approaches to complex sample analysis is limited due to the lack of enrichment capability. Recently, a biotin-PEG_2-4-glyoxalbenzoic acid (BPG) is developed for the simultaneous derivatization and enrichment of citrullinated peptides from biological samples for MS analysis.[26, 27] However, this method resulted in low identification rate of citrullinated peptides due to the poor fragmentation efficiency, and therefore is not amenable to MS analysis. To date, there has been no effective method for MS-based in-depth analysis of citrullinated peptides in complex biological samples.

In this study, a novel biotin thiol tag is designed for the simultaneous chemical derivatization and biotinylation of citrullinated peptides. Furthermore, we successfully develop a novel method, which integrates the biotin tag labeling strategy with mass spectrometry (MS)-based technology, enabling the large-scale identification of citrullinated proteins from biological samples and high-confident annotation of citrullination site. In the end, the new method is applied to mouse-tissue-specific protein citrullination analysis, exploring tissue-specific distribution patterns and biological functions of protein citrullination for the first time.

Materials and Methods

1. Materials

Methanol (MeOH), ethanol (EtOH), acetonitrile (ACN), formic acid (FA), trifluoroacetic acid (TFA), ammonium bicarbonate and Tris base were purchased from Fisher Scientific (Pittsburgh, PA). Trypsin gold, Lys-C and Trypsin/Lys-C Mix were purchased from Promega (Madison, WI). Tris (2-carboxyethyl) phosphine (TCEP) was purchased from EMD Millipore Sigma (Burlington, MA). Biotin-NHS Ester was purchased from Click Chemistry Tools (Scottsdale, AZ). Sodium dodecyl sulfate (SDS), cysteamine and streptavidin agarose were purchased from Sigma-Aldrich (St. Louis, MO). Phosphate-Buffered Saline (PBS) buffer was ordered from Mediatech, Inc (Manassas, VA). Distilled water mentioned in this work was Milli-Q water from a Millipore filtration system (Bedford, MA). All reagents were used without additional purification.

2. Synthesis of biotin thiol tag

A solution of biotin-NHS ester (100 mg, 0.29 mmol) and cysteamine (34 mg, 0.44 mmol, 1.5 equiv) in CH_2Cl_2 (5 mL) was added to N, N-Diisopropylethylamine (144 µL, 0.88 mmol. 3 equiv) and stirred at 40 °C for 24h. The crude was purified using a CombiFlash system and a

gradient from 0 to 20% of B, where A was dichloromethane and B was methanol. Fractions containing pure product (as detected by UV) were collected to yield the grey solid (60 mg, 0.20 mmol; 68% yield). ^{1}H NMR data was obtained from a Varian Inova 500 MHz NMR spectrometer. ^{13}C NMR data was obtained by a Bruker Avance III HD 400 MHz NMR spectrometer. The spectra were recorded in CD_3OD solution (10 mg cm^{-3}) with a probe temperature of ca. 300K and referenced to TMS. ^{1}H NMR (500 MHz, CD_3OD) δ 8.17 (m, ^{1}H), 4.56 (dd, J = 7.7, 5.0 Hz, ^{1}H), 4.37 (dd, J = 7.8, 4.5 Hz, ^{1}H), 3.57-3.53 (m, ^{1}H), 3.42-3.40 (m, ^{2}H), 3.28 (dt, J = 9.9, 5.3 Hz, ^{1}H), 3.00 (dd, J = 12.7, 5.0 Hz, ^{1}H), 2.89 (q, J = 6.4 Hz, ^{1}H), 2.77 (d, J = 12.7 Hz, ^{1}H), 2.67 (t, J = 6.8 Hz, ^{2}H), 2.29 (t, J = 7.4 Hz, ^{2}H), 1.85-1.63 (m, ^{4}H), 1.55-1.49 (m, ^{2}H). ^{13}C NMR (101 MHz, CD_3OD) δ 174.8, 163.2, 62.0, 60.2, 55.6, 42.5, 39.6, 35.3, 28.4, 28.1, 25.4, 23.1. Formula: $C_{12}H_{22}N_3O_2S_2$; $[M+H]^+$: m/z 304.1153 Da.

3. Method development with citrullinated peptide standard

3.1. Chemical derivatization of citrullinated peptide standard

A citrullinated peptide standard SAVRA{Cit}SSVPGVR was obtained from Genscript (New Jersey, USA), and dissolved in water to a concentration of 1 mg/ml. 2,3-butanedione solution was freshly prepared by transferring 1µl of 2,3-butanedione (11038, Sigma-Aldrich) into 114 µl 12.5% Trifluoroacetic acid (TFA) solution. 30µl biotin thiol tag solution (10 mg/ml) was dried with SpeedVac, and then reconstituted with 40µl 12.5% TFA. 1µl citrullinated peptide standard and 10 µl diluted 2,3-butanedione solution was subsequently added to initiate the chemical derivatization reaction. The mixture was incubated in the dark at 37°C with shaking. After 6 h, the reaction was stopped by drying the mixture out with SpeedVac.

To remove excess reactants, strong cation exchange (SCX) chromatography was performed using TopTips (TT200SEA, Poly LC) containing PolySULFOETHYL A beads. Briefly,

SCX beads were pre-equilibrated with 100µl loading buffer containing 50% ACN/0.2% FA/10mM ammonium formate for three times. The dried peptide mixture mentioned above was then reconstituted in 200µl loading buffer and applied to the SCX beads twice followed by continuous washing of the beads with 100µl loading buffer for 10 times. The derivatized citrullinated peptide standard was eluted by rinsing the beads with 50µl 25% ACN/0.4 M ammonium formate for 3 times. The eluent was finally dried in the SpeedVac and stored for further use. All centrifugation steps were performed at 2,000 rpm for 2 min.

3.2. Enrichment and release of biotin tag-modified citrullinated peptide standard using streptavidin beads

The derivatized citrullinated peptide standard prepared above was firstly reconstituted in 50µl 50% ACN/H_2O. One micro liter was added into 50µl H_2O containing 400 µg tryptic peptides digested from mouse brain. After brief vortex and centrifugation, 1 µl peptide solution was spotted onto stainless steel plate for MALDI analysis.

The enrichment process was performed as previously described with some modifications.[28] Briefly, 75µl streptavidin beads were pre-washed with 1ml PBS 1× buffer for 5 times. The leftover peptide solution was diluted with 950 µl PBS 1× buffer, which was further loaded onto the pre-washed streptavidin agarose and incubated at room temperature for 2 h with rotation. The beads were subsequently washed 4 times each with 1ml PBS 1× buffer, 1ml 5 % ACN/PBS 1× buffer and 1ml water. The bound peptides were finally released with 300µl 80%ACN/H_2O contain 0.2% TFA and 0.1% FA for four times. The first release was performed in room temperature for 5min, while other three release processes were conducted at 95°C for 5min with shaking. The eluents were combined and dried in the SpeedVac. Enriched peptides were desalted with C_{18} columns (Zip-tips) and eluted in 100 µl 80% ACN/0.2% FA solution, which was completely dried in the

vacuum. The enriched peptides were further reconstituted in 50µl 50% ACN/H_2O, and 1 µl solution was spotted onto stainless plate for MALDI analysis.

3.3. MS-based fragmentation patterns of biotin tag-modified citrullinated peptide standard

The derivatized citrullinated peptide standard was reconstituted in 0.1% FA, 50% ACN/H_2O and directly injected into an Orbitrap Fusion Lumos Tribrid mass spectrometer. Full MS scan was performed from *m/z* 300 to *m/z* 1500 using a resolution of 60,000 and RF lens of 30%. AGC target was set to $2e^5$ and the maximum injection time was 100ms. The precursor ion of biotin tag-modified citrullinated peptide standard was isolated for HCD, ETD and EThcD fragmentation, respectively. Different normalized collisional energies (NCEs) were tested upon HCD fragmentation, while ETD fragmentation was performed using various ETD reaction time. For EThcD conditions, precursor ions were repeatedly selected and fragmented with different EThcD reaction times or normalized collision energies (NCE), respectively. Tandem MS spectra of biotin tag-modified citrullinated peptide standard fragmented by three fragmentation techniques were collected for further analysis. Fragment ions in tandem MS spectra were manually annotated based on their accurate mass.

4. Method optimization for global protein citrullination analysis on biological samples

4.1. Enzymatic protein digestion

A freshly sacrificed mouse brain was homogenized with a probe ultrasonicator in 4% Sodium dodecyl sulfate (SDS)/50mM Tris base buffer (~pH 8, adjusted with HCl). After homogenization, the supernatant was collected after centrifugation at 16,000 rcf for 15 mins. Protein concentration was determined by BCA protein assay reagent (Thermo Scientific, Fair Lawn, NJ). Each 800µg mouse protein extract was reduced by 10mM dithiothreitol (DTT) for 30 min at room temperature and then alkylated with 50mM iodoacetic acid (IAA) for another 30 min

in the dark. Protein was further precipitated by adding cold acetone (-20°C) to a final concentration of 80% (v/v). After incubation in -20°C overnight, the sample solution was centrifuged at 16,000 rcf for 15 min and the supernatant was discarded. To completely remove SDS, the pellet was washed with cold 80% acetone for another two times, and then dried in the air for 12 min. The dried pellet was reconstituted with 150µl 5 M guanidine hydrochloride/50mM Tris-base buffer (pH 8, adjusted with HCl). After reconstitution, the buffer solution was diluted tenfold with 50mM Tris-base buffer to reduce the concentration of guanidine HCl to 0.5 M.

Reconstituted protein was further digested by three different enzymes (Promega, Madison, WI), including Lys-C, Trypsin gold and Trypsin/Lys-C Mix, respectively. Enzyme (Protein:enzyme, 100:1, w/w) was added to sample and incubated at 37°C for 6 h. The secondary digestion was performed by adding the same amount of enzyme and incubating samples at 37°C for another 12 h. The digestion process was quenched by adding 10% TFA to reduce the pH to <3. Digested peptides were desalted with C_{18} columns (Sep-Pak, waters, Milford, MA) and eluted with 80% ACN/0.2% FA, which were dried out with SpeedVac. Peptide concentration was determined by colorimetric peptide assay (23275, Thermo Scientific, Fair Lawn, NJ). Each 400µg of peptide was transferred to an Eppendorf tube and dried for further analysis.

4.2. Chemical derivatization

Thirty micro liter of biotin thiol tag (10 mg/ml) solution was added to each sample tube containing 400µg peptide and dried with SpeedVac. Then, the biotin thiol tag and peptide were reconstituted in 40µl 12.5% TFA solution followed by adding 10µl of 2,3-butanedione solution prepared as mentioned above. The reactants were incubated at 37°C for 6h with shaking in the dark and were further dried out with SpeedVac. To remove leftover reactants, SCX chromatography was performed as mentioned above. Peptides were finally eluted with 150 µl 25%

ACN/0.4M ammonium formate and dried out in SpeedVac. A secondary drying process was conducted to reduce the concentration of ammonium formate in samples by adding another 400µl water to each sample and drying out in SpeedVac.

4.3. Enrichment of biotinylated citrullinated peptides with streptavidin beads

To enable complete dissolution of peptides recovered from SCX, 300µl 50% ACN/H_2O was added to each sample. After brief vortex and centrifugation, each sample was dried in SpeedVac to less than 100µl and further diluted with 900µl PBS 1× buffer. The enrichment process was performed as previously described. Enriched peptides were desalted with C_{18} columns (Zip-tips,) and eluted in 100 µl 80% ACN/0.2% FA solution, which was finally dried in the vacuum to almost dry.

4.4. Mass spectrometry (MS) analysis

MS analysis was performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer coupled with Dionex UltiMate 3000 UPLC system (Thermo Fisher Scientific, San Jose, CA). The mobile phase was composed of 0.1% formic acid in water (A) and 0.1% formic acid in ACN (B). The flow rate was set as 300 nl/min. Each sample was resuspended in 15 µl 0.1% formic acid 3% ACN in water and underwent three LC-MS/MS run. For each run, 2µl of sample was loaded onto a customized C18 column filled with 1.7µm Ethylene Bridged Hybrid packing materials (130 Å, Waters) and separated with the following gradient: 3% B for the first 18.3 min upon sample trapping; 3%-30% B for 18.3-120 min; 30%-75% B for 120-120.5 min; 75% B for 120.5-130 min; 75%-95% B for 130.0-130.5 min; 95% B for 130.5-140.0 min; 95%-3% B for 140.0-140.5 min and equilibrated at 3% B for 15 min. MS data was acquired in the positive ion mode with the spray voltage of 2 kV. MS^1 spectra were collected from m/z 350-1500 by orbitrap at a resolution of 60,000. Automatic gain control (AGC) target and maximum injection time was set to $2e^5$ and 100

ms, respectively. Ions with the charge of 2-6 were included for screening, while a dynamic exclusion time of 45s was used to avoid repeatedly sequencing the same precursor ion in a short time.

Three fragmentation methods were tested, including stepped HCD fragmentation, HCD product ion-triggered ETD (HCD-pd-ETD) and HCD product ion-triggered EThcD (HCD-pd-EThcD). For stepped HCD fragmentation, the top 15 precursors were fragmented with an isolation window of 1 Da, a resolving power of 15K and AGC target of 5×10^4. For HCD-pd-ETD and HCD-pd-EThcD methods, m/z 227.0848 and 304.1147 were listed as HCD product ions to trigger the subsequent ETD or EThcD fragmentation.

5. Mouse-tissue specific protein citrullination analysis

Mouse tissues from three mice were collected and stored at -80°C freezer until use. Tissues were homogenized with a probe ultrasonicator in 4% sodium dodecyl sulfate (SDS)/50mM Tris base buffer (~pH 8, adjusted with HCl). Protein extract was digested by Trypsin/Lys-C Mix and desalted following the procedure previously described. For each tissue, 400µg of peptide was used for global protein citrullination analysis. The chemical derivatization, SCX cleaning and streptavidin enrichment processes were performed as previously described. Enriched citrullinated peptides from each tissue was finally detected by MS using stepped HCD fragmentation technique.

6. Database search

MS data from mouse tissues was searched against mouse protein database (Downloaded from Uniprot website on 12/30/2018) using the software of MaxQuant. The search parameters were defined as follows: 20 ppm and 4.5 ppm were set as the first search peptide tolerance and the main search peptide tolerance, respectively; 2 ppm was used for the isotope match tolerance. a minimum of six amino-acid peptide length and up to three missed cleavages were allowed for

peptide identification; carbamidomethylation (C) was set as a static modification, while oxidation (M,+15.995 Da) and biotin tag-labeled citrullination (R, +354.1071 Da) were defined as dynamic modifications. A neutral loss of biotin tag (303.1075 Da) and two diagnostic ions of 227.0848 Da and 304.1147 Da were included in the search. A maximum of three modifications were allowed for each peptide. The 1% false-discovery rate (FDR) was applied to filter both peptide and protein identification. Peptides that were found as reverse or potential contaminant hits were filtered out and citrullination site localization probability threshold was set to 0.75. Gene ontology (GO) enrichment analysis of identified citrullinated proteins was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8.

Results and discussion

Synthesis and protection of biotin thiol tag

The novel biotin thiol tag was easy to be synthesized in just one step using biotin-NHS ester and cysteamine (**Figure 1**). However, we observed the dimerization of synthesized biotin thiol tag upon purification and drying processes, and more biotin tag dimers generated during long-term storage (**Figure S1**). To deal with this issue, tris(2-carboxyethyl)phosphine (TCEP) was used to prevent the biotin thiol tag from oxidation. Different with other reducing agents, such as DTT and 2-mercaptoethanol, TCEP does not has a free thiol group and thereby will not interfere with the biotin tag labeling reaction. To test the protection effect of TCEP, the biotin tag powder was firstly reconstituted with methanol/water (50:50, v/v) to a concentration of 10 mg/ml. Then, neutral TCEP was added into the biotin tag solution to a final concentration of 10mM for the reduction of biotin thiol tag dimer. Following that, the biotin tag solution was briefly sonicated and placed at room temperature until the solution was clear. MS analysis of the clear solution suggested that

most of the biotin tag dimer was reduced into biotin thiol tag, which could be stored at -80°C for long-term use (**Figure S2**).

Citrullinated peptide standard-assisted method development

We noticed that free thiol group, together with 2,3-butanedione, could specifically modify the ureido group at the side chain of peptidylcitrulline in low pH aqueous solution. Thus, we designed a bifunctional biotin thiol tag for the concurrent chemical derivatization and biotinylation of citrullinated peptide (**Figure 2**). The chemical derivatization efficiency of the novel biotin tag was first examined by using the citrullinated peptide standard (SAVRACitSSVPGVR). We noticed that the citrullinated peptide standard could be efficiently derivatized after 6 h reaction, which was shown in **Figure 3**.

Next, we tested whether the biotin tag-labeled citrullinated peptide standard can be enriched from peptide mixture. A small amount of derivatized citrullinated peptide standard was spiked into 400μg tryptic peptides digested from mouse brain, and was further enriched with streptavidin beads. After enrichment, as shown in **Figure 4**, the derivatized citrullinated peptide standard could be successfully enriched and released from streptavidin beads though no cleavable group was present in the spacer arm of the biotin tag.

We further investigated the mass spectrometric fragmentation patterns of the biotin tag-modified citrullinated peptide standard using higher-energy collisional dissociation (HCD), electron-transfer dissociation (ETD) and electron-transfer/higher-energy collision dissociation (EThcD) fragmentation techniques, respectively. HCD fragmentation yielded a series of b and y peptide fragment ions including some ions resulting from the neutral loss of biotin tag group. Remarkably, two dominant ions at m/z 227 and 304 were observed, which were generated from the biotin head group and the cleavage of C-S bond linking biotin tag with imidazolone moiety,

respectively (**Fig. S3**). ETD fragmentation, in contrast to HCD fragmentation, did not generate ions at m/z 227 and 304, but numerous c and z fragment ions were observed, especially these fragment ions with intact modification group crucial for citrullination site assignment (**Fig. S4**). EThcD fragmentation, compared to HCD and ETD fragmentation, produced two diagnostic ions at m/z 227 and 304, b and y ions as well as c and z ions in one single spectrum (**Fig. S5**). These results suggested that the biotin tag-modified citrullinated peptide could generate high-quality tandem MS spectrum for citrullinated peptide identification and citrullination site assignment.

Enzymes and MS detection method for global protein citrullination analysis

Following the method development with the citrullinated peptide standard, a standard workflow was established for the large-scale analysis of citrullinated proteins in complex biological samples (**Figure 5**). We firstly tested the performance of our method using mouse brain, which has been reported to show high level of citrullination modification. Three different MS fragmentation and detection strategies were compared, including stepped-HCD, HCD product ion-triggered ETD (HCD-pd-ETD) and HCD product ion-triggered EThcD (HCD-pd-EThcD). For those two product ion-triggered methods, the diagnostic ions of biotin head group (m/z 227 and 304) were used as the HCD product ions. We found that all methods were suited for citrullination analysis though stepped-HCD fragmentation method identified the maximum numbers of citrullination sites primarily due to the shorter duty cycle time of HCD (**Fig. 6**). Actually, most citrullination sites identified by HCD-pd-ETD and HCD-pd-EThcD were also identified by HCD fragmentation. Hence, we chose stepped HCD fragmentation for the large-scale protein citrullination analysis.

Furthermore, we noticed that a large number of citrullinated peptides with C-terminal citrulline residue were identified with high confidence, which is in contrast with a common

understanding about protein citrullination that trypsin cannot cleave after the citrulline residue. To investigate this contradictory result, we applied our method to analyze three peptide samples digested from the same mouse brain protein extract by trypsin, Lys C and trypsin/Lys C mixture, respectively (**Fig. 7a-b**). We found that different citrullinated peptides containing the same citrullination site were identified from the Lys C-digested sample and trypsin or Lys C/trypsin-digested sample (**Fig. S6**). Meanwhile, more than 50% of the citrullinated peptides identified from the sample treated by trypsin or Lys C/trypsin mixture, had the citrulline residue at the C-terminus, while few citrullinated peptides were identified with C-terminal citrulline residue from the Lys C-digested sample (**Fig. 7c-d**). These results suggested that some citrullination sites were trypsin-cleavable *in vitro* though the mechanism is still uncertain.

Mouse tissue-specific protein citrullination analysis

Next, we applied our method to mouse tissue-specific protein citrullination analysis. In total, 1551 unique citrullinated proteins and 2972 citrullination sites were identified from mouse brain and organs with high confidence (**Fig. 8a**). Citrullination motif analysis revealed that there was no conserved amino acid sequence flanking citrullination sites (**Fig. 8b**). Furthermore, we found that protein citrullination showed distinct tissue-specific distribution patterns. More citrullinated proteins were identified from brain in comparison to other organs. However, as shown in **Figure 8c**, no apparent distribution pattern was observed from multiple brain regions, while poor overlap of citrullinated proteins between organs was noticed. Gene ontology (GO) cellular component analysis indicated that citrullinated proteins are mainly present in cytoplasm, extracellular exosomes and membranes as well as mitochondrion, while biological process analysis showed the involvement of citrullinated proteins in diverse cellular processes, such as oxidation-reduction process, translation and transport (**Figure S7**).

With the improved method reported here, we detected not only previously known citrullination sites, but also new citrullination sites from the reported citrullinated proteins. Taking myelin basic protein (MBP) as an example, only 4 citrulline residues are asserted by similarity in the Uniprot database. However, we confidently identified 10 citrullination sites on MBP from mouse brain tissues. Besides arginine 157 and 163, 8 new citrullination sites were assigned with their high-quality tandem MS spectra (**Fig. 9**). Another interesting citrullinated protein is the glial fibrillary acidic protein (GFAP), which is considered as an astrocyte-specific protein marker and involved in the astrocyte-neuron interactions.[29] Our method pinpointed 23 citrullination sites in GFAP with high confidence, most of which have not been reported previously (**Figure S8**).

More importantly, our method enabled the identification of many important citrullinated proteins for the first time. For example, apolipoprotein E (ApoE) is an important apolipoprotein associated with lipid metabolism and *Aβ* metabolism in the brain. ApoE4, one type of ApoE isoforms encoded by ε4 allele of the APOE gene, is considered as a strong risk factor of Alzheimer's disease (AD). In this study, we identified two new citrullination sites located at R214 and R236 of protein apolipoprotein E from mouse brain for the first time (**Fig. 10**). Another important citrullinated protein we found in the mouse brain was microtubule-associated protein tau (Tau). It is well-known that hyperphosphorylation of the tau protein contributed to the pathogenesis of AD. Here, we detected two citrullination modifications at its R534 and R698, which have not been reported to date (**Fig. 11**). Besides those proteins, two citrullination sites at R57 and R153 were detected in NAD-dependent protein deacetylase sirtuin-2, which functions as an essential enzyme for the deacetylation of histones, tubulin and many transcription factors, and therefore plays a critical role in many biological processes (**Figure S9**).[30, 31]

We further observed that citrullination modification was colocalized with some arginine methylation modifications at some arginine residues, especially the omega-n-methylarginine modification (**Fig. 12a**). 36 identified citrullination sites were overlapped with the omega-n-methylarginine sites in UniProt database, such as the R175 and R181 at MBP, R11 and R20 at GFAP. Additionally, we observed that 4 types of arginine modifications occurred at the R203 and R213 of the heterogeneous nuclear ribonucleoproteins A2/B1(Hnrnpa2b1), suggesting the presence of potential crosstalk between these arginine PTMs (**Fig. 12b**).

Conclusions

In summary, we report a novel biotin cysteamine tag for the specific chemical derivatization of peptidylcitrulline. By integrating the biotin tag-assisted chemical derivatization strategy, streptavidin-aided enrichment technique with bottom-up proteomics, we developed a novel biotin tag-assisted MS-based method, which, to our knowledge, is the first method enabling large-scale identification of citrullinated proteins in complex biological samples. We expect that this method would provide the scientific community a simple, powerful and reliable tool to unravel the roles of citrullination in biological and disease pathways, significantly expanding our understanding of this important protein PTM.

Acknowledgements

This research is supported in part by the National Institutes of Health through grants-R21 AG060242, U01CA231081, RF1AG052324, and R01 DK071801. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531 to LL) and the University of Wisconsin-Madison, Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation.

References

(1) Tarcsa, E.; Marekov, L. N.; Mei, G.; Melino, G.; Lee, S. C.; Steinert, P. M., Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin. *J Biol Chem* 1996, 271, (48), 30709-16.

(2) Denman, R. B., PAD: the smoking gun behind arginine methylation signaling? *Bioessays* 2005, 27, (3), 242-6.

(3) Ishida-Yamamoto, A.; Senshu, T.; Eady, R. A.; Takahashi, H.; Shimizu, H.; Akiyama, M.; Iizuka, H., Sequential reorganization of cornified cell keratin filaments involving filaggrin-mediated compaction and keratin 1 deimination. *J Invest Dermatol* 2002, 118, (2), 282-7.

(4) Muller, S.; Radic, M., Citrullinated Autoantigens: From Diagnostic Markers to Pathogenetic Mechanisms. *Clin Rev Allergy Immunol* 2015, 49, (2), 232-9.

(5) Bradford, C. M.; Ramos, I.; Cross, A. K.; Haddock, G.; McQuaid, S.; Nicholas, A. P.; Woodroofe, M. N., Localisation of citrullinated proteins in normal appearing white matter and lesions in the central nervous system in multiple sclerosis. *J Neuroimmunol* 2014, 273, (1-2), 85-95.

(6) Haag, S.; Schneider, N.; Mason, D. E.; Tuncel, J.; Andersson, I. E.; Peters, E. C.; Burkhardt, H.; Holmdahl, R., Identification of new citrulline-specific autoantibodies, which bind to human arthritic cartilage, by mass spectrometric analysis of citrullinated type II collagen. *Arthritis Rheumatol* 2014, 66, (6), 1440-9.

(7) Sebbag, M.; Chapuy-Regaud, S.; Auger, I.; Petit-Texeira, E.; Clavel, C.; Nogueira, L.; Vincent, C.; Cornelis, F.; Roudier, J.; Serre, G., Clinical and pathophysiological significance of the autoimmune response to citrullinated proteins in rheumatoid arthritis. *Joint Bone Spine* 2004, 71, (6), 493-502.

(8) Pruijn, G. J., Citrullination and carbamylation in the pathophysiology of rheumatoid arthritis. *Front Immunol* 2015, 6, 192.

(9) Wang, Y.; Li, M.; Stadler, S.; Correll, S.; Li, P.; Wang, D.; Hayama, R.; Leonelli, L.; Han, H.; Grigoryev, S. A.; Allis, C. D.; Coonrod, S. A., Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J Cell Biol* 2009, 184, (2), 205-13.

(10) Ishigami, A.; Ohsawa, T.; Hiratsuka, M.; Taguchi, H.; Kobayashi, S.; Saito, Y.; Murayama, S.; Asaga, H.; Toda, T.; Kimura, N.; Maruyama, N., Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease. *J Neurosci Res* 2005, 80, (1), 120-8.

(11) Ishigami, A.; Masutomi, H.; Handa, S.; Nakamura, M.; Nakaya, S.; Uchida, Y.; Saito, Y.; Murayama, S.; Jang, B.; Jeon, Y. C.; Choi, E. K.; Kim, Y. S.; Kasahara, Y.; Maruyama, N.; Toda, T., Mass spectrometric identification of citrullination sites and immunohistochemical detection of citrullinated glial fibrillary acidic protein in Alzheimer's disease brains. *J Neurosci Res* 2015, 93, (11), 1664-74.

(12) Martinod, K.; Demers, M.; Fuchs, T. A.; Wong, S. L.; Brill, A.; Gallant, M.; Hu, J.; Wang, Y.; Wagner, D. D., Neutrophil histone modification by peptidylarginine deiminase 4 is critical for deep vein thrombosis in mice. *Proc Natl Acad Sci* U S A 2013, 110, (21), 8674-9.

(13) Mastronardi, F. G.; Wood, D. D.; Mei, J.; Raijmakers, R.; Tseveleki, V.; Dosch, H. M.; Probert, L.; Casaccia-Bonnefil, P.; Moscarello, M. A., Increased citrullination of histone H3 in multiple sclerosis brain and animal models of demyelination: a role for tumor necrosis factor-induced peptidylarginine deiminase 4 translocation. *J Neurosci* 2006, 26, (44), 11387-96.

(14) Nicholas, A. P., Dual immunofluorescence study of citrullinated proteins in Alzheimer diseased frontal cortex. *Neurosci Lett* 2013, 545, 107-11.

(15) Nicholas, A. P.; King, J. L.; Sambandam, T.; Echols, J. D.; Gupta, K. B.; McInnis, C.; Whitaker, J. N., Immunohistochemical localization of citrullinated proteins in adult rat brain. *J Comp Neurol* 2003, 459, (3), 251-66.

(16) Nicholas, A. P.; Whitaker, J. N., Preparation of a monoclonal antibody to citrullinated epitopes: its characterization and some applications to immunohistochemistry in human brain. *Glia* 2002, 37, (4), 328-36.

(17) Raats, J. M.; Wijnen, E. M.; Pruijn, G. J.; van den Hoogen, F. H.; van Venrooij, W. J., Recombinant human monoclonal autoantibodies specific for citrulline-containing peptides from phage display libraries derived from patients with rheumatoid arthritis. *J Rheumatol* 2003, 30, (8), 1696-711.

(18) Bicker, K. L.; Subramanian, V.; Chumanevich, A. A.; Hofseth, L. J.; Thompson, P. R., Seeing citrulline: development of a phenylglyoxal-based probe to visualize protein citrullination. *J Am Chem Soc* 2012, 134, (41), 17015-8.

(19) Kunieda, K.; Yamauchi, H.; Kawaguchi, M.; Ieda, N.; Nakagawa, H., Development of a fluorescent probe for detection of citrulline based on photo-induced electron transfer. *Bioorg Med Chem Lett* 2018, 28, (5), 969-973.

(20) De Ceuleneer, M.; De Wit, V.; Van Steendam, K.; Van Nieuwerburgh, F.; Tilleman, K.; Deforce, D., Modification of citrulline residues with 2,3-butanedione facilitates their detection by liquid chromatography/mass spectrometry. *Rapid Commun Mass Spectrom* 2011, 25, (11), 1536-42.

(21) Holm, A.; Rise, F.; Sessler, N.; Sollid, L. M.; Undheim, K.; Fleckenstein, B., Specific modification of peptide-bound citrulline residues. *Anal Biochem* 2006, 352, (1), 68-76.

(22) Choi, M.; Song, J. S.; Kim, H. J.; Cha, S.; Lee, E. Y., Matrix-assisted laser desorption ionization-time of flight mass spectrometry identification of peptide citrullination site using Br signature. *Anal Biochem* 2013, 437, (1), 62-7.

(23) Fert-Bober, J.; Venkatraman, V.; Hunter, C. L.; Liu, R.; Crowgey, E. L.; Pandey, R.; Holewinski, R. J.; Stotland, A.; Berman, B. P.; Van Eyk, J. E., Mapping Citrullinated Sites in Multiple Organs of Mice Using Hypercitrullinated Library. *J Proteome Res* 2019, 18, (5), 2270-2278.

(24) Raijmakers, R.; van Beers, J. J.; El-Azzouny, M.; Visser, N. F.; Bozic, B.; Pruijn, G. J.; Heck, A. J., Elevated levels of fibrinogen-derived endogenous citrullinated peptides in synovial fluid of rheumatoid arthritis patients. *Arthritis Res Ther* 2012, 14, (3), R114.

(25) Jin, Z.; Fu, Z.; Yang, J.; Troncosco, J.; Everett, A. D.; Van Eyk, J. E., Identification and characterization of citrulline-modified brain proteins by combining HCD and CID fragmentation. *Proteomics* 2013, 13, (17), 2682-91.

(26) Tutturen, A. E.; Holm, A.; Fleckenstein, B., Specific biotinylation and sensitive enrichment of citrullinated peptides. *Anal Bioanal Chem* 2013, 405, (29), 9321-31.

(27) Tutturen, A. E.; Fleckenstein, B.; de Souza, G. A., Assessing the citrullinome in rheumatoid arthritis synovial fluid with and without enrichment of citrullinated peptides. *J Proteome Res* 2014, 13, (6), 2867-73.

(28) Schiapparelli, L. M.; McClatchy, D. B.; Liu, H. H.; Sharma, P.; Yates, J. R., 3rd; Cline, H. T., Direct detection of biotinylated proteins by mass spectrometry. *J Proteome Res* 2014, 13, (9), 3966-78.

(29) Eng, L. F., Glial fibrillary acidic protein (GFAP): the major protein of glial intermediate filaments in differentiated astrocytes. *J Neuroimmunol* 1985, 8, (4-6), 203-14.

(30) Liu, G.; Park, S. H.; Imbesi, M.; Nathan, W. J.; Zou, X.; Zhu, Y.; Jiang, H.; Parisiadou, L.; Gius, D., Loss of NAD-Dependent Protein Deacetylase Sirtuin-2 Alters Mitochondrial Protein Acetylation and Dysregulates Mitophagy. *Antioxid Redox Signal* 2017, 26, (15), 849-863.

(31) Pais, T. F.; Szego, E. M.; Marques, O.; Miller-Fleming, L.; Antas, P.; Guerreiro, P.; de Oliveira, R. M.; Kasapoglu, B.; Outeiro, T. F., The NAD-dependent deacetylase sirtuin 2 is a suppressor of microglial activation and brain inflammation. *EMBO J* 2013, 32, (19), 2603-16.

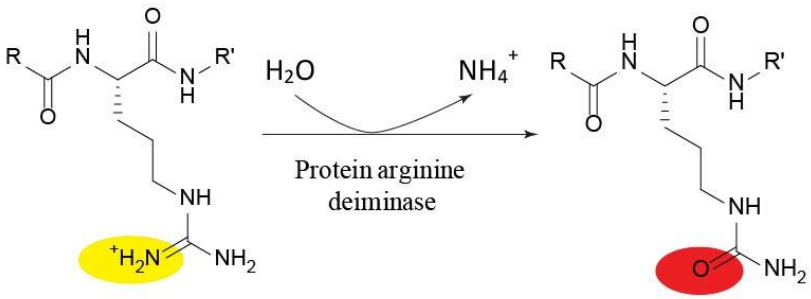


Figure 1. Protein citrullination catalyzed by PAD enzymes. PAD: protein arginine deiminase.

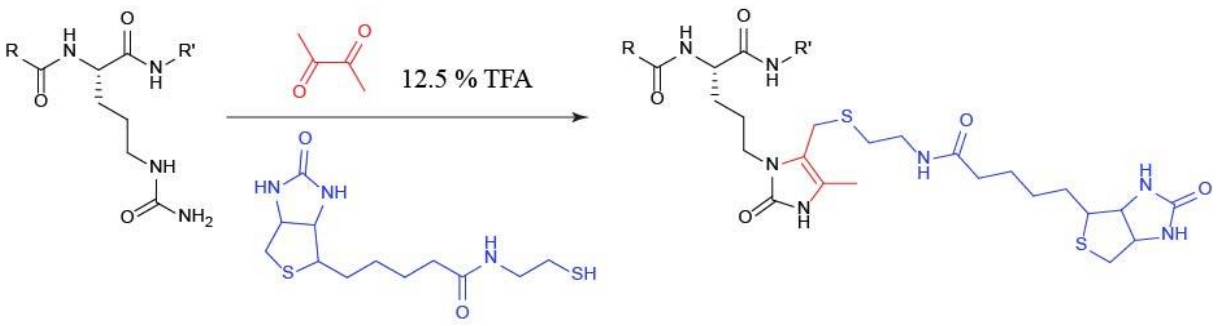


Figure 2. Chemical derivatization of citrullinated peptides by biotin thiol tag and 2,3-butanedione in low aqueous solution.

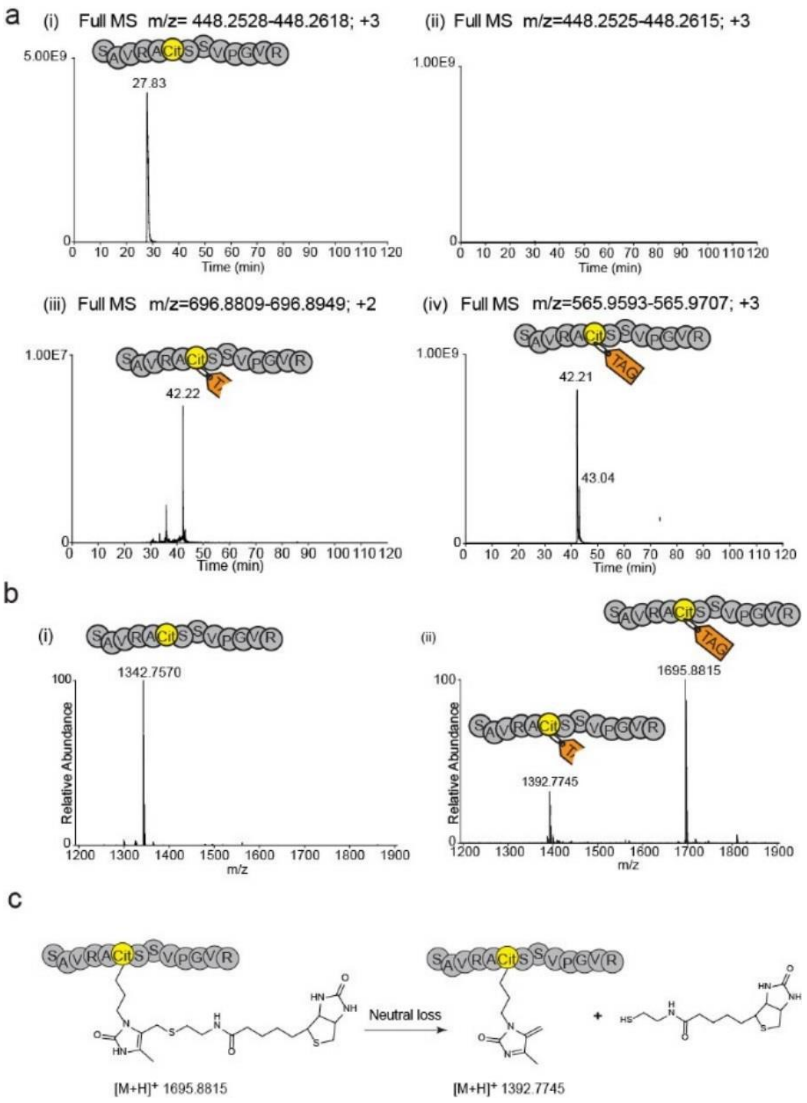


Figure 3. Biotin tag-based chemical derivatization of citrullinated peptide standard. (a) Test the derivatization reaction with ESI mass spectrometry. Standard citrullinated peptide before derivatization (i); Citrullinated peptide standard was completely modified after 6h derivatization (ii & iv); Peak at m/z 1392.7745 could either be generated by the reaction of 2, 3-butanedione with citrullinated peptide standard (less than 1%), or by the neutral loss from biotin tag modified citrullinated peptide standard (about 1%) (iii). (b) Test the derivatization reaction with MALDI orbitrap mass spectrometry. (c) Peak at m/z 1392.7735 resulted from the neutral loss of peak 1695.8798 upon MALDI analysis.

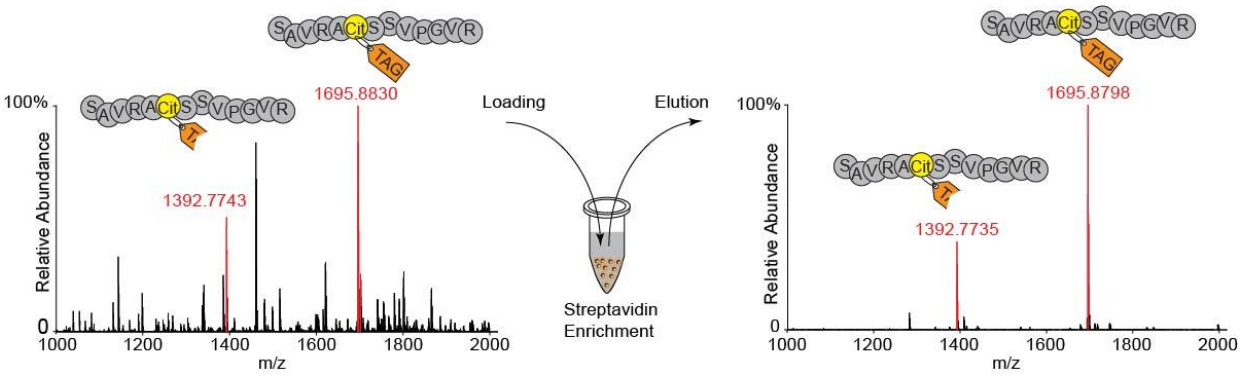


Figure 4. Enrichment of biotin tag-labeled citrullinated peptide standard using streptavidin beads. MALDI-MS analysis of biotin tag-labeled citrullinated peptide standard mixed with peptide mixture before streptavidin enrichment (left); MALDI-MS analysis of eluent after streptavidin enrichment (right) showing enrichment of citrullinated peptide peaks (labeled in red) and removal of chemical background peaks;

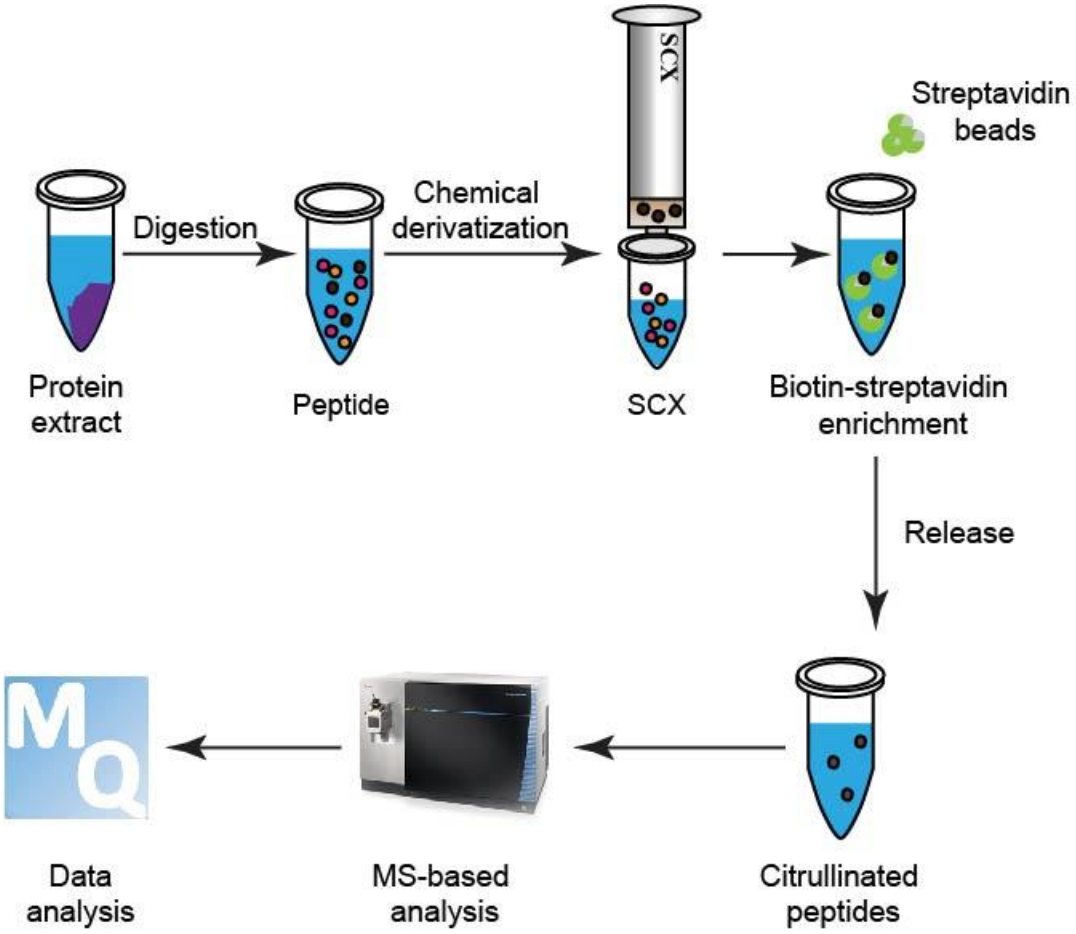


Figure 5. A standard workflow for global protein citrullination analysis of complex biological samples.

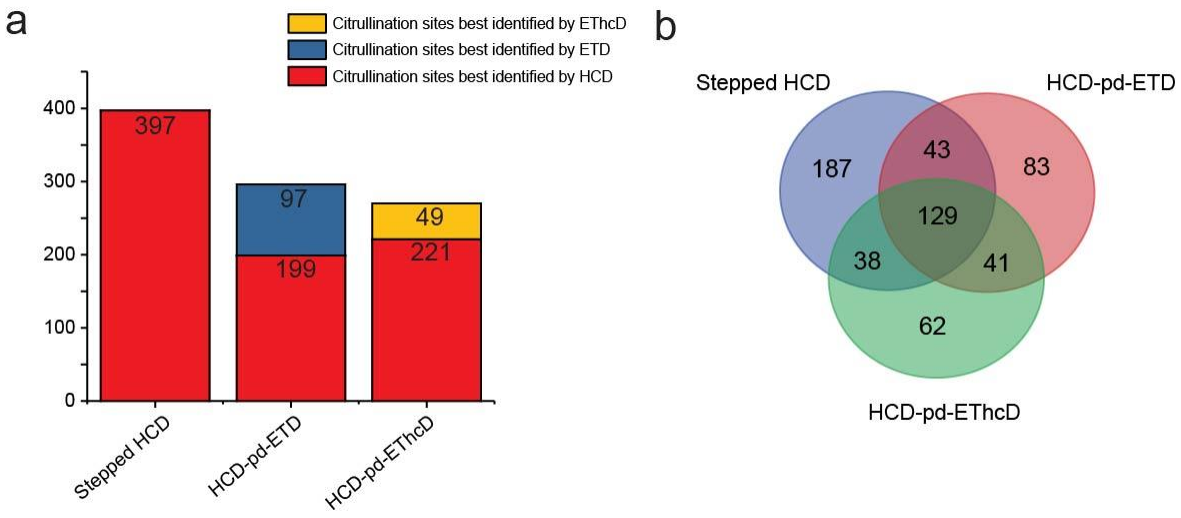


Figure 6. Optimization of MS detection method for large-scale protein citrullination analysis. (a) Histogram of the high-confident citrullination sites identified by HCD, ETD and EThcD techniques, respectively, when different MS methods were used. (b) Venn diagram showing the overlap of citrullination sites identified by three different MS methods.

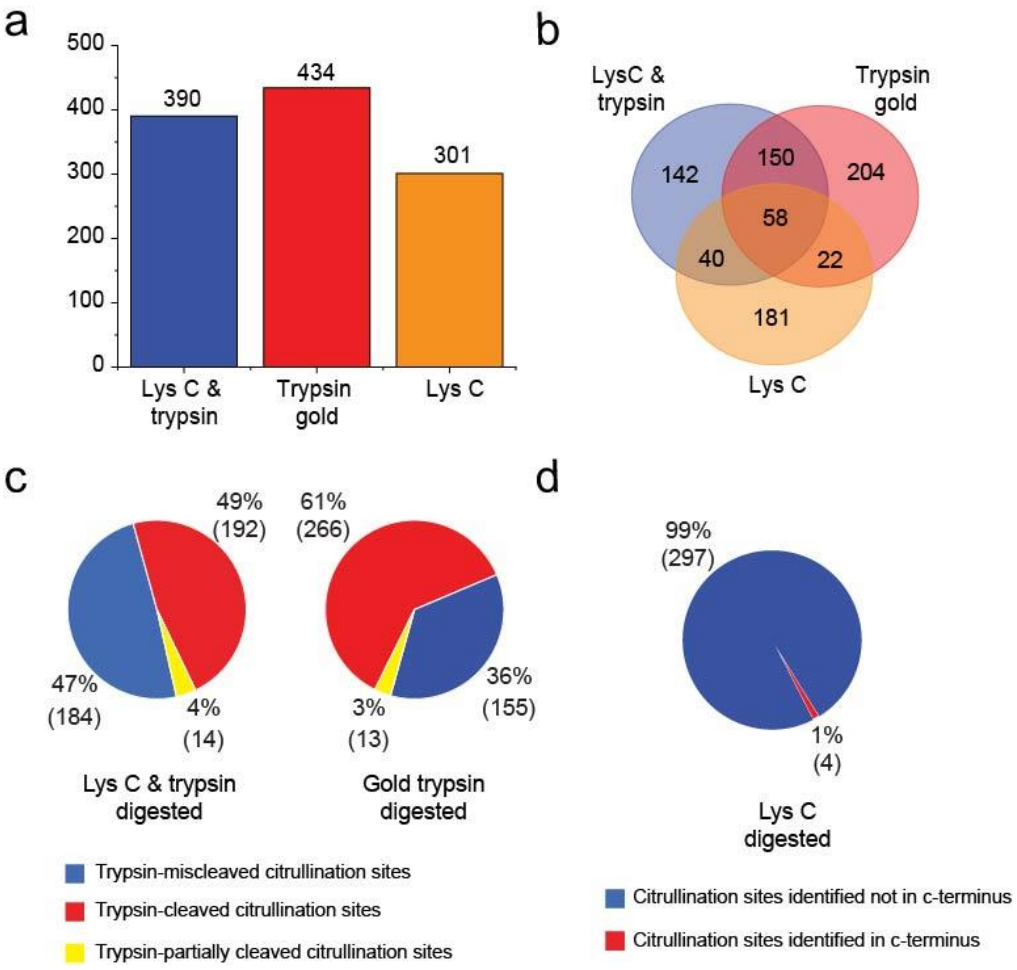


Figure 7. Comparison of enzymes used for large-scale protein citrullination analysis. (a) Histogram of the citrullination sites identified from Lys C & trypsin, trypsin gold and Lys C-digested samples, respectively. (b) Venn diagram showing the overlap of citrullination sites identified from samples digested by different enzyme. (c) Pie charts showing the ratio of cleaved, miscleaved and partially cleaved citrullination sites from samples digested by Lys C & trypsin (left) and trypsin gold (right). (d) Pie chart indicating that most citrullination sites identified from Lys C-digested sample do not locate at peptide C-terminus.

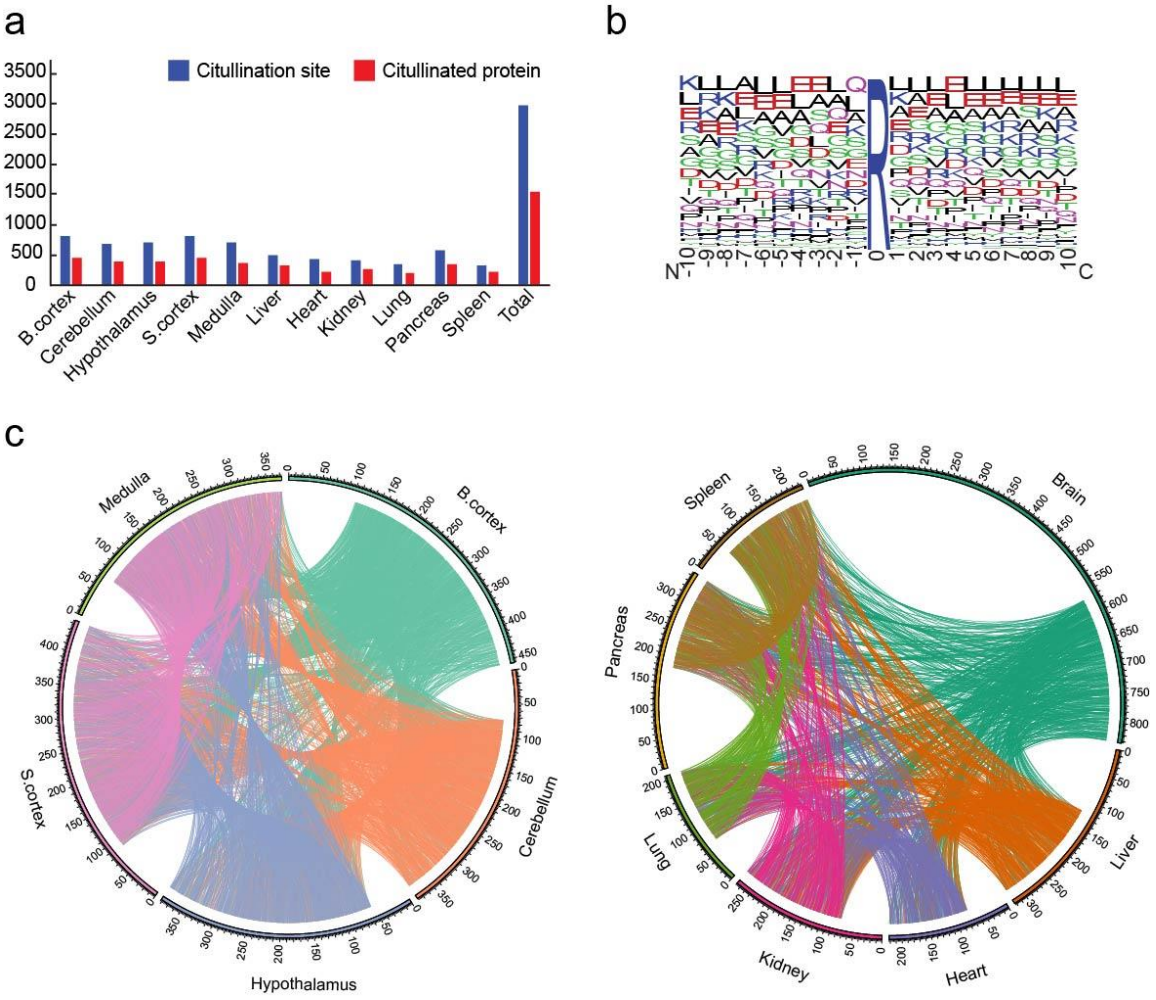


Figure 8. Mouse tissue-specific protein citrullination analysis. (a) Histogram of the citrullination sites and citrullinated proteins identified from mouse brain regions and multiple organs. (b) Citrullination motif analysis using the identified citrullination sites from mouse tissues. (c) Circo plots indicating the overlap of citrullinated proteins identified from different mouse brain regions (left), and the overlap of citrullinated proteins identified from mouse brain and other organs (right).

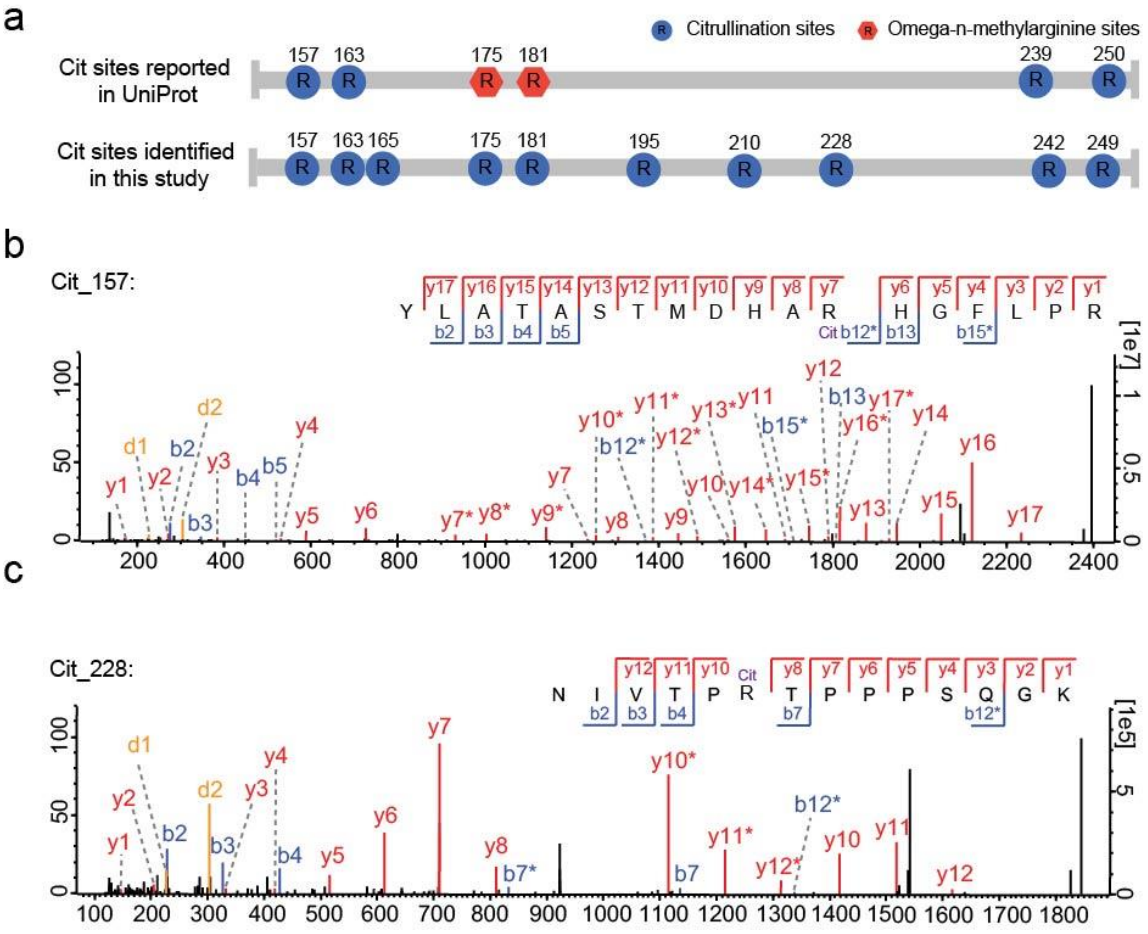


Figure 9. Citrullination sites identified from mouse myelin basic protein (MBP) protein. (a) Citrullination sites of mouse MBP protein reported in UniProt database and identified in this study. Tandem MS spectra of identified citrullinated peptides containing citrullination sites at R157 (b) and R228 (c). Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. Fragment ions resulting from the neutral loss of biotin tag were labeled with superscript star.

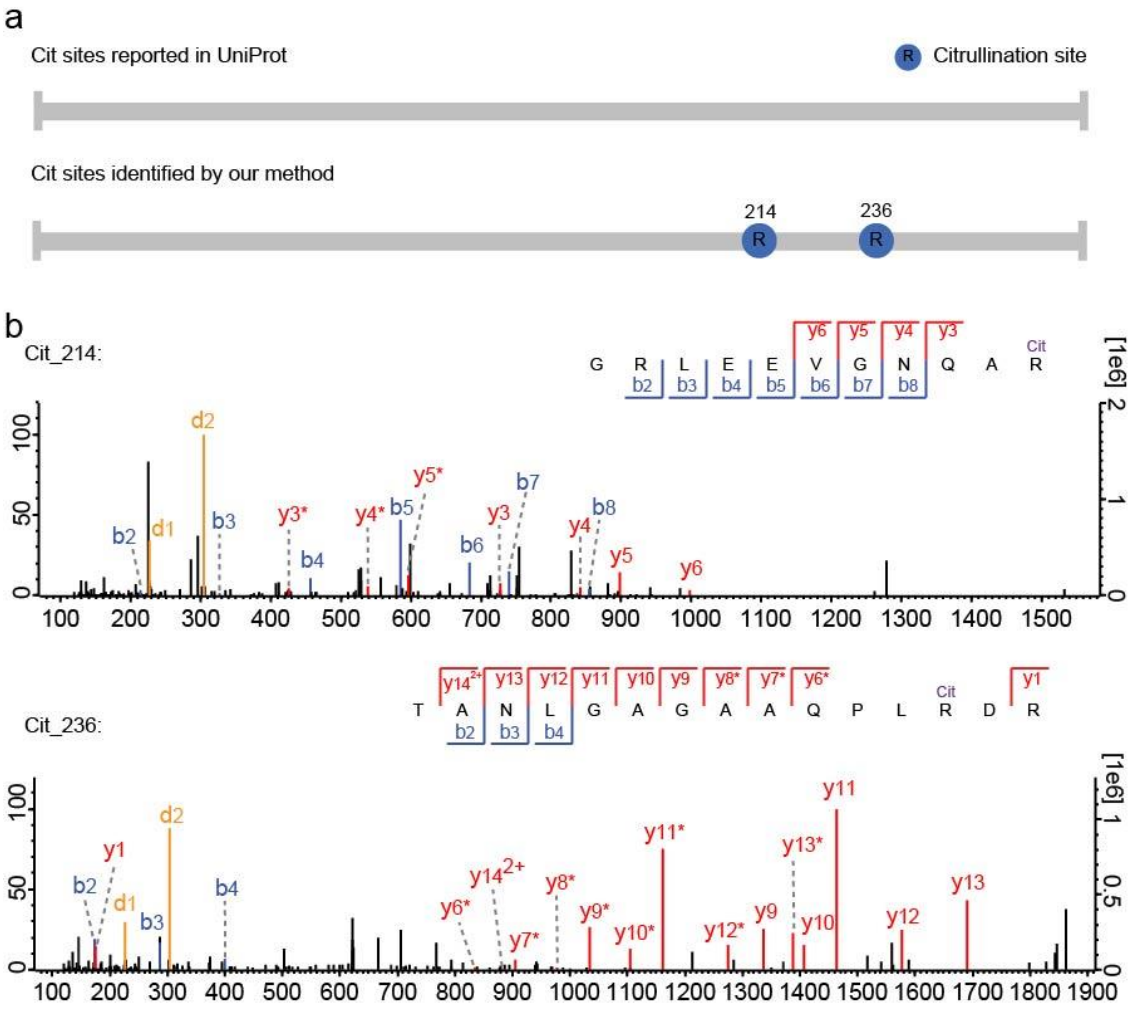


Figure 10. Two new citrullination sites identified from mouse apolipoprotein E protein. (a) Citrullination sites of mouse apolipoprotein E reported in UniProt database and identified in this study. (b) Tandem MS spectra of identified citrullinated peptides containing two new citrullination sites at R214 and R236, respectively. Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. Fragment ions resulting from the neutral loss of biotin tag were labeled with superscript star.

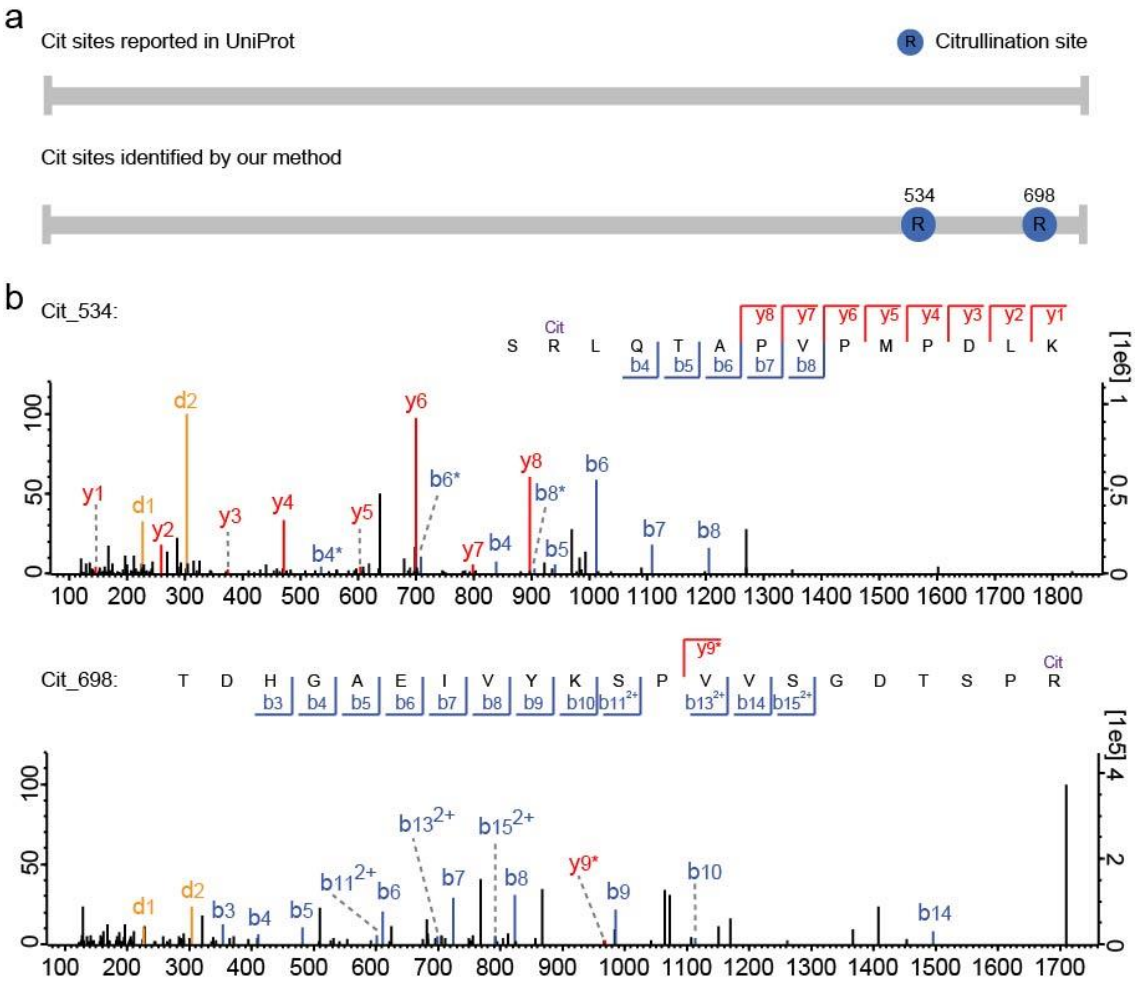


Figure 11. Two new citrullination sites identified from mouse microtubule-associated protein tau (Tau). (a) Citrullination sites of mouse Tau protein reported in UniProt database and identified in this study. (b) Tandem MS spectra of identified citrullinated peptides containing two new citrullination sites at R534 and R698, respectively. Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. Fragment ions resulting from the neutral loss of biotin tag were labeled with superscript star.

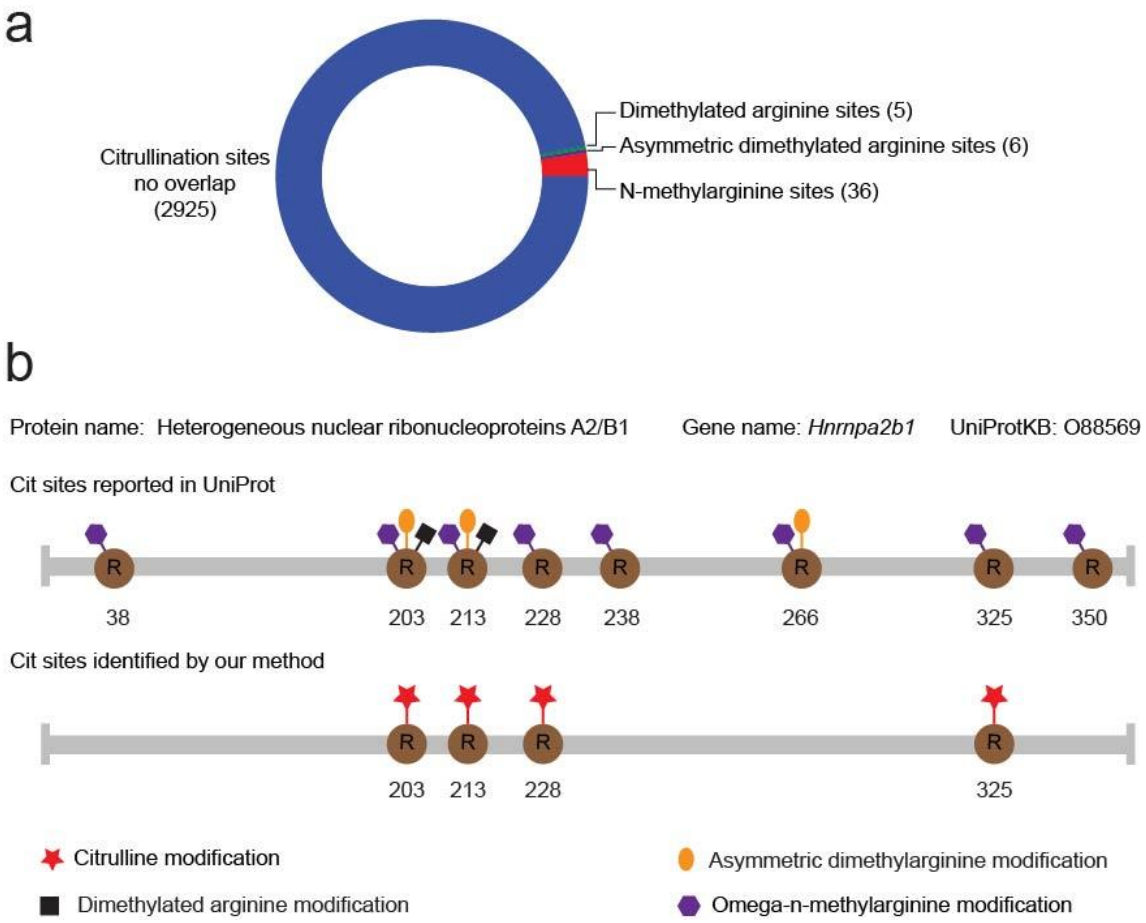


Figure 12. Potential crosstalk between citrullination and other three arginine modifications. (a) Overlap of identified citrullination sites with arginine sites reported UniProt database with dimethylated arginine, asymmetric dimethylated arginine and n-methylarginine, respectively. (b) Four citrullination sites were identified on heterogeneous nuclear ribonucleoproteins A2/B1. Besides citrullination, three other arginine modifications were reported on the site of R203 and R213. N-methylarginine modification and citrullination concurrently presented on four arginine sites.

Supplemental Information

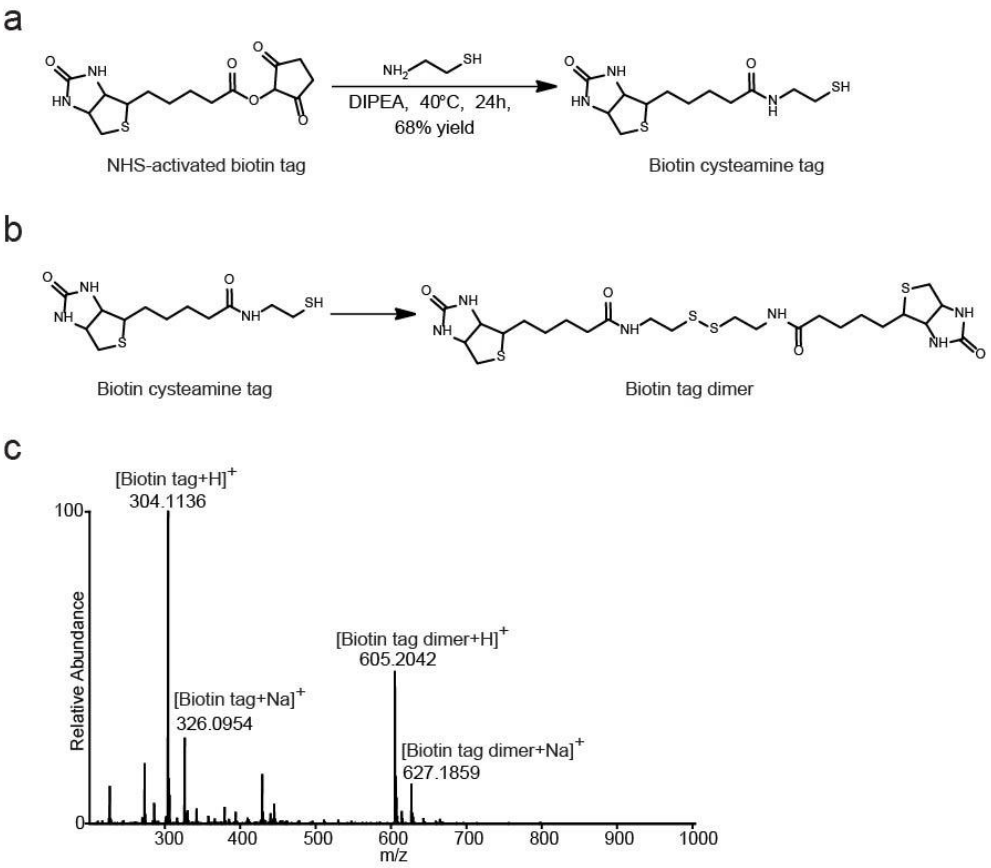


Figure S1. Synthesis of the biotin cysteamine tag. (a) One step synthesis of biotin cysteamine tag. (b) The formation of biotin tag dimer upon tag synthesis and storage. (c) MALDI-Orbitrap mass spectrometry analysis of the biotin tag indicated the presence of biotin tag dimer.

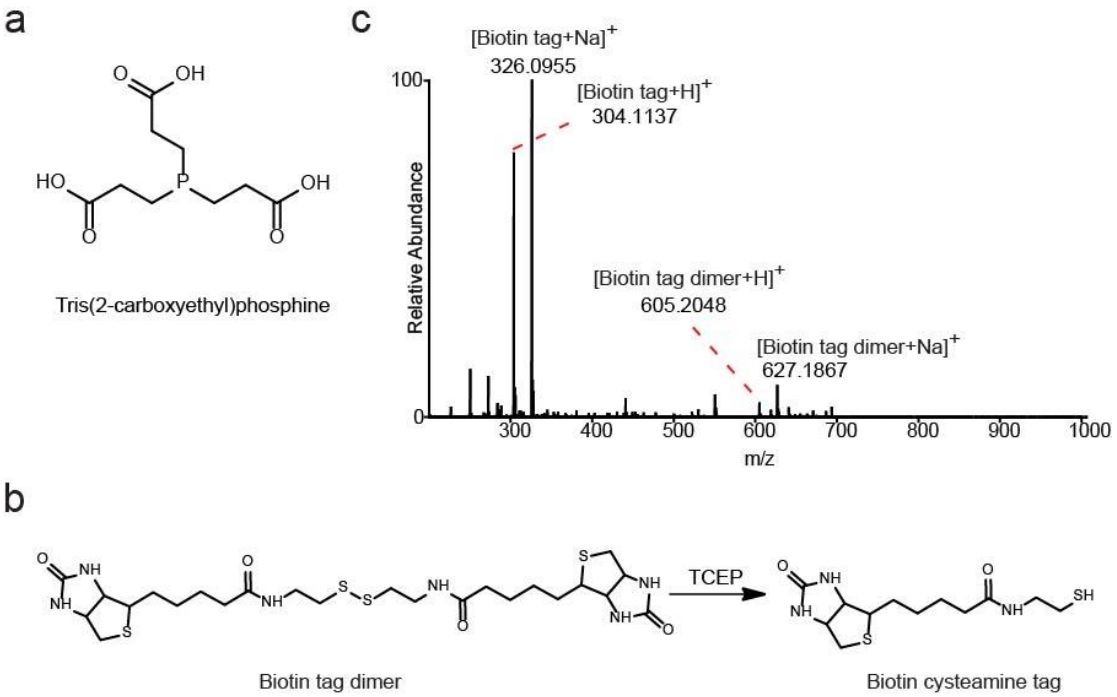


Figure S2. Protection of the biotin cysteamine tag with tris (2-carboxyethyl) phosphine (TCEP). (a) Tris (2-carboxyethyl) phosphine (TCEP), as a reducing reagent, could protect the biotin tag from the formation of dimer. (b & c) The protection effect of TCEP was examined by the MALDI-LTQ orbitrap mass spectrometry analysis.

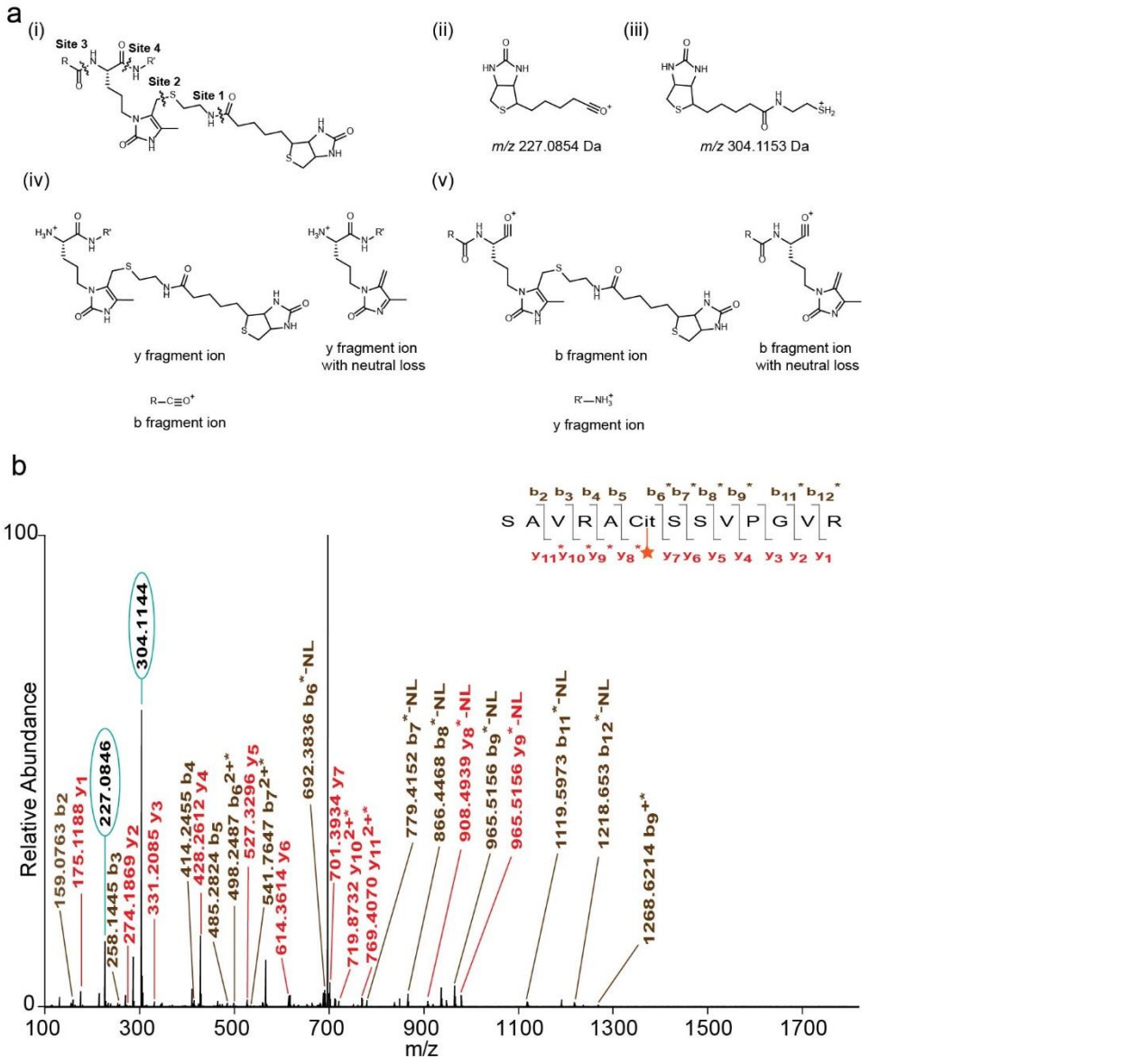


Figure S3. HCD fragmentation pattern of biotin tag-modified citrullinated peptide. (a) Cleavage sites of biotin tag-modified citrullinated peptide upon HCD fragmentation (i); Two diagnostic ions at m/z 227 (ii) and m/z 304 (iii) yielded from the modification group; Structural examples of b and y fragment ions, and fragment ion produced from the neutral loss of biotin tag (iv & v). (b) Tandem MS spectrum of biotin tag-modified citrullinated peptide standard fragmented by HCD fragmentation technique. Star represents the intact chemical modification group.

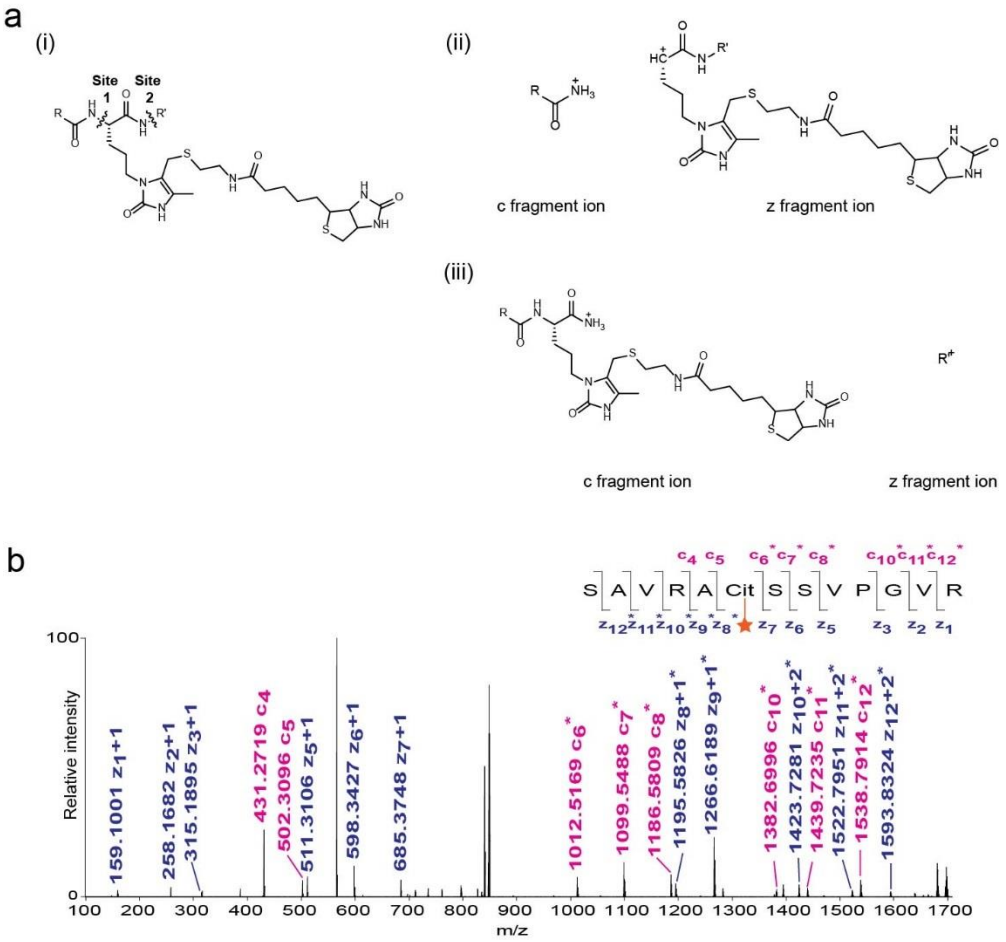


Figure S4. ETD fragmentation pattern of biotin tag-modified citrullinated peptide. (a) Cleavage sites of biotin tag-modified citrullinated peptide upon ETD fragmentation and no diagnostic ions were observed (i); Structural examples of c and z fragment ions (ii & iii). (b) Tandem MS spectrum of biotin tag-modified citrullinated peptide standard fragmented by ETD fragmentation technique. Star represents the intact chemical modification group.

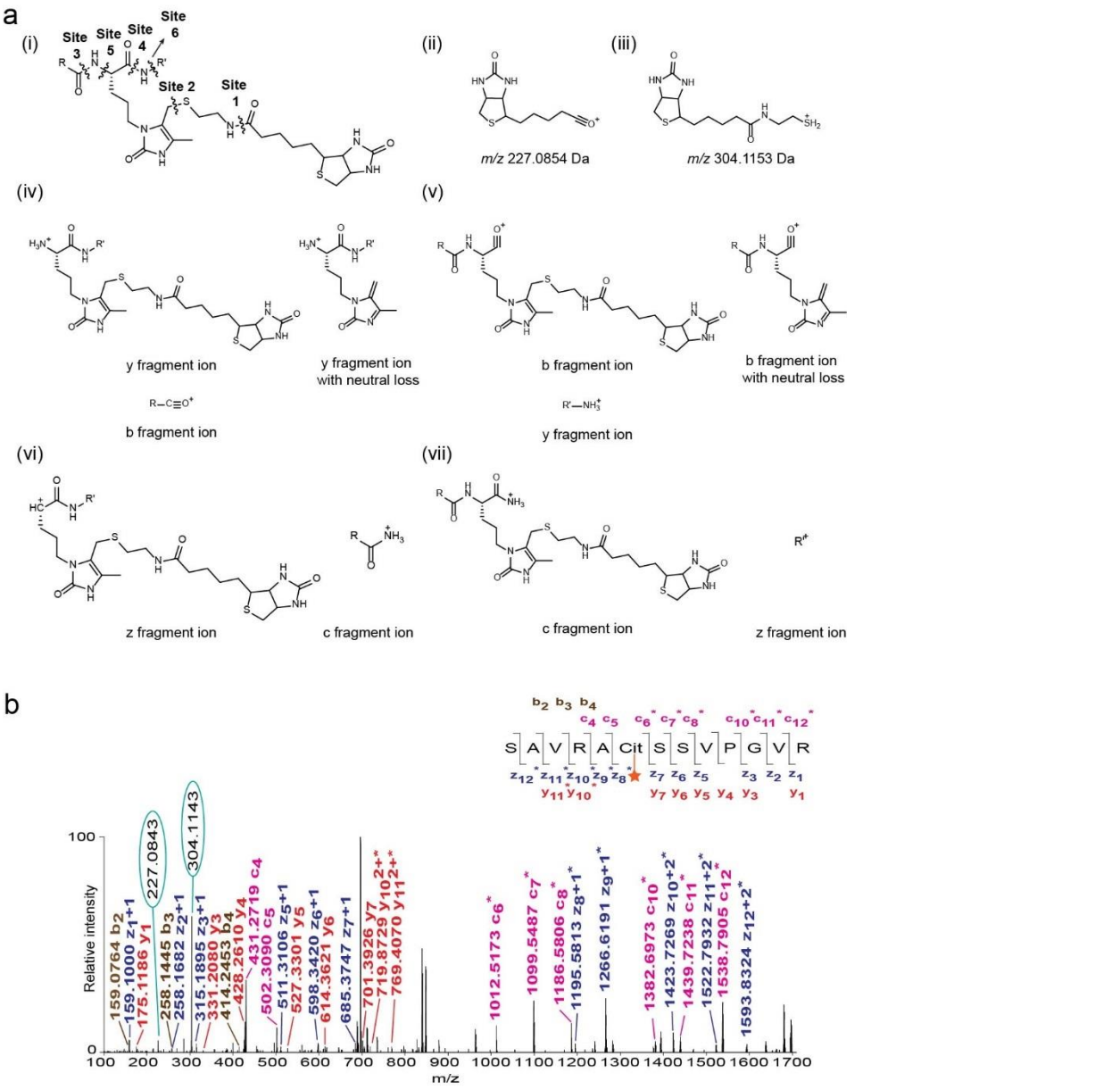


Figure S5. EThcD fragmentation pattern of biotin tag-modified citrullinated peptide. (a) Cleavage sites of biotin tag-modified citrullinated peptide upon EThcD fragmentation (i). Two diagnostic ions at m/z 227 (ii) and m/z 304 (iii), fragment ions generated by EThcD fragmentation technique (iv-vii). (b) Tandem MS spectrum of biotin tag-modified citrullinated peptide standard fragmented by EThcD fragmentation technique. Star represents the intact chemical modification group.

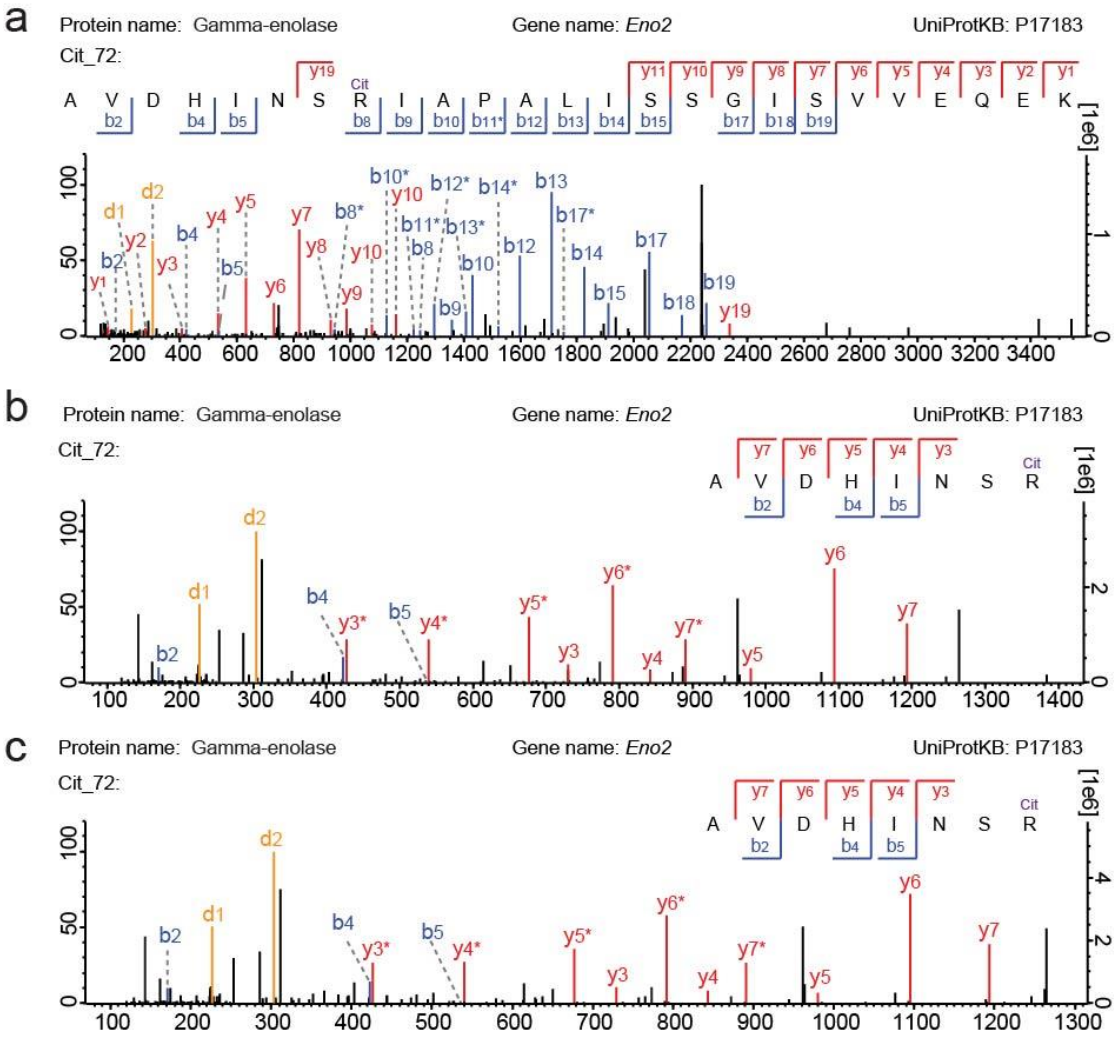


Figure S6. Representative of citrullinated peptides containing the same citrullination site identified from peptide mixture digested by different enzymes from the same brain extract.

(a) Lys C-digested peptide mixture. (b) Lys C & trypsin gold-digested peptide mixture. (c) Trypsin gold-digested peptide mixture.

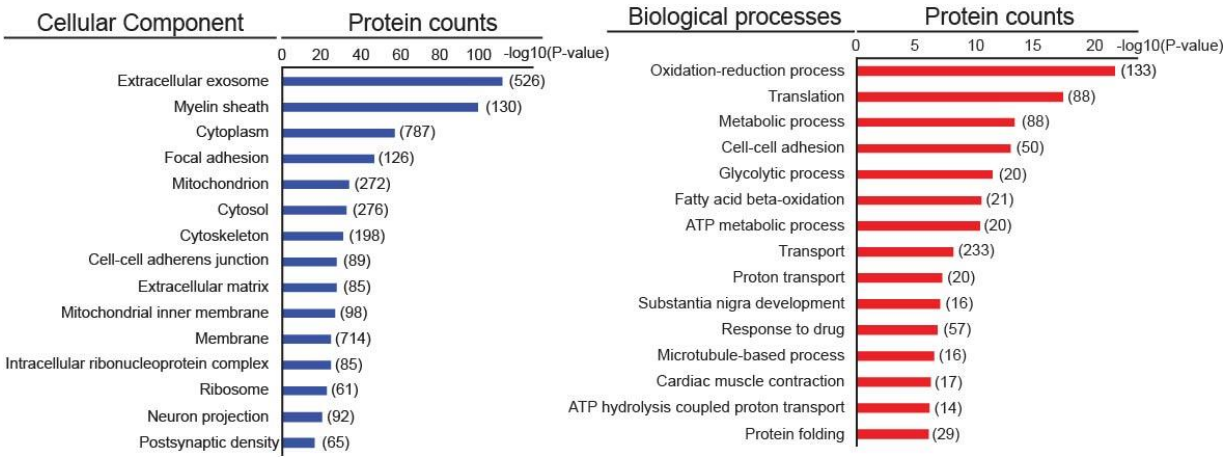


Figure S7. Gene ontology (GO) cellular component and biological processes analysis of citrullinated proteins identified from mouse tissues.

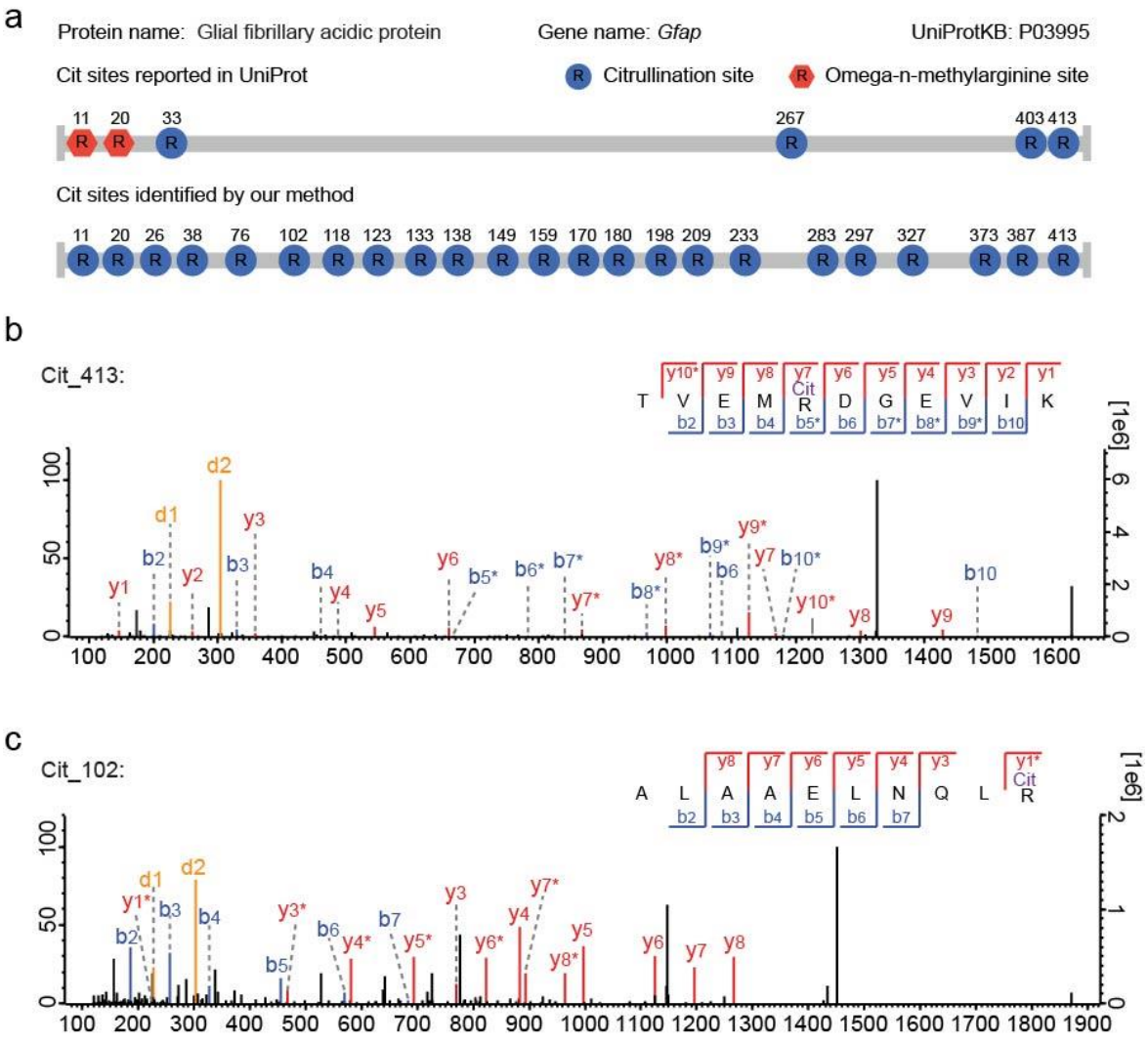


Figure S8. Citrullination sites identified from mouse glial fibrillary acidic protein (GFAP). (a) Citrullination sites of mouse GFAP reported in UniProt database and identified in this study. Tandem MS spectra of identified citrullinated peptides containing citrullination sites at R413 (b) and R102 (c). Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. Fragment ions resulting from the neutral loss of biotin tag were labeled with superscript star.

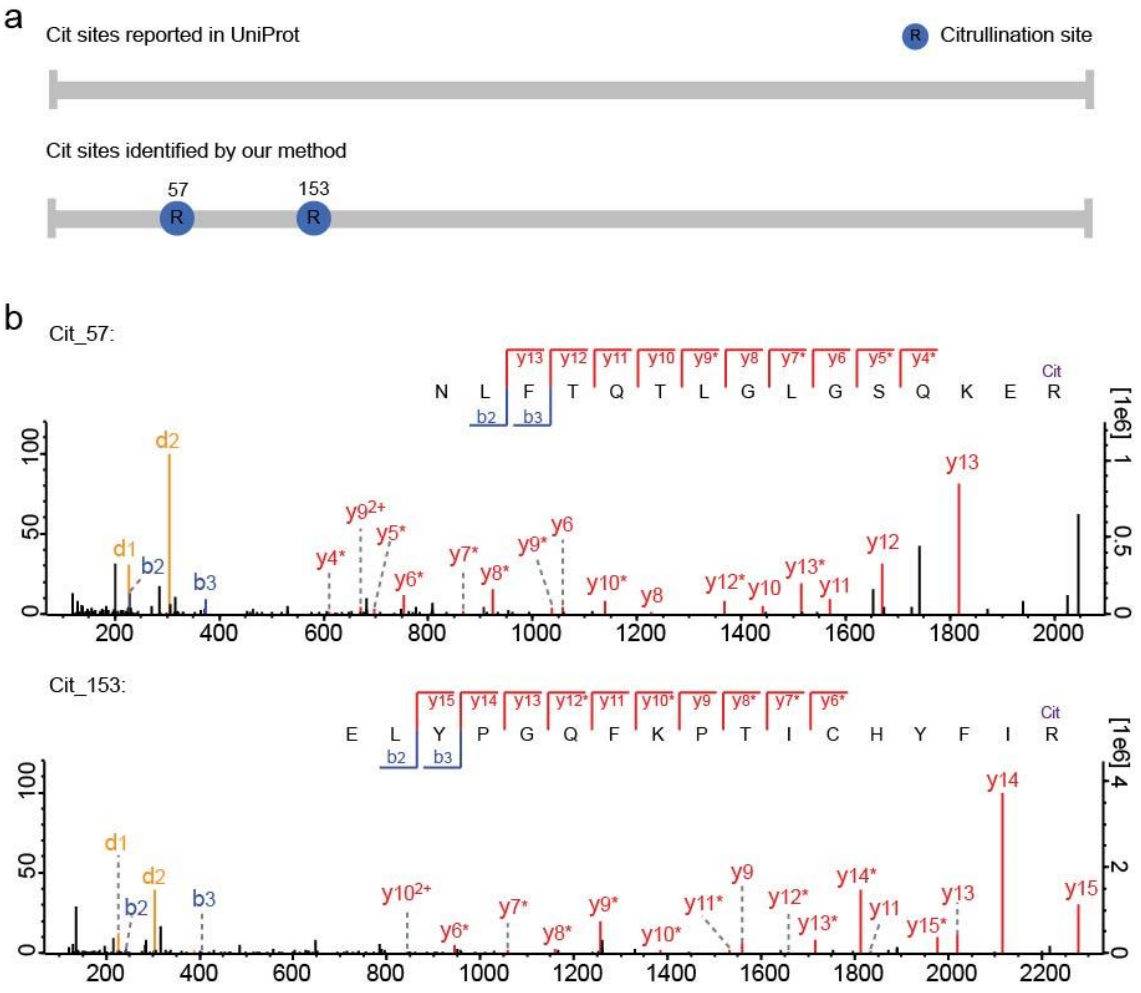


Figure S9. Two new citrullination sites identified from mouse NAD-dependent protein deacetylase sirtuin-2. (a) Citrullination sites of mouse NAD-dependent protein deacetylase sirtuin-2 reported in UniProt database and identified in this study. (b) Tandem MS spectra of identified citrullinated peptides containing two new citrullination sites at R57 and R153, respectively. Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. Fragment ions resulting from the neutral loss of biotin tag were labeled with superscript star.

Chapter 3

MS-based Methods for Quantitative Analysis of Citrullinated Proteins in Biological Samples

This project is performed in Prof. Lingjun Li lab. Yatao Shi leads the project and is responsible for the experiment design and MS-based analysis.

Abstract

Protein citrullination plays an important role in many serious diseases. However, a method for global quantitative analysis of citrullinated proteins in biological samples is currently lacking, significantly impeding our understanding about the pathogenic role of protein citrullination in diseases. Thus, there is an urgent need to develop novel methods enabling the comparative analysis of citrullinated proteins from various specimens, especially samples from healthy control and patients. MS-based isotopic labeling and isobaric labeling strategies have become widespread tools for quantitative proteomic studies. Here, we report the development and evaluation of two novel MS-based quantitative methods for the relative quantitation of citrullinated proteins. The first method involves combination of the recently developed biotin tag-assisted MS-based method with the stable isotope dimethyl labeling strategy, for simultaneous qualitative and quantitative analysis of citrullinated proteins. This method was initially evaluated by using citrullinated peptide standard. We further tested the performance of the new method using real biological sample and obtained reliable quantitative accuracy. To expand the utility and high throughput capability, we also demonstrated that the biotin tag-assisted MS-based method could be integrated with our 12-plex DiLeu isobaric tagging reagents, for multiplex quantitative analysis of citrullinated proteins from different samples. These novel MS-based quantitative methods will provide the scientific community a powerful tool to unravel the roles of citrullination in biological and disease pathways.

Introduction

Protein citrullination is initially known as a disease-related protein post-translational modification (PTM) in rheumatoid arthritis (RA).[1-5] Recently, more evidence suggests its involvement in many devastating diseases, such as multiple sclerosis,[6,7] lupus,[8] Parkinson's disease,[9,10] Alzheimer's disease,[11] diabetes and cancers.[12-14] Although important, the pathogenic role of protein citrullination in these diseases is still unclear. One major challenge is the lack of method for large-scale and high throughput protein citrullination analysis of biological samples, which further limits the development of quantitative strategies for protein citrullination analysis. Therefore, there is an urgent need to develop novel quantitative approaches to probe the expression level changes of citrullinated proteins in different biological states.

Prior studies in terms of quantitative protein citrullination analysis suffer from low throughput capability. For example, highly expressed citrullinated proteins, such as fibrinogen and vimentin, and their corresponding antibodies, termed as anti-citrullinated protein antibodies (ACPA) or anti–cyclic citrullinated peptide (anti-CCP) antibodies, are found in the plasma of RA patients.[15,16] Based on this, an array of anti-CCP antibody enzyme-linked immunosorbent assay (ELISA) kits have been developed and widely used in clinic as a powerful tool for the early diagnosis and prognosis of RA.[17,18] Some fluorescence probes have recently been reported that enable visualization and relative quantification of citrullinated proteins from different biological samples.[19,20] However, the fluorescence probe-based method does not allow the identification of citrullinated proteins at a large scale, nor does it reveal the dynamic changes of individual citrullinated protein in different samples, limiting its application to clinical sample analysis.

Benefiting from its high throughput capability, mass spectrometry (MS)-based quantitative proteomics, including label-free or label-based, has been widely applied to the relative

quantification of protein expression levels in different specimens.[21-23] The label-free quantitative method has the advantages of unlimited number of samples and easy sample preparation process, but suffering from poor quantitative accuracy and reproducibility, whereas label-based methods require multiplex labeling of samples and enable comparative analysis of analyzing multiple labeled samples in one LC-MS run, leading to improved quantitative accuracy and decreased run-to-run variability. Stable isotopic labeling strategy, such as stable isotope labeling by amino acids in cell culture (SILAC)[24] and stable-isotope dimethyl labeling[25], and isobaric labeling strategy including tandem mass tags (TMT),[26] isobaric tags for relative and absolute quantitation (iTRAQ)[27] and dimethylated leucine (DiLeu) tags[28], are two commonly used label-based strategies for MS-based quantitative proteomics.

In this work, I developed two novel methods, by integrating the newly developed biotin tag-assisted MS-based citrullination analysis method with either stable-isotope dimethyl labeling or our DiLeu isobaric tag labeling, for simultaneous qualitative and quantitative protein citrullination analysis. Two quantitative methods were tested and evaluated using the citrullinated peptide standard and reliable quantitation accuracy can be obtained. We also demonstrate the utility of novel quantitative method using stable-isotope dimethyl labeling strategy for complex biological sample analysis.

Materials and Methods

Materials

Formaldehyde solution, formaldehyde-d2 solution and borane pyridine, 1.0 M triethylammonium bicarbonate buffer (TEAB) and 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium tetrafluoroborate (DMTMM), Sodium dodecyl sulfate (SDS) and streptavidin agarose were purchased from Sigma-Aldrich (St. Louis, MO). Phosphate-Buffered

Saline (PBS) buffer was ordered from Mediatech, Inc (Manassas, VA). Ammonium bicarbonate, Optima LC/MS grade formic acid (FA), Tris base, water, acetonitrile and formic acid (FA) were purchased from Fisher Scientific (Pittsburgh, PA). Omix C_{18} pipet tips were purchased from Agilent Technologies (Santa Clara, CA). SepPak C_{18} SPE cartridges were purchased from Waters (Milford, MA). N-Methylmorpholine (NMM) was purchased from TCI America (Tokyo, Japan). Trypsin was purchased from Promega (Madison, WI)

Optimization of buffer acidity for dimethyl labeling of standard citrullinated peptide

Four resuspension buffers were prepared with different acidity including neutral (H_2O), mild acidic (1.25 % FA, v/v), medium acidic (2.5 % FA, v/v) and high acidic (5 % FA, v/v). Forty micro liters of each resuspension buffer was individually added to dissolve 1 μg standard citrullinated peptide. After diluted to 1 % (v/v) with various buffers, 20 μL formaldehyde was added to each sample. To each sample, 20 μL of borane pyridine (30 mM in distinct buffers) was then added to initiate the labeling reaction. Following the incubation at 37 ℃ for 20 min, labeling reaction was quenched by adding 20 μL ammonium bicarbonate solution (200 mM in H_2O). Finally, samples were acidified with FA to pH < 3, desalted with Omix C_{18} pipet tips and dried in vacuo for later labeling efficiency test on a MALDI-LTQ-Orbitrap XL mass spectrometer.

Stable isotope dimethyl labeling of citrullinated peptide standard

Forty micro liters of H_2O was added to dissolve 1 μg standard citrullinated peptide. After diluted to 1 % (v/v) with H_2O, 20 μL formaldehyde or formaldehyde-d2 solution was added to samples for light or heavy labeling, respectively. To each sample, 20 μL of borane pyridine (30 mM) was then added to initiate the labeling reaction. Following incubation at 37 ℃ for 20 min, labeling was quenched by addition of 20 μL ammonium bicarbonate solution (200 mM). Labeled standard peptides were then combined in 1:1, 2:1 or 5:1 ratio (v/v, light/heavy). Samples were

acidified with FA to pH < 3, desalted with Omix C_{18} pipet tips and tested on MALDI-LTQ-Orbitrap XL mass spectrometer.

Biotin tag labeling of dimethyl labeled citrullinated peptide standard

The novel biotin thiol tag was easy to be synthesized in just one step using biotin-NHS ester and cysteamine (**Scheme S1**). The dimethyl labeled citrullinated peptide standard combined in different ratios (1:1, 2:1, 5:1, v/v, light/heavy) prepared above was further derivatized by the biotin thiol tag. Briefly, 30µl biotin thiol tag solution (10 mg/ml in 50% MeOH/H_2O) was added to citrullinated peptide standard and dried with SpeedVac, which was further reconstituted in 40µl 12.5% TFA/H_2O. 10 µl diluted 2,3-butanedione solution (100 mM in 12.5% TFA/H_2O) was subsequently added to initiate the chemical derivatization reaction.

After derivatization, the leftover biotin tag was completely removed by strong cation exchange (SCX) chromatography using TopTips (TT200SEA, Poly LC) containing PolySULFOETHYL A beads. SCX beads were pre-equilibrated with 100µl loading buffer containing 50% ACN/0.2% FA/10mM ammonium formate for three times. Then, derivatized citrullinated peptide standard was reconstituted in 200µl loading buffer and applied to the SCX beads twice followed by continuous washing of the beads with 100µl loading buffer for 10 times. All centrifugation steps were performed at 2,000 rpm for 2 min. Bound peptides were eluted with 150 µl 25% ACN/0.4M ammonium formate for two times and dried out in SpeedVac, which were finally tested on a MALDI-LTQ-Orbitrap XL mass spectrometer.

Enzymatic protein digestion

Mouse brain was homogenized with a probe ultrasonicator in 4% sodium dodecyl sulfate (SDS)/50mM Tris base buffer (~pH 8, adjusted with HCl). After homogenization, the supernatant

was collected after centrifugation at 16,000 rcf for 15 min. Protein concentration was determined by BCA protein assay reagent (Thermo Scientific, Fair Lawn, NJ).

Protein extract was reduced by 10mM dithiothreitol (DTT) for 30 min at room temperature and then alkylated with 50mM iodoacetic acid (IAA) for another 30 min in the dark. Protein was further precipitated by adding cold acetone (-20°C) to a final concentration of 80% (v/v). After incubation in -20°C overnight, the sample solution was centrifugated at 16,000 rcf for 15 mins and the supernatant was discarded. To completely remove SDS, the pellet was washed with cold 80% acetone for another two times, and then dried in the air for 12 mins. The dried pellet was reconstituted with 150µl 5 M guanidine hydrochloride/50mM Tris-base buffer (pH 8, adjusted with HCl). After reconstitution, the buffer solution was diluted ten-fold with 50mM Tris-base buffer to reduce the concentration of guanidine HCl to 0.5 M. Mass spec grade Trypsin/Lys-C Mix (Promega, Madison, WI; Protein: enzyme, 100:1, w/w) was added to samples and incubated at 37°C for 6 h. The secondary digestion was performed by adding the same amount of Trypsin/Lys-C Mix and incubating samples at 37°C for another 12 h. The digestion process was quenched by adding 10% TFA to reduce the pH to <3. Digested peptides were desalted with C_{18} columns (Sep-Pak, waters, Milford, MA) and eluted with 0.8ml 80% ACN/0.2% FA for two times, which were dried out with SpeedVac. Peptide concentration was determined by colorimetric peptide assay (23275, Thermo Scientific, Fair Lawn, NJ).

Evaluation of the quantitative method combining dimethyl labeling and biotin tag labeling using tryptic digest from mouse brain

One set of samples containing 300 µg, 400 µg and 500 µg tryptic peptides prepared above was lightly dimethyl-labeled, while another sample group containing 300 µg, 200 µg and 100 µg tryptic peptides was heavily dimethyl-labeled. After labeling, differentially labeled peptides were

then combined in 1:1, 2:1 or 5:1 ratio (light/heavy, w/w) to an equal total amount of 600 µg. Each sample was then acidified with FA to pH < 3, desalted with a SepPak C_{18} SPE cartridge and dried in vacuo.

The combined dimethyl-labeled tryptic peptides are further chemically derivatized by the novel biotin thiol tag. The labeling and SCX cleaning processes are similar with that of citrullinated peptide standard mentioned above. For the enrichment process, 300µl 50% ACN/H_2O was added to each peptide sample recovered from SCX cleaning. After brief vortex and centrifugation, each sample was dried in SpeedVac to less than 100µl and further diluted with 900µl PBS 1× buffer. 75µl streptavidin beads were pre-washed with 1ml PBS 1× buffer for 5 times. The peptide sample was diluted with 950 µl PBS 1× buffer and was further loaded onto the pre-washed streptavidin agarose. The binding process was performed at room temperature for 2 h with rotation. The beads were subsequently washed 4 times each with 1ml PBS 1× buffer, 1ml 5 % ACN/PBS 1× buffer and 1ml water. The bound peptides were finally released with 300µl 80%ACN/H_2O contain 0.2% TFA and 0.1% FA for four times. The first release was performed in room temperature for 5min, while other three release processes were conducted at 95°C for 5min with shaking. The eluents were combined and dried in the SpeedVac. Enriched peptides were desalted with C_{18} columns (Zip-tips) and eluted in 100 µl 80% ACN/0.2% FA solution, which was finally dried in the vacuum to almost dry.

Reversed-phase separation and mass spectrometry

MS analysis was performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer coupled with Dionex UltiMate 3000 UPLC system (Thermo Fisher Scientific, San Jose, CA). The mobile phase was composed of 0.1% formic acid in water (A) and 0.1% formic acid in ACN (B). The flow rate was set as 300 nl/min. Each sample was resuspended in 15 µl 0.1% formic acid 3%

ACN in water and underwent three LC-MS/MS run. For each run, 2µl of sample was loaded onto a customized C18 column filled with 1.7µm Ethylene Bridged Hybrid packing materials (130 Å, Waters) and separated with the following gradient: 3% B for the first 18.3 min upon sample trapping; 3%-30% B for 18.3-120 min; 30%-75% B for 120-120.5 min; 75% B for 120.5-130 min; 75%-95% B for 130.0-130.5 min; 95% B for 130.5-140.0 min; 95%-3% B for 140.0-140.5 min and equilibrated at 3% B for 15 min. MS data was acquired in the positive ion mode with the spray voltage of 2 kV. MS1 spectra were collected from m/z 350-1500 by orbitrap at a resolution of 60 K. Automatic gain control (AGC) target and maximum injection time was set to $2e^5$ and 100 ms, respectively. Ions with the charge of 2-6 were included for screening, while a dynamic exclusion time of 45s was used to avoid repeatedly sequencing the same precursor ion in a short time. For the MS^2 fragmentation, the top 15 precursors were selected for stepped HCD fragmentation with an isolation window of 1 Da, a resolving power of 15K and AGC target of 5×10^4.

Data analysis of dimethyl labeled mouse brain digest

Mass spectra were processed with MaxQuant version 1.5.2.8 using default settings if not stated otherwise. Raw files were searched against the mouse UniProt FASTA database (December 2018) using the Andromeda search engine integrated into the MaxQuant environment with trypsin selected as the enzyme and three missed cleavages allowed. Oxidized methionine (M, +15.995 Da) and biotin tag-labeled citrullination (R, +354.1072 Da) were selected as variable modifications, and carbamidomethyl (C, +57.02146 Da) as fixed modification. Multiplicity was set to two with dimethLys0/dimethNter0 specified as light labels, and dimethLys4/dimethNter4 as heavy labels. Proteins and peptides were identified with a target-decoy approach in revert mode and false-discovery rate (FDR) at the protein or peptide level was set to 0.01. Peptides that were found as

reverse or potential contaminant hits were filtered out and citrullination site localization probability threshold was set to 0.75.

DiLeu labeling of citrullinated peptide standard

DiLeu tags were suspended in anhydrous DMF and combined with DMTMM and NMM at 0.6× molar ratios to DiLeu. The mixture was then vortexed at room temperature for 60 min. Following centrifugation, the supernatant was immediately mixed with the citrullinated peptide standard. DiLeu labeling was performed by adding the activated label solution to citrullinated peptide standard at a ratio 10:1 (w/w) and vortexing at room temperature for 2 h. The reaction was quenched by adding 5% NH_2OH to the final concentration of 0.25% and dried under SpeedVac. Citrullinated peptide standard labeled with different DiLeu channels were then mixed with experiment-specific ratios. (1:1:1:1 and 1:1:2:5).

Results and discussion

In this study, we developed a novel method, which combines the stable isotope dimethyl labeling technique and the biotin tag-assisted, MS-based method established in Chapter 2, for simultaneous qualitative and quantitative analysis of citrullinated protein in biological samples. The reaction condition of stable isotope dimethyl labeling was optimized. Furthermore, the quantitation accuracy of the novel method was evaluated by using citrullinated peptide standard and peptide mixtures digested from mouse brain protein extract. In addition, we also tested the integration of DiLeu isobaric labeling technique with the biotin tag-assisted, MS-based method for multiplex and high-throughput quantitative protein citrullination analysis.

Optimization of stable isotope dimethyl labeling of citrullinated peptides

Stable isotope dimethyl labeling technique is to label primary amines at the N-terminus and the side chain of lysine in proteins or peptides by using cyanoborohydride and formaldehyde

as well as their stable isotope-labeled forms, for 2-plex and 3-plex quantitative proteomic analysis (**Figure 1**).[29] Compared to other isotopic labeling strategies, such as stable isotope labeling with amino acids in cell culture (SILAC)[30] and isotope-coded affinity tag (ICAT),[31,32] stable isotope dimethyl labeling is much simpler, faster and more cost-effective, and thereby is one of the most commonly used isotopic labeling strategies for MS-based quantitative analysis.

However, it has been reported that the free amine groups at the N-terminus and the side chain of lysine could be differentially dimethyl labeled depending on the reaction condition used, especially the acidity of the reaction buffer.[33] Hence, we firstly optimized the dimethyl labeling reaction using the citrullinated peptide standard. Three different reaction buffers were investigated including neutral H_2O, 1.25% FA/H_2O and 2.5% FA/H_2O. As shown in **Figure 2**, the citrullinated peptide standard could be completely dimethyl labeled in neutral H_2O without apparent side reaction observed. In low acidic aqueous solution, numerous peaks resulting from side reactions were detected though the dimethyl labeled citrullinated peptide remained to be the dominant product. Therefore, the neutral H_2O was finally used as the reaction buffer for the dimethyl labeling of citrullinated peptides.

Given that the stable isotope dimethyl labeling and biotin thiol tag labeling target the primary amines and the ureido group at the side chain of citrulline, respectively, we proposed a novel method, which combines these two labeling strategies, for the simultaneous qualitative and quantitative analyses of citrullinated proteins from different biological specimens (**Figure 3)**.

Evaluation of the quantitative method using citrullinated peptide standard

The new method combining stable isotope dimethyl labeling and biotin thiol tag labeling was initially evaluated by using the citrullinated peptide standard. The citrullinated peptide standard was lightly and intermediately dimethyl labeled, respectively, and further combined in

three different ratios (1:1, 2:1 and 5:1, w/w) for subsequent biotin thiol tag labeling. As shown in **Figure 4**, the differentially dimethyl labeled citrullinated peptide standard could be completely derivatized by the biotin thiol tag and no apparent side reaction was observed. Moreover, reliable quantitation accuracy was acquired, demonstrating the feasibility of combining the stable isotope dimethyl labeling technique and biotin thiol tag labeling approach.

Performance of the quantitative method on complex biological samples

Although the quantitation accuracy of the new method has been demonstrated by the citrullinated peptide standard, we moved further to assess its quantitative performance with real biological samples. As depicted in **Figure 5**, different amounts of citrullinated peptides digested from the same mouse brain protein extract were lightly and intermediately dimethyl labeled, respectively, and further combined in three different ratios (1:1, 2:1 and 5:1, w/w) for subsequent biotin thiol tag labeling, streptavidin enrichment and MS analysis.

The quantitative results were summarized in **Figure 6**. It was no surprise that the number of identified citrullination sites and proteins decreased primarily due to the increased complexity of full MS spectrum caused by dimethyl labeling. In addition, fewer identified citrullinated proteins could be quantified when sample difference increases, which is another drawback of isotopic labeling-based quantitative strategy. Nevertheless, reliable quantitative accuracy was still obtained for these quantifiable citrullinated proteins. These results suggested the applicability of the novel method for comprehensive analysis of citrullinated proteins in different biological samples.

DiLeu isobaric tag labeling for quantitative protein citrullination analysis

Stable isotope dimethyl labeling, as the commonly used isotopic labeling strategy, is mainly amenable to 2-plex and 3-plex quantitative analysis, and suffers from the low throughput

capability. Although 5-plex dimethyl labeling method has recently been reported, its application to biological sample analysis is significantly limited due to the increased complexity of full MS spectra.

Our lab has developed our own multiplex isobaric tags, DiLeu isobaric tandem-mass tagging reagents, which can be synthesized with ease using readily available isotopic reagents, for high-throughput MS-based relative quantification of peptides and proteins (**Figure S2**). Compared to isotopic labeling, the DiLeu isobaric labeling tag has numerous advantages including: (1) Combination of multiple samples for a single MS analysis reduces the amount of material from each sample, which is essential for precious sample analysis; (2) Parallel sample analysis reduces labor work and saves the sample preparation and instrument analysis time; (3) Multiplex labeling enables the collection of biological replicate data; (4) Simultaneous MS analysis results in greater overlap between samples across different conditions. Meanwhile, similar with the dimethyl labeling reagent, the DiLeu isobaric tags also react with the free amine groups in peptides or proteins, and thereby can be used as an alternative strategy for multiplex quantitative protein citrullination analysis. The proposed method, which integrates the DiLeu isobaric labeling with the biotin tag-assisted, MS-based method, was shown in **Figure S3**.

We firstly investigated the DiLeu isobaric labeling of citrullinated peptide standard. As shown in **Figure S4**, the citrullinated peptide standard could be labeled by DiLeu isobaric tag, resulting in a mass increment of 145 Da. However, about 30% of citrullinated peptide standard was doubly labeled by DiLeu isobaric tag. Besides the primary amine group at the N-terminus, the ureido group at the side chain of citrulline was the functional group available for DiLeu labeling. This result indicated that the ureido group of citrulline could be partially labeled upon DiLeu labeling.

Current commercial isobaric labeling reagents, such as tandem mass tag (TMT) and isobaric tag for relative and absolute quantitation (iTRAQ), employ the similar reaction mechanism as the DiLeu isobaric tag for the peptide labeling. The side reaction is unavoidable upon isobaric labeling of citrullinated peptides. Due to the advantages of isobaric labeling technique, it is still worth developing an isobaric labeling-based method for quantitative protein citrullination analysis.

Four aliquots of citrullinated peptide standard were labeled by 4-plex DiLeu isobaric labeling reagents and combined for MALDI-MS analysis. As depicted in **Figure S5**, most citrullinated peptide standard was DiLeu tag-labeled and produced a single peak in the full MS spectrum, while peak from the side reaction was still visible. The 4-plex DiLeu tag labeled citrullinated peptide standard was further fragmented by HCD fragmentation technique. Four reporter ions were detected and their relative intensities were consistent with the initial ratio of citrullinated peptide standard in each sample (1:1:1:1).

We next tested the quantitation accuracy of the method integrating DiLeu isobaric tag labeling technique and the biotin thiol tag labeling approach. Citrullinated peptide standard prepared in two different ratios (1:1:1:1 and 1:1:2:5) was labeled by 4-plex DiLeu isobaric tag followed by the biotin thiol tag labeling. As illustrated in **Figure S6a**, the DiLeu tagged citrullinated peptide standard could be effectively labeled by the novel biotin thiol tag, indicating the successful coupling of the DiLeu isobaric labeling technique with the biotin thiol tag labeling approach. Moreover, the citrullinated peptide standard concurrently labeled by DiLeu isobaric tag and biotin thiol tag was chosen for HCD fragmentation. In the tandem MS spectrum, numerous fragment ions yielded from the peptide backbone fragmentation, enabling *de novo* sequencing of citrullinated peptide standard. In addition, 4 reporter ions emerged in the low mass range and their

relative signal intensities were close to the theoretical ratios of citrullinated peptide standard in samples prior to labeling. These results implicated the possibility of the integration of DiLeu isobaric tag labeling technique and the biotin thiol tag labeling approach for comprehensive protein citrullination though systematic evaluation of this approach using biological samples is still required.

Conclusions

In this study, we developed the first method, which combines biotin tag-assisted, MS-based method and the stable isotope dimethyl labeling technique, for the simultaneous qualitative and quantitative analysis of citrullinated proteins in biological samples. The quantitation accuracy of this method was evaluated and demonstrated by using the citrullinated peptide standard and complex biological samples. Furthermore, to increase the multiplexing capability, we performed preliminary study on the application of DiLeu isobaric labeling technique for comprehensive protein citrullination analysis. Reliable quantitation accuracy was obtained though more systematic evaluation is required. In summary, the novel method we developed in this chapter will be useful for the unraveling of physiological role of protein citrullination, especially its role in the onset and progression of diseases.

References

(1) Bizzaro, N.; Mazzanti, G.; Tonutti, E.; Villalta, D.; Tozzoli, R., Diagnostic accuracy of the anti-citrulline antibody assay for rheumatoid arthritis. *Clin Chem* 2001, 47, (6), 1089-93.

(2) De Rycke, L.; Nicholas, A. P.; Cantaert, T.; Kruithof, E.; Echols, J. D.; Vandekerckhove, B.; Veys, E. M.; De Keyser, F.; Baeten, D., Synovial intracellular citrullinated proteins colocalizing with peptidyl arginine deiminase as pathophysiologically relevant antigenic determinants of rheumatoid arthritis-specific humoral autoimmunity. *Arthritis Rheum* 2005, 52, (8), 2323-30.

(3) Raijmakers, R.; van Beers, J. J.; El-Azzouny, M.; Visser, N. F.; Bozic, B.; Pruijn, G. J.; Heck, A. J., Elevated levels of fibrinogen-derived endogenous citrullinated peptides in synovial fluid of rheumatoid arthritis patients. *Arthritis Res Ther* 2012, 14, (3), R114.

(4) Farid, S.; Azizi, G.; Mirshafiey, A., Anti-citrullinated protein antibodies and their clinical utility in rheumatoid arthritis. *Int J Rheum Dis* 2013, 16, (4), 379-86.

(5) Kim, S.; Hwang, J.; Xuan, J.; Jung, Y. H.; Cha, H. S.; Kim, K. H., Global metabolite profiling of synovial fluid for the specific diagnosis of rheumatoid arthritis from other inflammatory arthritis. *PLoS One* 2014, 9, (6), e97501.

(6) Musse, A. A.; Li, Z.; Ackerley, C. A.; Bienzle, D.; Lei, H.; Poma, R.; Harauz, G.; Moscarello, M. A.; Mastronardi, F. G., Peptidylarginine deiminase 2 (PAD2) overexpression in transgenic mice leads to myelin loss in the central nervous system. *Dis Model Mech* 2008, 1, (4-5), 229-40.

(7) Yon, M.; Ackerley, C. A.; Mastronardi, F. G.; Groome, N.; Moscarello, M. A., Identification of a mitogen-activated protein kinase site in human myelin basic protein in situ. *J Neuroimmunol* 1996, 65, (1), 55-9.

(8) Elkon, K. B.; Wiedeman, A., Type I IFN system in the development and manifestations of SLE. *Curr Opin Rheumatol* 2012, 24, (5), 499-505.

(9) Nicholas, A. P., Dual immunofluorescence study of citrullinated proteins in Parkinson diseased substantia nigra. *Neurosci Lett* 2011, 495, (1), 26-9.

(10) Acharya, N. K.; Nagele, E. P.; Han, M.; Coretti, N. J.; DeMarshall, C.; Kosciuk, M. C.; Boulos, P. A.; Nagele, R. G., Neuronal PAD4 expression and protein citrullination: possible role in production of autoantibodies associated with neurodegenerative disease. *J Autoimmun* 2012, 38, (4), 369-80.

(11) Nicholas, A. P., Dual immunofluorescence study of citrullinated proteins in Alzheimer diseased frontal cortex. *Neurosci Lett* 2013, 545, 107-11.

(12) Mohanan, S.; Cherrington, B. D.; Horibata, S.; McElwee, J. L.; Thompson, P. R.; Coonrod, S. A., Potential role of peptidylarginine deiminase enzymes and protein citrullination in cancer pathogenesis. *Biochem Res Int* 2012, 2012, 895343.

(13) Rondas, D.; Crevecoeur, I.; D'Hertog, W.; Ferreira, G. B.; Staes, A.; Garg, A. D.; Eizirik, D. L.; Agostinis, P.; Gevaert, K.; Overbergh, L.; Mathieu, C., Citrullinated glucose-regulated protein 78 is an autoantigen in type 1 diabetes. *Diabetes* 2015, 64, (2), 573-86.

(14) McElwee, J. L.; Mohanan, S.; Griffith, O. L.; Breuer, H. C.; Anguish, L. J.; Cherrington, B. D.; Palmer, A. M.; Howe, L. R.; Subramanian, V.; Causey, C. P.; Thompson, P. R.; Gray, J. W.; Coonrod, S. A., Identification of PADI2 as a potential breast cancer biomarker and therapeutic target. *BMC Cancer* 2012, 12, 500.

(15) Takizawa, Y.; Suzuki, A.; Sawada, T.; Ohsaka, M.; Inoue, T.; Yamada, R.; Yamamoto, K., Citrullinated fibrinogen detected as a soluble citrullinated autoantigen in rheumatoid arthritis synovial fluids. *Ann Rheum Dis* 2006, 65, (8), 1013-20.

(16) Hueber, W.; Hassfeld, W.; Smolen, J. S.; Steiner, G., Sensitivity and specificity of anti-Sa autoantibodies for rheumatoid arthritis. *Rheumatology (Oxford)* 1999, 38, (2), 155-9.

(17) Forslind, K.; Ahlmen, M.; Eberhardt, K.; Hafstrom, I.; Svensson, B.; Group, B. S., Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: role of antibodies to citrullinated peptides (anti-CCP). *Ann Rheum Dis* 2004, 63, (9), 1090-5.

(18) Zendman, A. J.; van Venrooij, W. J.; Pruijn, G. J., Use and significance of anti-CCP autoantibodies in rheumatoid arthritis. *Rheumatology (Oxford)* 2006, 45, (1), 20-5.

(19) Bicker, K. L.; Subramanian, V.; Chumanevich, A. A.; Hofseth, L. J.; Thompson, P. R., Seeing citrulline: development of a phenylglyoxal-based probe to visualize protein citrullination. *J Am Chem Soc* 2012, 134, (41), 17015-8.

(20) Kunieda, K.; Yamauchi, H.; Kawaguchi, M.; Ieda, N.; Nakagawa, H., Development of a fluorescent probe for detection of citrulline based on photo-induced electron transfer. *Bioorg Med Chem Lett* 2018, 28, (5), 969-973.

(21) Zhu, W.; Smith, J. W.; Huang, C. M., Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol* 2010, 2010, 840518.

(22) Frost, D. C.; Greer, T.; Li, L., High-resolution enabled 12-plex DiLeu isobaric tags for quantitative proteomics. *Anal Chem* 2015, 87, (3), 1646-54.

(23) Yu, Q.; Shi, X.; Greer, T.; Lietz, C. B.; Kent, K. C.; Li, L., Evaluation and Application of Dimethylated Amino Acids as Isobaric Tags for Quantitative Proteomics of the TGF-beta/Smad3 Signaling Pathway. *J Proteome Res* 2016, 15, (9), 3420-31.

(24) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002, 1, (5), 376-86.

(25) Fu, Q.; Li, L., De novo sequencing of neuropeptides using reductive isotopic methylation and investigation of ESI QTOF MS/MS fragmentation pattern of neuropeptides with N-terminal dimethylation. *Anal Chem* 2005, 77, (23), 7783-95.

(26) Chaves, D. F.; Carvalho, P. C.; Lima, D. B.; Nicastro, H.; Lorenzeti, F. M.; Siqueira-Filho, M.; Hirabara, S. M.; Alves, P. H.; Moresco, J. J.; Yates, J. R., 3rd; Lancha, A. H., Jr., Comparative proteomic analysis of the aging soleus and extensor digitorum longus rat muscles using TMT labeling and mass spectrometry. *J Proteome Res* 2013, 12, (10), 4532-46.

(27) Griffin, T. J.; Xie, H.; Bandhakavi, S.; Popko, J.; Mohan, A.; Carlis, J. V.; Higgins, L., iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. *J Proteome Res* 2007, 6, (11), 4200-9.

(28) Hao, L.; Thomas, S.; Greer, T.; Vezina, C. M.; Bajpai, S.; Ashok, A.; De Marzo, A. M.; Bieberich, C. J.; Li, L.; Ricke, W. A., Quantitative proteomic analysis of a genetically induced prostate inflammation mouse model via custom 4-plex DiLeu isobaric labeling. *Am J Physiol Renal Physiol* 2019, 316, (6), F1236-F1243.

(29) Liu, S.; Yu, F.; Yang, Z.; Wang, T.; Xiong, H.; Chang, C.; Yu, W.; Li, N., Establishment of Dimethyl Labeling-based Quantitative Acetylproteomics in Arabidopsis. *Mol Cell Proteomics* 2018, 17, (5), 1010-1027.

(30) Gonneaud, A.; Jones, C.; Turgeon, N.; Levesque, D.; Asselin, C.; Boudreau, F.; Boisvert, F. M., A SILAC-Based Method for Quantitative Proteomic Analysis of Intestinal Organoids. *Sci Rep* 2016, 6, 38195.

(31) Shiio, Y.; Aebersold, R., Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. *Nat Protoc* 2006, 1, (1), 139-45.

(32) Chan, J. C. Y.; Zhou, L.; Chan, E. C. Y., The Isotope-Coded Affinity Tag Method for Quantitative Protein Profile Comparison and Relative Quantitation of Cysteine Redox Modifications. *Curr Protoc Protein Sci* 2015, 82, 23 2 1-23 2 19.

(33) Qin, H.; Wang, F.; Zhang, Y.; Hu, Z.; Song, C.; Wu, R.; Ye, M.; Zou, H., Isobaric cross-sequence labeling of peptides by using site-selective N-terminus dimethylation. *Chem Commun (Camb)* 2012, 48, (50), 6265-7.

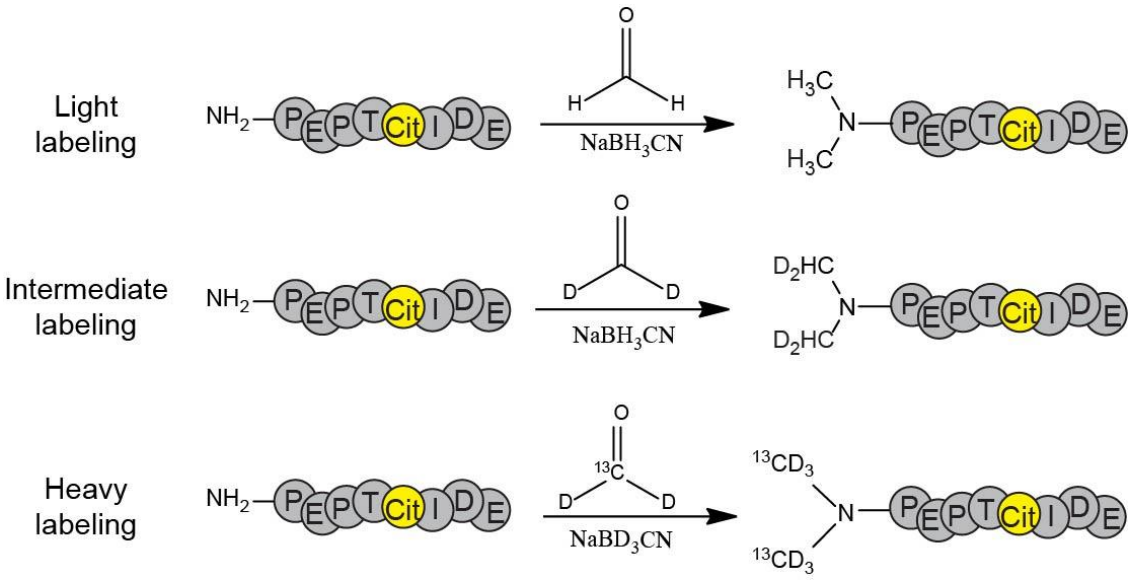


Figure 1. Principle of the stable isotope dimethyl labeling.

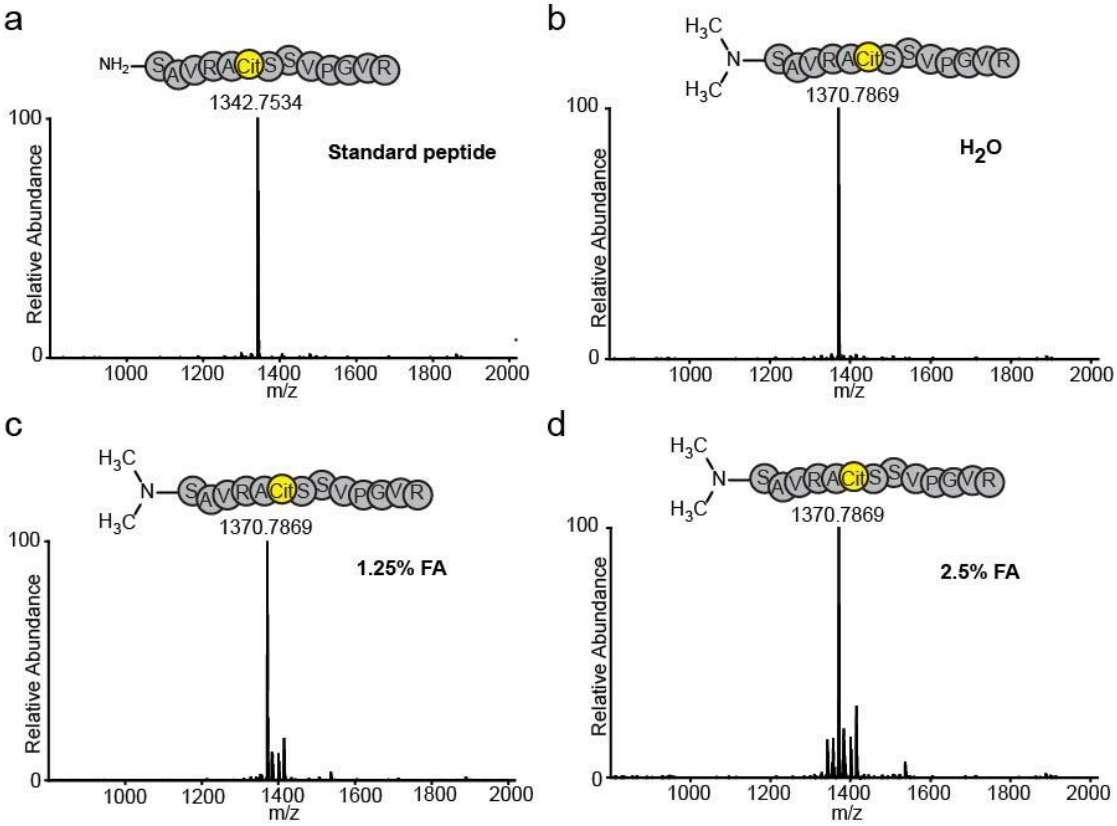


Figure 2. Optimization of buffer acidity for dimethyl labeling of standard citrullinated peptide. a. Standard citrullinated peptide before dimethyl labeling. **b.** Dimethyl labeling of citrullinated peptide standard using H_2O as buffer. **c.** Dimethyl labeling of citrullinated peptide standard using 1.25% FA/H_2O as buffer. **d.** Dimethyl labeling of citrullinated peptide standard using 2.5% FA/H_2O as buffer.

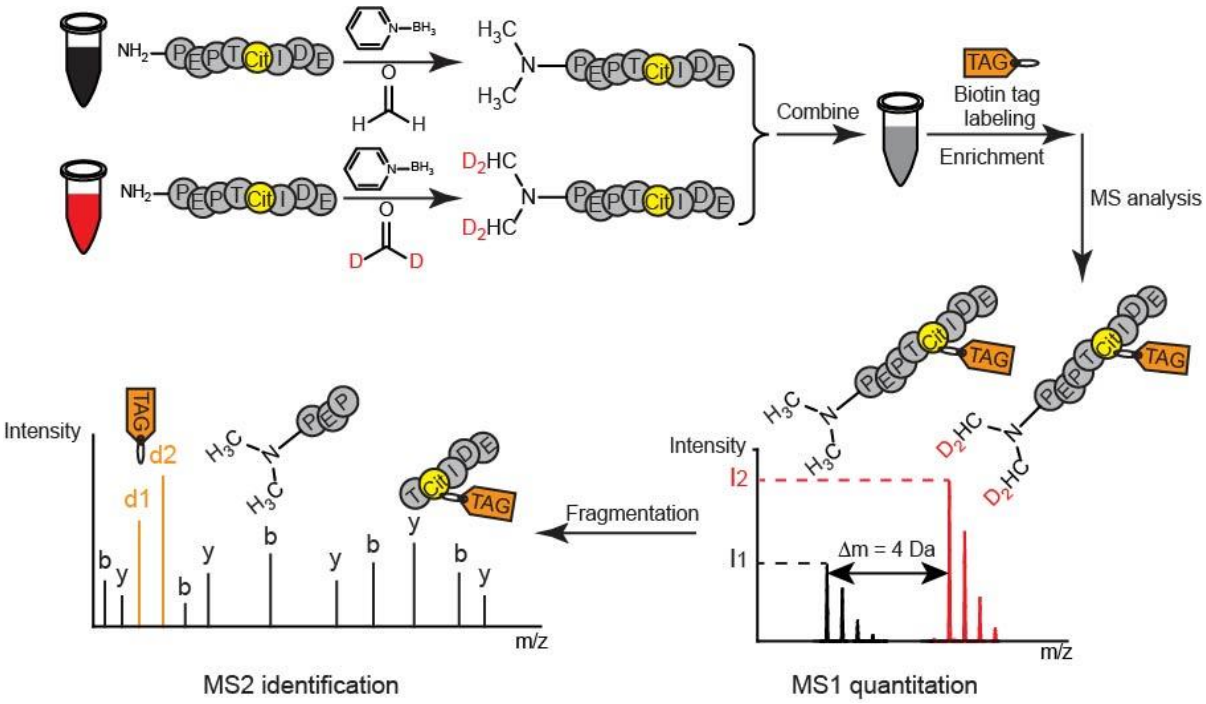


Figure 3. Schematic graph reveals the principle of the new method combining the biotin tag-assisted, MS-based method and stable isotopic dimethylation labeling approach.

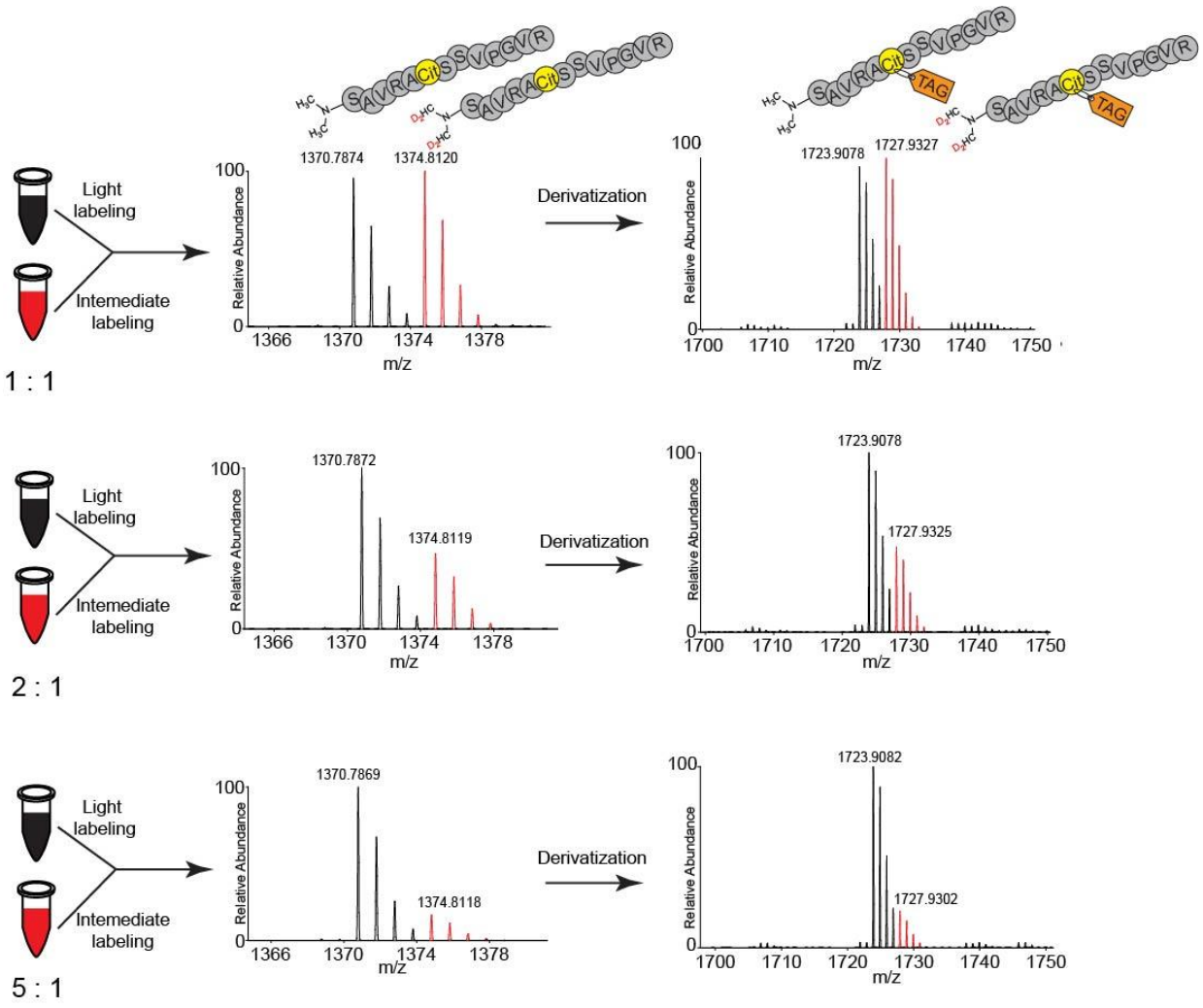


Figure 4. Evaluation of the quantitative method combining dimethyl labeling and biotin thiol tag labeling using the citrullinated peptide standard.

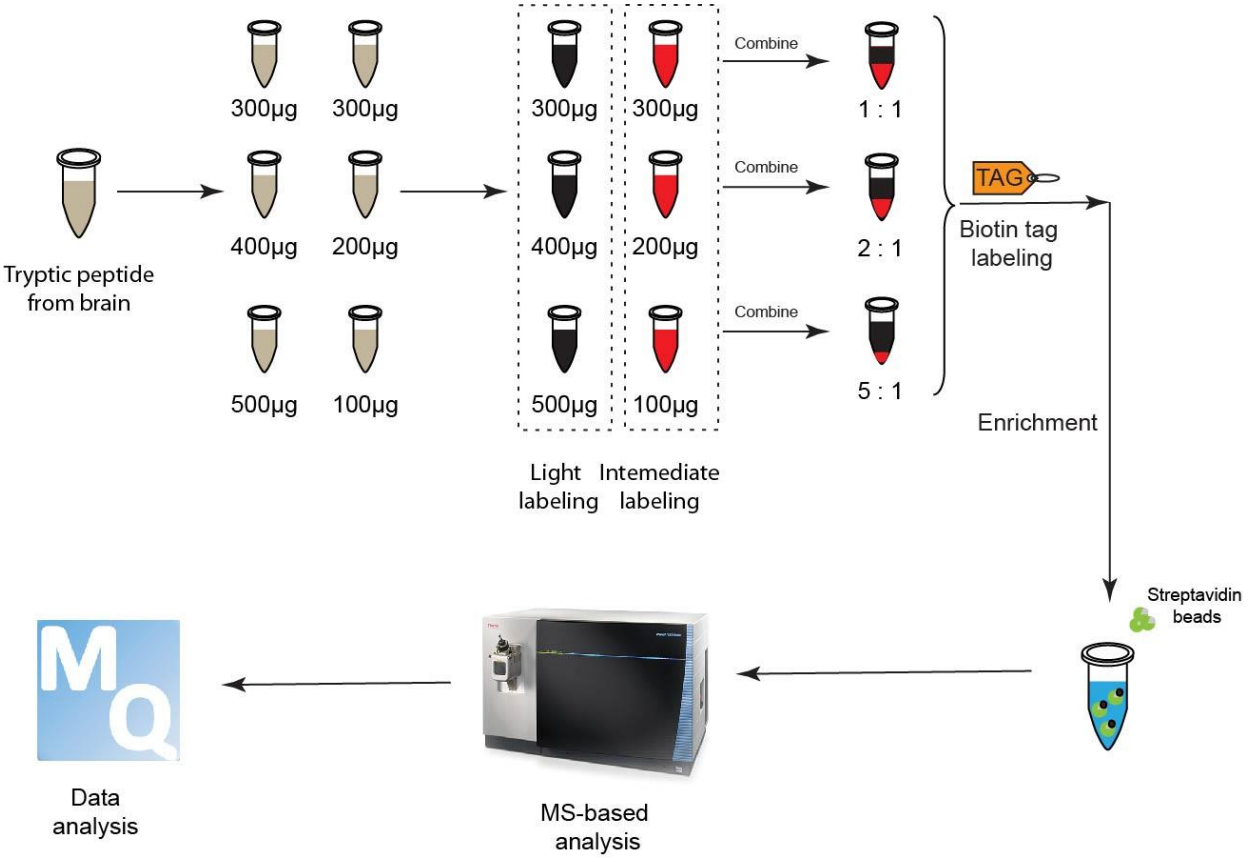


Figure 5. Workflow for testing the performance of the quantitative method combining dimethyl labeling and biotin thiol tag labeling on biological samples.

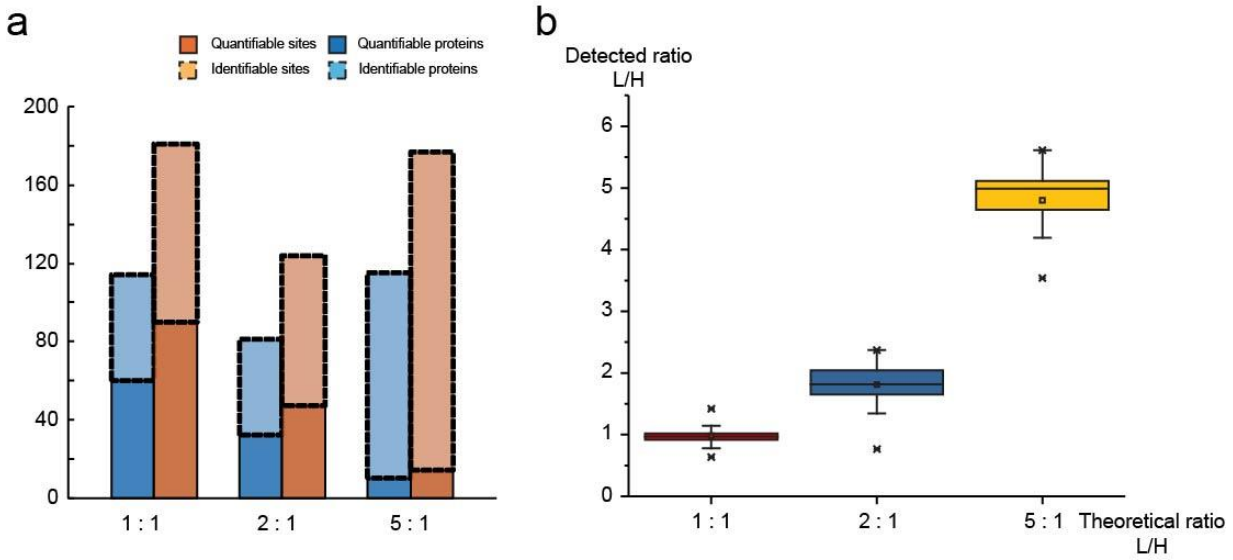
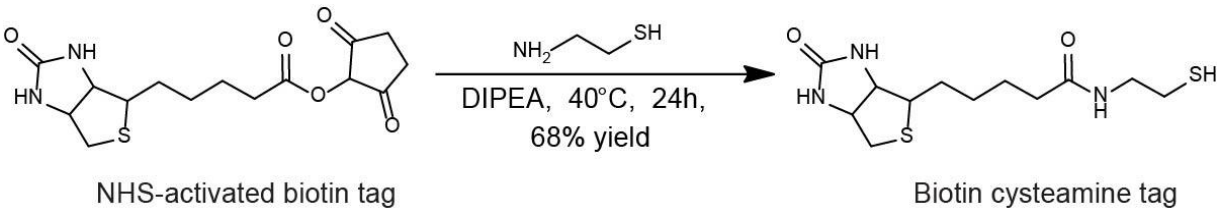


Figure 6 Performance of the novel quantitative method was evaluated using biological samples (a) Bar graph showing the numbers of identified and quantifiable citrullinated proteins and citrullination sites. **(b)** Box-plot showing that reliable relative quantitation accuracy was obtained for the quantifiable citrullinated proteins.

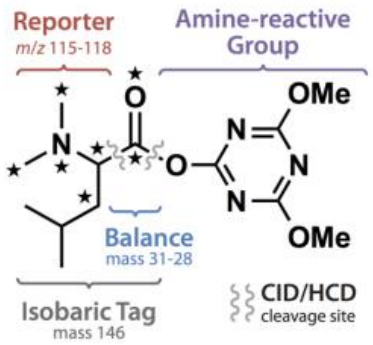
Supplemental Information



NHS-activated biotin tag

Biotin cysteamine tag

Scheme S2. Syntheses of the novel biotin thiol tag.



DiLeu Label	Stable Isotopes		Reporter Mass
	Reporter	Balance	
115a	●	●●	115.12476
115b	●	●●	115.13108
116a	●●	●	116.12812
116b	●●	●	116.13444
116c	●●	●	116.14028
117a	●●●	●	117.13147
117b	●●●	●	117.13731
117c	●●●	●	117.14363
118a	●●●●		118.13483
118b	●●●●		118.14067
118c	●●●●		118.14699
118d	●●●●		118.15283

● ^{13}C ● ^{2}H ● ^{15}N ● ^{18}O

Figure S2. Structure and the composition of 12-plex DiLeu isobaric labeling tag (Cited from Anal. Chem. 2015, 87, 3, 1646-1654).

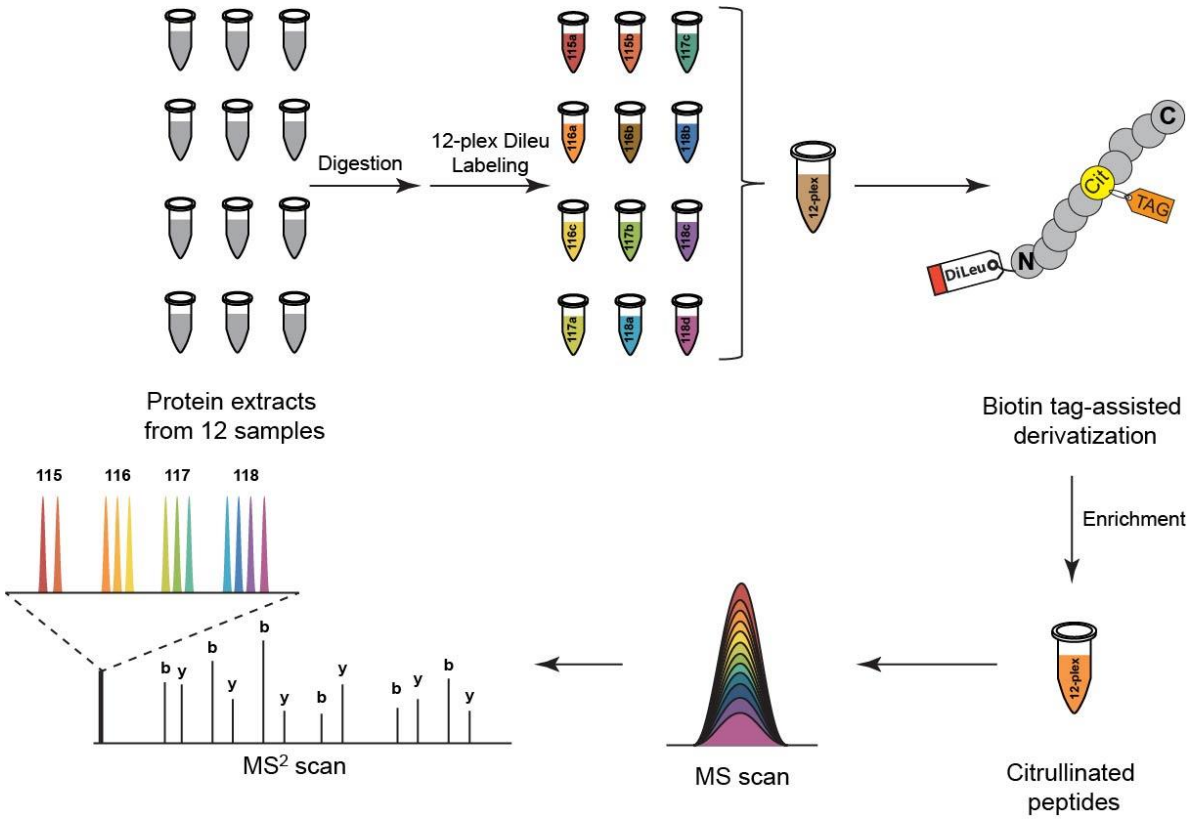


Figure S3. Schematic graph revealed the workflow of the quantitative method integrating the biotin tag-assisted, MS-based method with DiLeu isobaric labeling strategy.

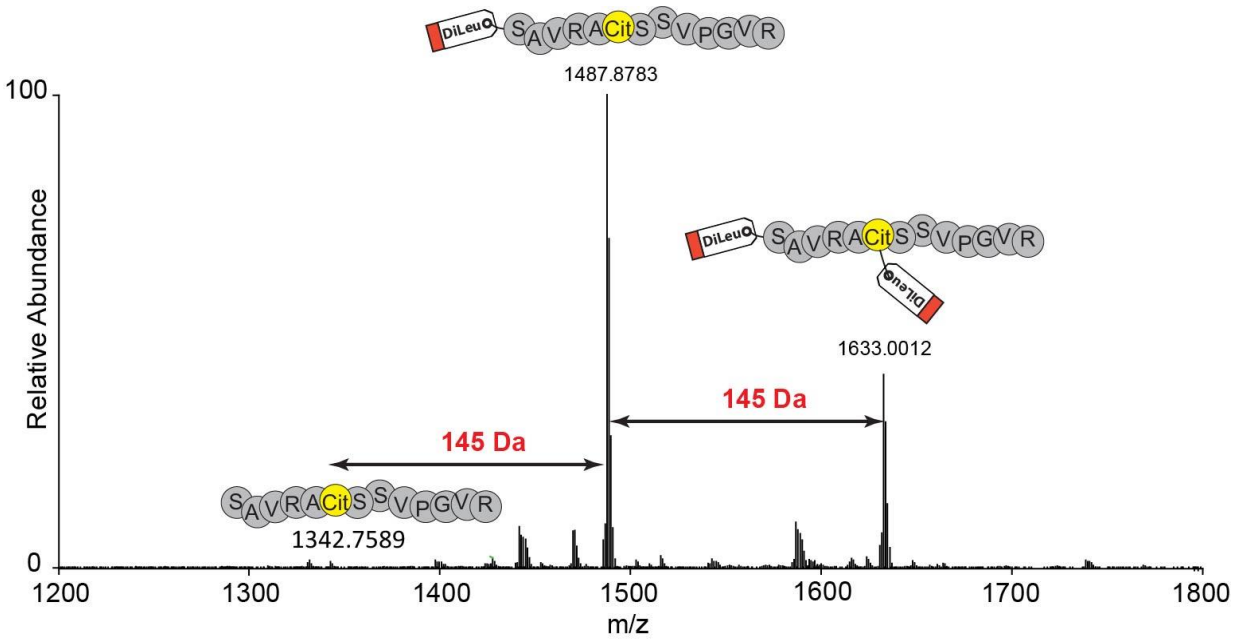


Figure S4. DiLeu isobaric labeling of citrullinated peptide standard.

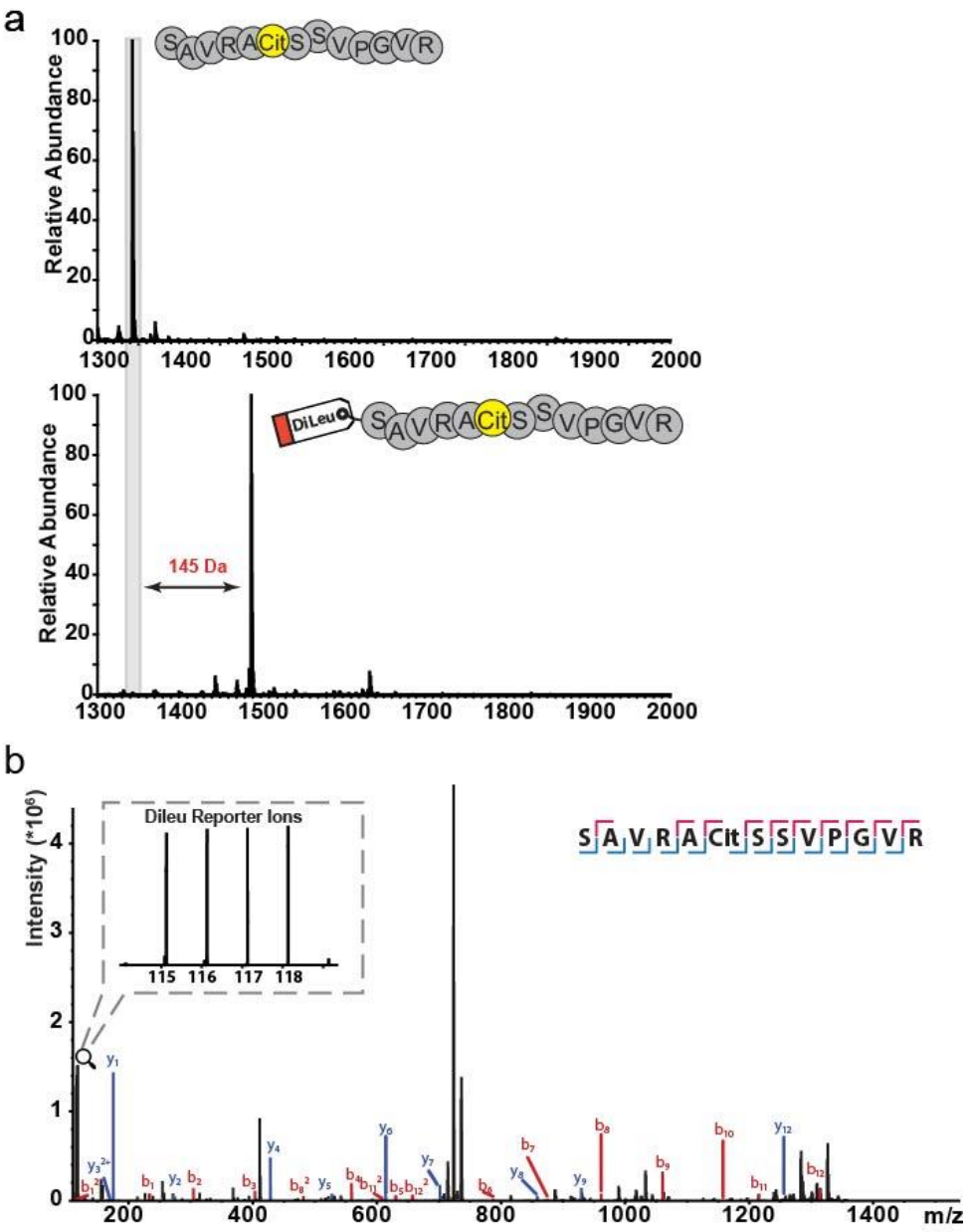


Figure S5. Evaluation of 4-plex DiLeu isobaric labeling of citrullinated peptide standard and the quantitation accuracy. (a) 4-plex DiLeu isobaric labeling of citrullinated peptide standard. **(b)** Tandem MS spectrum of 4-plex DiLeu labeled citrullinated peptide standard fragmented by HCD fragmentation technique.

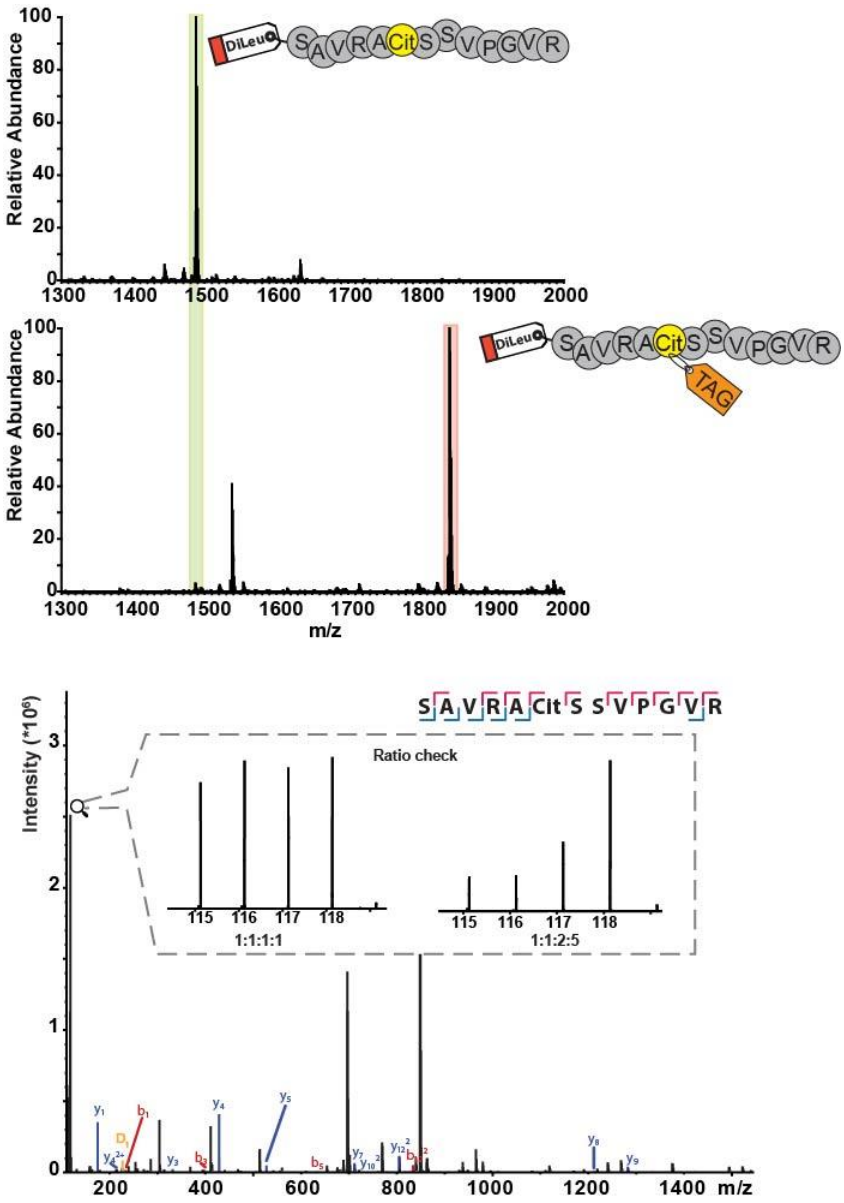


Figure S6. Evaluation of the integration of 4-plex DiLeu isobaric labeling and the biotin thiol tag labeling. (a) Labeling of DiLeu isobaric tag-labeled citrullinated peptide standard by the novel biotin thiol tag. **(b)** Tandem MS spectrum of citrullinated peptide standard concurrently labeled by DiLeu isobaric tag and biotin thiol tag fragmented by HCD fragmentation technique.

Chapter 4

Enabling Global Protein Homocitrullination Analysis by Using a Novel Biotin Tag-assisted Mass Spectrometry-based Method

This project is performed in Prof. Lingjun Li lab. Yatao Shi leads the project and is responsible for the biotin tag design and experiment design as well as MS-based analysis. The biotin tag is synthesized by Dr. Zhengqing Ye from Medicinal Chemistry Center (MCC), School of Pharmacy, University of Wisconsin–Madison.

Abstract

Protein homocitrullination is a protein post-translational modification (PTM) occuring at the ε-amino group of lysine. Different from other lysine-specific protein PTMs, protein homocitrullination is not catalyzed by endogenous enzymes, but regulated by the level of cyanate *in vivo*. Hence, alteration of protein homocitrullination has been reported in many diseases accompanied by the elevated cyanate levels. However, due to the lack of method for large-scale and high-throughput analysis, systematic exploration of this disease-related PTM, including its cellular distribution, physiological functions and roles during the onset and progression of diseases, has been rarely conducted. Here, I developed a novel MS-based method enabling in-depth profiling of homocitrullinated proteins and accurate annotation of homocitrullination site, which has been successfully applied to mouse-tissue specific protein homocitrullination analysis. In total, 840 homocitrullination sites and 599 unique homocitrullinated proteins were identified from mouse brain and multiple organs, most of which were reported for the first time. Our results indicate that homocitrullinated proteins are widely distributed with tissue-specific patterns that may play much more significant roles in cell signaling and pathogenesis than we have anticipated.

Introduction

Protein homocitrullination, also called protein carbamylation, is a type of non-enzymatic protein post-translational modification (PTM).[1] *In vivo*, this modification results from the chemical modification of the free amine group of the lysine residue by cyanate, which produced either from urea or the oxidation of thiocyanate catalyzed by myeloperoxidase (MPO).[2-4] In normal physiological conditions, the level of cyanate is too low to induce substantial protein homocitrullination. However, aberrant protein homocitrullination can occur when cyanate level is significantly enhanced caused by some pathogenic conditions, such as uremia, inflammation, and the exposure to cigarette smoke.[5-8] It has been reported that protein homocitrullination is associated with many diseases, such as rheumatoid arthritis (RA), kidney failure and cardiovascular diseases.[9-15]

Previous studies on protein homocitrullination analysis mainly rely on antibody-based detection techniques. Several anti-homocitrullination antibodies have been reported and have been commercially available for ELISA or WB analysis. For example, the anti-carbamyl-lysine antibody (HRP) (ab175576, abcam) and rabbit anti-carbamyl-lysine (CBL) polyclonal antibody (STA-078, Cell biolabs) are commercially available for these immunoassays.[16-18] However, given the similar structures of protein homocitrullination and citrullination, it has been revealed that some of these anti-homocitrullination antibodies are not specific to homocitrullinated protein detection and can react with citrullinated proteins.[19] In addition, the antibody-based detection methods suffer from low throughput and cross-reactivity, and thereby are not applicable for global

protein homocitrullination analysis of complex biological samples.

Different from the antibody-based detection approaches, mass spectrometry (MS) can effectively distinguish protein homocitrullination from citrullination according to their different location sites and mass shifts.[20-24] Protein homocitrullination occurs at lysine residue resulting in a mass shift of 43 Dalton, while protein citrullination is located at arginine residue with a mass increment of 0.984 Da. Moreover, due to the larger mass shift, protein deamidation and ^{13}C isotopic peaks will not interfere with the identification of homocitrullinated peptides and the accurate annotation of homocitrullination sites. The major challenge for MS-based protein homocitrullination analysis is the lack of effective method for the specific enrichment of homocitrullinated peptides from complex peptide mixture. To deal with this challenge, a biotin-conjugated phenylglyoxal tag was recently reported for the chemical derivatization and enrichment of homocitrullinated peptides for MS analysis.[25] However, its application to real biological sample is limited by the observed strong background reactivities.

In this study, I develop a novel biotin thiol tag for the chemical derivatization and enrichment of homocitrullinated peptides from complex biological samples. The biotin tag labeled homocitrullinated peptides are amenable to MS-based identification. A standard workflow is established for the large-scale analysis of homocitrullinated proteins in biological samples. Further application of this method to mouse tissue-specific homocitrullination analysis enabling the identification of homocitrullinated proteins and homocitrullination sites, most of which are reported for the first time.

Materials and Methods

Materials

Methanol (MeOH), ethanol (EtOH), acetonitrile (ACN), formic acid (FA), trifluoroacetic acid (TFA), ammonium bicarbonate and Tris base were purchased from Fisher Scientific (Pittsburgh, PA). Trypsin and Trypsin/Lys-C Mix were purchased from Promega (Madison, WI). Tris (2-carboxyethyl) phosphine (TCEP) was purchased from EMD Millipore Sigma (Burlington, MA). Biotin-NHS Ester was purchased from Click Chemistry Tools (Scottsdale, AZ). Sodium dodecyl sulfate (SDS), cysteamine and streptavidin agarose were purchased from Sigma-Aldrich (St. Louis, MO). Phosphate-Buffered Saline (PBS) buffer was ordered from Mediatech, Inc (Manassas, VA). Distilled water mentioned in this work was Milli-Q water from a Millipore filtration system (Bedford, MA). All reagents were used without additional purification.

Synthesis of biotin thiol tag

The novel biotin thiol tag was easy to be synthesized in just one step using biotin-NHS ester and cysteamine (**Scheme S1**). After synthesis, the biotin tag powder was reconstituted with methanol/water (50:50, v/v) to a concentration of 10 mg/ml. Then, neutral Tris (2-carboxyethyl) phosphine (TCEP) was added into the biotin tag solution to a final concentration of 10mM for the prevention of biotin thiol tag from oxidation.

Enzymatic protein digestion

A fresh mouse brain was homogenized with a probe ultrasonicator in 4% Sodium dodecyl sulfate (SDS)/50mM Tris base buffer (~pH 8, adjusted with HCl). After homogenization, the

supernatant was collected after centrifugation at 16,000 rcf for 15 mins. Protein concentration was determined by BCA protein assay reagent (Thermo Scientific, Fair Lawn, NJ). Each 800µg mouse protein extract was reduced by 10mM dithiothreitol (DTT) for 30 min at room temperature and then alkylated with 50mM iodoacetic acid (IAA) for another 30 min in the dark. Protein was further precipitated by adding cold acetone (-20°C) to a final concentration of 80% (v/v). After incubation in -20°C overnight, the sample solution was centrifugated at 16,000 rcf for 15 min and the supernatant was discarded. To completely remove SDS, the pellet was washed with cold 80% acetone for another two times, and then dried in the air for 12 mins. The dried pellet was reconstituted with 150µl 5 M guanidine hydrochloride/50mM Tris-base buffer (pH 8, adjusted with HCl). After reconstitution, the buffer solution was diluted tenfold with 50mM Tris-base buffer to reduce the concentration of guanidine HCl to 0.5 M.

Reconstituted protein was further digested by three different enzymes (Promega, Madison, WI), including Lys-C, Trypsin gold and Trypsin/Lys-C Mix, respectively. Enzyme (Protein:enzyme, 100:1, w/w) was added to sample and incubated at 37°C for 6 h. The secondary digestion was performed by adding the same amount of enzyme and incubating samples at 37°C for another 12 h. The digestion process was quenched by adding 10% TFA to reduce the pH to <3. Digested peptides were desalted with C_{18} columns (Sep-Pak, waters, Milford, MA) and eluted with 80% ACN/0.2% FA, which were dried out with SpeedVac. Peptide concentration was determined by colorimetric peptide assay (23275, Thermo Scientific, Fair Lawn, NJ). Each 400µg of peptide was transferred to an Eppendorf tube and dried for further analysis.

Chemical derivatization

Thirty micro liter of biotin thiol tag (10 mg/ml) solution was added to each sample tube containing 400µg peptide and dried with SpeedVac. Then, the biotin thiol tag and peptide were reconstituted in 40µl 12.5% TFA solution followed by adding 10µl of 2,3-butanedione solution (100mM 12.5% TFA solution). The reactants were incubated at 37°C for 6h with shaking in the dark and were further dried out with SpeedVac.

To remove excess reactants, strong cation exchange (SCX) chromatography was performed using TopTips (TT200SEA, Poly LC) containing PolySULFOETHYL A beads. Briefly, SCX beads were pre-equilibrated with 100µl loading buffer containing 50% ACN/0.2% FA/10mM ammonium formate for three times. The dried peptide mixture dried above was then reconstituted in 200µl loading buffer and applied to the SCX beads twice followed by continuous washing of the beads with 100µl loading buffer for 10 times. All centrifugation steps were performed at 2,000 rpm for 2 min. Peptides were finally eluted with 150 µl 25% ACN/0.4M ammonium formate and dried out in SpeedVac. A secondary drying process was conducted to reduce the concentration of ammonium formate in samples by adding another 400µl water to each sample and drying out in SpeedVac.

Enrichment of biotin tag-modified homocitrullinated peptides

To enable complete dissolution of peptides recovered from SCX, 300µl 50% ACN/H_2O was added to each sample. After brief vortex and centrifugation, each sample was dried in SpeedVac to less than 100µl and further diluted with 900µl PBS 1× buffer.

The enrichment process was performed as previously described with some modifications.[26] Briefly, 75µl streptavidin beads were pre-washed with 1ml PBS 1× buffer for 5 times. The peptide

solution prepared above was loaded onto the pre-washed streptavidin agarose and incubated at room temperature for 2 h with rotation. The beads were subsequently washed 4 times each with 1ml PBS 1× buffer, 1ml 5 % ACN/PBS 1× buffer and 1ml water. The bound peptides were finally released with 300µl 80%ACN/H_2O contain 0.2% TFA and 0.1% FA for four times. The first release was performed in room temperature for 5 min, while other three release processes were conducted at 95°C for 5 min with shaking. The eluents were combined and dried in the SpeedVac. Enriched peptides were desalted with C_{18} columns (Zip-tips) and eluted in 100 µl 80% ACN/0.2% FA solution, which was dried in the vacuum to almost dry.

Enzymatic digestion of protein extract from mouse tissues

Mouse tissues from three mice were collected and stored at -80°C refrigerator until use. Tissues were homogenized with a probe ultrasonicator in 4% Sodium dodecyl sulfate (SDS)/50mM Tris base buffer (~pH 8, adjusted with HCl). Protein extract was digested by Trypsin/Lys-C Mix and desalted following the procedure previously described. For each tissue, 400µg of peptide was used for global protein homocitrullination analysis. The chemical derivatization, SCX cleaning and streptavidin enrichment processes were performed as previously described. Enriched homocitrullinated peptides from each tissue was finally analyzed by MS using stepped HCD fragmentation technique.

Mass spectrometry (MS) analysis

MS analysis was performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer coupled with Dionex UltiMate 3000 UPLC system (Thermo Fisher Scientific, San Jose, CA). The mobile phase was composed of 0.1% formic acid in water (A) and 0.1% formic acid in ACN (B).

The flow rate was set as 300 nl/min. Each sample was resuspended in 15 µl 0.1% formic acid 3% ACN in water and underwent three LC-MS/MS run. For each run, 2µl of sample was loaded onto a customized C_{18} column filled with 1.7µm Ethylene Bridged Hybrid packing materials (130 Å, Waters) and separated with the following gradient: 3% B for the first 18.3 min upon sample trapping; 3%-30% B for 18.3-120 min; 30%-75% B for 120-120.5 min; 75% B for 120.5-130 min; 75%-95% B for 130.0-130.5 min; 95% B for 130.5-140.0 min; 95%-3% B for 140.0-140.5 min and equilibrated at 3% B for 15 min. MS data was acquired in the positive ion mode with the spray voltage of 2 kV. MS^1 spectra were collected from m/z 350-1500 by orbitrap at a resolution of 60 K. Automatic gain control (AGC) target and maximum injection time was set to $2e^5$ and 100 ms, respectively. Ions with the charge of 2-6 were included for screening, while a dynamic exclusion time of 45s was used to avoid repeatedly sequencing the same precursor ion in a short time. For the MS^2 fragmentation, the top 15 precursors were selected for stepped HCD fragmentation with an isolation window of 1 Da, a resolving power of 15K and AGC target of 5×10^4.

Database search

MS data from mouse tissues was searched against mouse protein database (Downloaded from Uniprot website on 12/30/2018) using the software of MaxQuant. The search parameters were defined as follows: 20 p.p.m and 4.5 ppm were set as the first search peptide tolerance and the main search peptide tolerance, respectively; 2 ppm was used for the isotope match tolerance. a minimum of six amino-acid peptide length and up to three missed cleavages were allowed for peptide identification; carbamidomethylation (C, +57.02146 Da) was set as a static modification, while oxidation (M, +15.995 Da) and biotin tag-labeled Homocitrullination (K, +396.129 Da) were

defined as dynamic modifications. A neutral loss of biotin tag (303.1075 Da) and two diagnostic ions of 227.0848 Da and 304.1147 Da were included in the search. A maximum of three modifications were allowed for each peptide. The 1% false-discovery rate (FDR) was applied to filter both peptide and protein identification. Peptides that were found as reverse or potential contaminant hits were filtered out and homocitrullination site localization probability threshold was set to 0.75. Gene ontology (GO) enrichment analysis of identified homocitrullinated proteins was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8.

Results and Discussion

Homocitrullinated proteins are low-abundance proteins in biological samples. Due to the lack of specific enrichment method, it is still a great challenge to perform MS-based large-scale and high throughput analysis of homocitrullinated proteins in biological samples, limiting our understanding about this important protein PTM. To fill in this knowledge gap, in this study, I developed a novel MS-based method, enabling the specific enrichment and large-scale characterization of homocitrullinated proteins in biological samples.

Method development for protein homocitrullination analysis

Protein homocitrullination represents the chemical modification of the free amine group at the lysine residue by cyanate, which produced either from urea or the conversion from thiocyanate (**Figure 1**). Protein homocitrullination results in the formation of a ureido group at the side chain of lysine residue, while protein citrullination leads to a ureido group at the side chain of arginine

residue. Hence, citrullinated and homocitrullinated proteins are characterized by the presence of one or more ureido groups, which explains the poor specificity of antibody-dependent methods for protein citrullination and homocitrullination detection.

We previously found that our customized biotin thiol tag, together with 2, 3-butanedione, could specifically modify the ureido group of citrullinated peptides in low pH aqueous solution. Therefore, we propose that homocitrullinated peptide can also be chemically derivatized by the novel biotin thiol tag and 2, 3-butanedione as shown in **Figure 2**. After biotin tag labeling, a biotin head group could be specifically induced into homocitrullinated peptides in biological samples, providing an opportunity to enrich homocitrullinated peptides by using streptavidin beads.

We directly tested this strategy using the tryptic peptides digested from mouse brain. A simple workflow is depicted in **Figure 3**, which is the same with that of protein citrullination analysis. From 400μg tryptic peptides from mouse brain, we successfully identified 29 unique homocitrullinated proteins and 31 homocitrullination sites with high confidence. A representative of identified homocitrullinated peptide was shown in **Figure 4**. Two diagnostic ions at m/z 227 and 304 produced from the fragmentation of biotin tag could be clearly observed. Furthermore, numerous b and y fragment ions were used to *de novo* sequencing the peptide, while fragment ions containing the intact biotin tag modification group or fragment ions resulting from the neutral loss of biotin tag enabled the accurate annotation of homocitrullination sites.

Comparison of enzymes used for homocitrullination analysis

Three different enzymes, including trypsin gold, trypsin/Lys-C mixture and LysC, were applied to prepare peptide mixture for protein homocitrullination analysis, respectively. As shown

in **Figure 5** and **Table S1**, a large number of unique homocitrullinated peptides and were identified from each 400 μg peptides prepared by different enzymes. Most of the identified homocitrullination sites were confidently annotated with a localization probability more than 0.75.

Similar with protein citrullination, a common understanding about protein homocitrullination is that trypsin and Lys-C cannot cleave after the homocitrulline residue.[27] However, we identified many high-confident homocirullinated peptides with C-terminal homocitrulline residue, suggesting that some homocitrullination sites are trypsin or Lys-C cleavable. A homocitrullinated peptide with C-terminal homocitrulline residue, which was identified from trypsin/Lys-C digested sample, was shown as an example in **Figure S1**. In total, about 50% of homocitrullinated peptides identified from the trypsin-digested or trypsin/Lys-C-digested sample, had the homocitrulline residue at the C-terminus, while less homocitrullinated peptides were identified with C-terminal homocitrulline residue from the Lys C-digested sample (**Figure 6**). This result suggested that homocitrulline residue was more prone to trypsin than Lys-C.

Mouse-tissue specific protein homocitrullination analysis

I further applied the novel method to mouse tissue-specific protein homocitrullination analysis. In total, I was able to identify 840 homocitrullination sites and 599 unique homocitrullinated proteins from mouse brain and Organs with high confidence (**Fig. 7**). As shown in **Figure S2**, the overlap of homocitrullinated proteins identified from the brain and Organs were poor. Though fewer homocitrullinated proteins and homocitrullination sites were identified from mouse S.cortex in comparison to other four brain regions, no apparent distribution pattern was

observed. In contrast, the distribution patterns of protein homocitrullination in organs showed significant difference. More homocitrullinated proteins and homocitrullination sites were identified from heart and spleen than other organs.

Enabled by the novel method reported here, I identified numerous homocitrullination sites and homocitrullinated proteins that have not been reported. For example, we identified a new homocitrullination site located at the K78 of mouse histone H4. A high-quality tandem MS spectrum is shown in **Fig. 8**. The homocitrullination site was mis-cleaved by either trypsin or Lys-C. Two diagnostic ions at *m/z* 227 and 304 are observed in the tandem spectrum, demonstrating the biotin tag labeling of this homocitrullinated peptide. A series of y fragment ions are annotated based on their accurate mass, enabling the de novo peptide sequencing and the precise assignment of the homocitrullination modification at K78.

In UniProt database, five types of lysine modifications have been marked at the K78 of mouse histone H4. 2-hydroxyisobutyrylation and succinylation have been reported, while another three modifications, including butyrylation, propionylation and succinylation, are asserted by similarity. However, the occurrence of protein homocitrullination at the K78 of mouse histone H4 was reported for the first time in our study. This result suggests that protein homocitrullination may have potential crosstalk with other lysine modifications. Although several protein citrullination sites on histones have been explored, the protein homocitrullination on histone is rarely reported.[17] In this study, we totally identified 9 new homocitrullination sites on different types of histones (**Fig S3**), which is crucial for the further study in terms of the epigenetic role of this PTM.

Besides nuclear proteins, we detected some important homocitrullinated proteins in organelles. For example, we identified a high-confident homocitrullination site located at the K165 of malate dehydrogenase in mitochondria, which is an important protein enzyme involved in many biological processes, such as tricarboxylic acid (TCA) cycle, malate metabolic process and internal protein amino acid acetylation. (**Fig. 9**). However, the functional role of protein homocitrullination in mitochondria have not been investigated.

Conclusions

In this study, we developed the first biotin tag-assisted, MS-based method for the large-scale and high throughput identification of homocitrullinated proteins in complex biological samples. The biological application of this novel method indicated that some homocitrullination sites could be cleaved by the commonly used enzymes, such as trypsin and Lys-C although different cleavage efficiency was observed. With the novel method, numerous homocitrullinated protein substrates and homocitrullination sites were identified for the first time. We hope this method could provide a viable solution to global analysis of protein homocitrullination.

Acknowledgements

This research is supported in part by the National Institutes of Health through grants-R21 AG060242, U01CA231081, RF1AG052324, and R01 DK071801. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531 to LL) and the University of Wisconsin-Madison, Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation. LL acknowledges a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

References

(1) Wang, Z.; Nicholls, S. J.; Rodriguez, E. R.; Kummu, O.; Horkko, S.; Barnard, J.; Reynolds, W. F.; Topol, E. J.; DiDonato, J. A.; Hazen, S. L., Protein carbamylation links inflammation, smoking, uremia and atherogenesis. *Nat Med* 2007, 13, (10), 1176-84.

(2) Kraus, L. M.; Kraus, A. P., Jr., Carbamoylation of amino acids and proteins in uremia. *Kidney Int Suppl* 2001, 78, S102-7.

(3) Wever, R.; Kast, W. M.; Kasinoedin, J. H.; Boelens, R., The peroxidation of thiocyanate catalysed by myeloperoxidase and lactoperoxidase. *Biochim Biophys Acta* 1982, 709, (2), 212-9.

(4) Sirpal, S., Myeloperoxidase-mediated lipoprotein carbamylation as a mechanistic pathway for atherosclerotic vascular disease. *Clin Sci (Lond)* 2009, 116, (9), 681-95.

(5) Stim, J.; Shaykh, M.; Anwar, F.; Ansari, A.; Arruda, J. A.; Dunea, G., Factors determining hemoglobin carbamylation in renal failure. *Kidney Int* 1995, 48, (5), 1605-10.

(6) Lhotta, K.; Schlogl, A.; Uring-Lambert, B.; Kronenberg, F.; Konig, P., Complement C4 phenotypes in patients with end-stage renal disease. *Nephron* 1996, 72, (3), 442-6.1.

(7) Mun, K. C.; Golper, T. A., Impaired biological activity of erythropoietin by cyanate carbamylation. *Blood Purif* 2000, 18, (1), 13-7.

(8) Loke, W. M.; Lam, K. M.; Chong, W. L.; Chew, S. E.; Quek, A. M.; Lim, E.; Seet, R. C., Products of 5-lipoxygenase and myeloperoxidase activities are increased in young male cigarette smokers. *Free Radic Res* 2012, 46, (10), 1230-7.

(9) Shi, J.; Knevel, R.; Suwannalai, P.; van der Linden, M. P.; Janssen, G. M.; van Veelen, P. A.; Levarht, N. E.; van der Helm-van Mil, A. H.; Cerami, A.; Huizinga, T. W.; Toes, R. E.; Trouw, L. A., Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. *Proc Natl Acad Sci* U S A 2011, 108, (42), 17372-7.

(10) Ok, E.; Basnakian, A. G.; Apostolov, E. O.; Barri, Y. M.; Shah, S. V., Carbamylated low-density lipoprotein induces death of endothelial cells: a link to atherosclerosis in patients with kidney disease. *Kidney Int* 2005, 68, (1), 173-8.

(11) Kalim, S.; Karumanchi, S. A.; Thadhani, R. I.; Berg, A. H., Protein carbamylation in kidney disease: pathogenesis and clinical implications. *Am J Kidney Dis* 2014, 64, (5), 793-803.

(12) Asselbergs, F. W.; Reynolds, W. F.; Cohen-Tervaert, J. W.; Jessurun, G. A.; Tio, R. A., Myeloperoxidase polymorphism related to cardiovascular events in coronary artery disease. *Am J Med* 2004, 116, (6), 429-30.

(13) Baldus, S.; Heitzer, T.; Eiserich, J. P.; Lau, D.; Mollnau, H.; Ortak, M.; Petri, S.; Goldmann, B.; Duchstein, H. J.; Berger, J.; Helmchen, U.; Freeman, B. A.; Meinertz, T.; Munzel, T., Myeloperoxidase enhances nitric oxide catabolism during myocardial ischemia and reperfusion. *Free Radic Biol Med* 2004, 37, (6), 902-11.

(14) Zheng, L.; Nukuna, B.; Brennan, M. L.; Sun, M.; Goormastic, M.; Settle, M.; Schmitt, D.; Fu, X.; Thomson, L.; Fox, P. L.; Ischiropoulos, H.; Smith, J. D.; Kinter, M.; Hazen, S. L., Apolipoprotein A-I is a selective target for myeloperoxidase-catalyzed oxidation and functional impairment in subjects with cardiovascular disease. *J Clin Invest* 2004, 114, (4), 529-41.

(15) Thukkani, A. K.; McHowat, J.; Hsu, F. F.; Brennan, M. L.; Hazen, S. L.; Ford, D. A., Identification of alpha-chloro fatty aldehydes and unsaturated lysophosphatidylcholine molecular species in human atherosclerotic lesions. *Circulation* 2003, 108, (25), 3128-33.

(16) Verheul, M. K.; van Veelen, P. A.; van Delft, M. A. M.; de Ru, A.; Janssen, G. M. C.; Rispens, T.; Toes, R. E. M.; Trouw, L. A., Pitfalls in the detection of citrullination and carbamylation. *Autoimmun Rev* 2018, 17, (2), 136-141.

(17) Joshi, A. D.; Mustafa, M. G.; Lichti, C. F.; Elferink, C. J., Homocitrullination Is a Novel Histone H1 Epigenetic Mark Dependent on Aryl Hydrocarbon Receptor Recruitment of Carbamoyl Phosphate Synthase 1. *J Biol Chem* 2015, 290, (46), 27767-78.

(18) Koro, C.; Bielecka, E.; Dahl-Knudsen, A.; Enghild, J. J.; Scavenius, C.; Brun, J. G.; Binder, V.; Hellvard, A.; Bergum, B.; Jonsson, R.; Potempa, J.; Blom, A. M.; Mydel, P., Carbamylation of immunoglobulin abrogates activation of the classical complement pathway. *Eur J Immunol* 2014, 44, (11), 3403-12.

(19) Shi, J.; Willemze, A.; Janssen, G. M.; van Veelen, P. A.; Drijfhout, J. W.; Cerami, A.; Huizinga, T. W.; Trouw, L. A.; Toes, R. E., Recognition of citrullinated and carbamylated proteins by human antibodies: specificity, cross-reactivity and the 'AMC-Senshu' method. *Ann Rheum Dis* 2013, 72, (1), 148-50.

(20) Lugli, E. B.; Correia, R. E.; Fischer, R.; Lundberg, K.; Bracke, K. R.; Montgomery, A. B.; Kessler, B. M.; Brusselle, G. G.; Venables, P. J., Expression of citrulline and homocitrulline residues in the lungs of non-smokers and smokers: implications for autoimmunity in rheumatoid arthritis. *Arthritis Res Ther* 2015, 17, 9.

(21) Scinocca, M.; Bell, D. A.; Racape, M.; Joseph, R.; Shaw, G.; McCormick, J. K.; Gladman, D. D.; Pope, J.; Barra, L.; Cairns, E., Antihomocitrullinated fibrinogen antibodies are specific to rheumatoid arthritis and frequently bind citrullinated proteins/peptides. *J Rheumatol* 2014, 41, (2), 270-9.

(22) Van Driessche, G.; Vandenberghe, I.; Jacquemotte, F.; Devreese, B.; Van Beeumen, J. J., Mass spectrometric identification of in vivo carbamylation of the amino terminus of Ectothiorhodospira mobilis high-potential iron-sulfur protein, isozyme 1. *J Mass Spectrom* 2002, 37, (8), 858-66.

(23) Lapko, V. N.; Smith, D. L.; Smith, J. B., In vivo carbamylation and acetylation of water-soluble human lens alphaB-crystallin lysine 92. *Protein Sci* 2001, 10, (6), 1130-6.

(24) Mori, D.; Matsui, I.; Shimomura, A.; Hashimoto, N.; Matsumoto, A.; Shimada, K.; Yamaguchi, S.; Oka, T.; Kubota, K.; Yonemoto, S.; Sakaguchi, Y.; Takahashi, A.; Shintani, Y.; Takashima, S.; Takabatake, Y.; Hamano, T.; Isaka, Y., Protein carbamylation exacerbates vascular calcification. *Kidney Int* 2018, 94, (1), 72-90

(25) Lewallen, D. M.; Bicker, K. L.; Subramanian, V.; Clancy, K. W.; Slade, D. J.; Martell, J.; Dreyton, C. J.; Sokolove, J.; Weerapana, E.; Thompson, P. R., Chemical Proteomic Platform To Identify Citrullinated Proteins. *ACS Chem Biol* 2015, 10, (11), 2520-8.

(26) Schiapparelli, L. M.; McClatchy, D. B.; Liu, H. H.; Sharma, P.; Yates, J. R., 3rd; Cline, H. T., Direct detection of biotinylated proteins by mass spectrometry. *J Proteome Res* 2014, 13, (9), 3966-78.

(27) Ruel, C.; Morani, M.; Bruneel, A.; Junot, C.; Taverna, M.; Fenaille, F.; Tran, N. T., A capillary zone electrophoresis method for detection of Apolipoprotein C-III glycoforms and other related

artifactually modified species. *J Chromatogr A* 2018, 1532, 238-245.

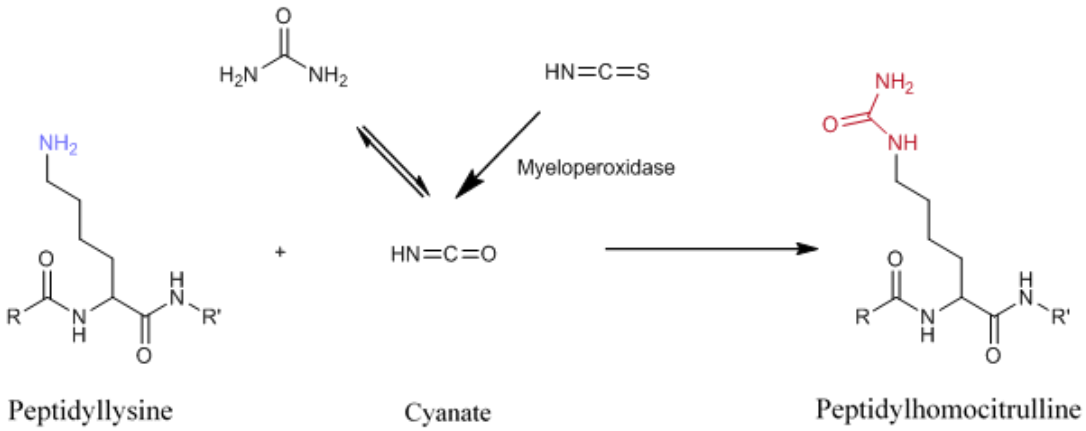


Figure 1. The formation of protein homocitrullination *in vivo*.

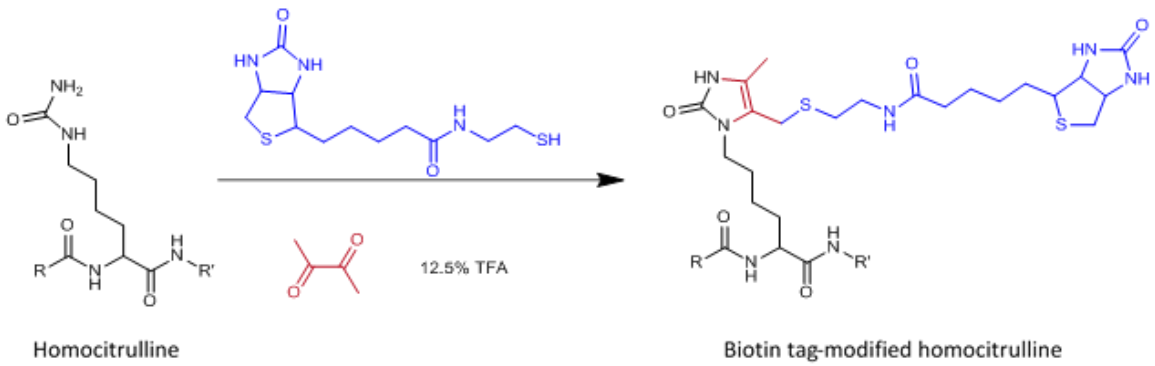


Figure 2. Chemical derivatization of homocitrullinated peptides with 2,3-butanedione and the novel biotin thiol tag in low pH aqueous solution.

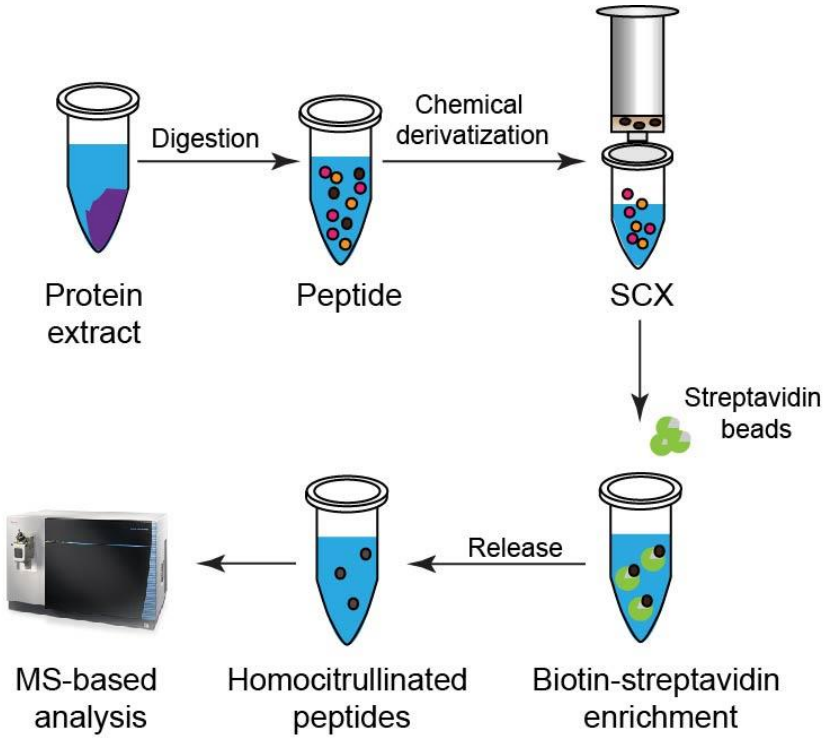


Figure 3. Workflow for the large-scale analysis of homocitrullinated proteins in biological samples.

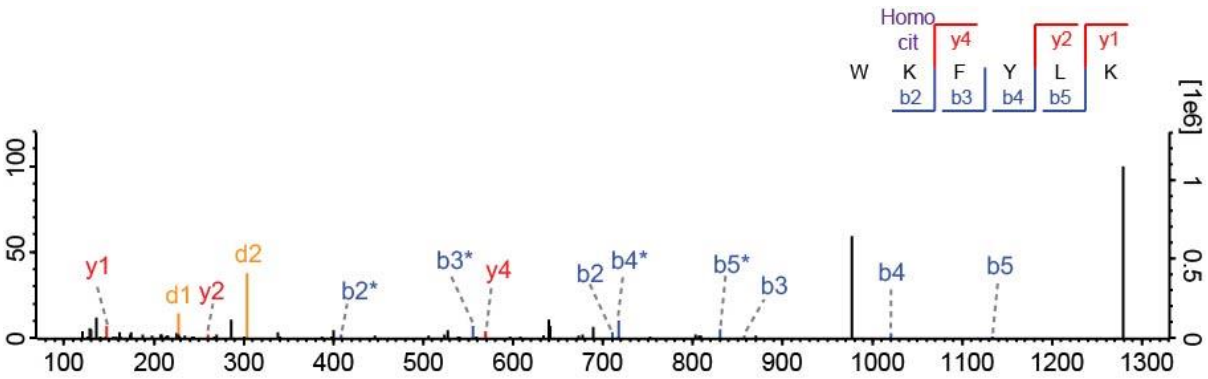


Figure 4. Tandem MS spectrum of the unique peptide WKFYLK after fragmentation by stepped HCD. Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. * denotes fragment ions resulting from the neutral loss of biotin tag.

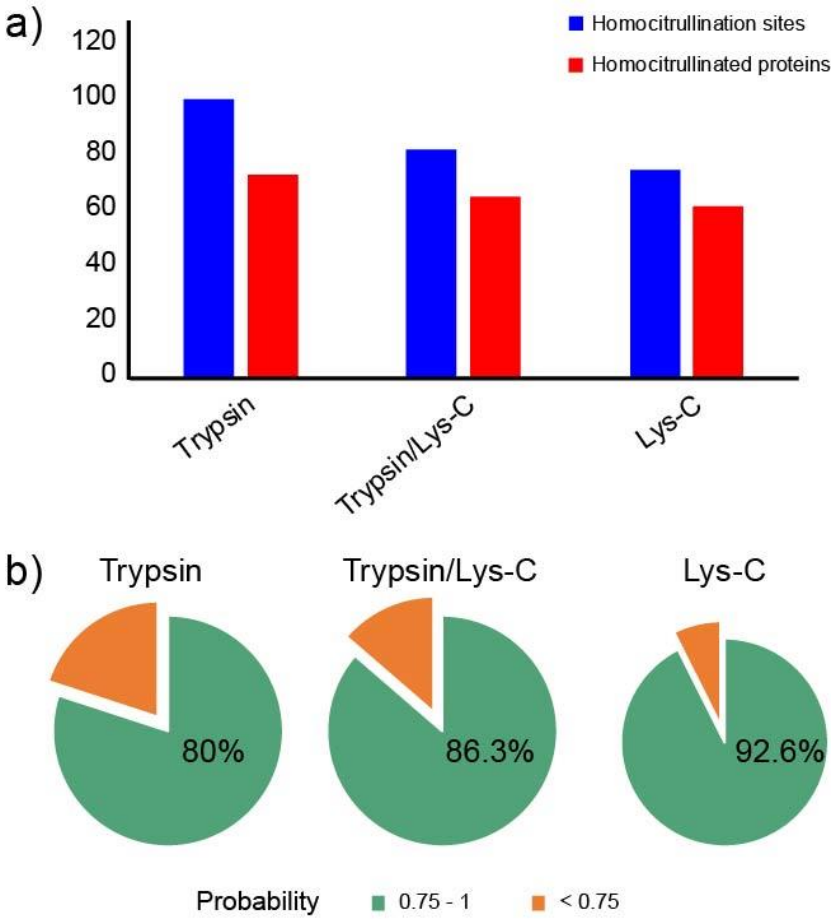


Figure 5. Number of unique homocitrullinated peptides and homocitrullination sites (a). Localization probability pie chart of identified homocitrullination sites (b).

a)

	Total identified homocit sites	Cleaved homocit sites	Misscleaved homocit sites
Trypsin	100	47	53
Trypsin/Lys-C	82	39	43
Lys-C	75	16	59

b)

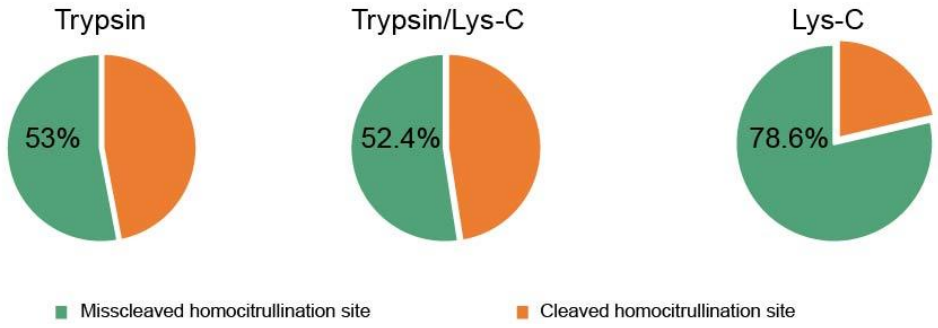
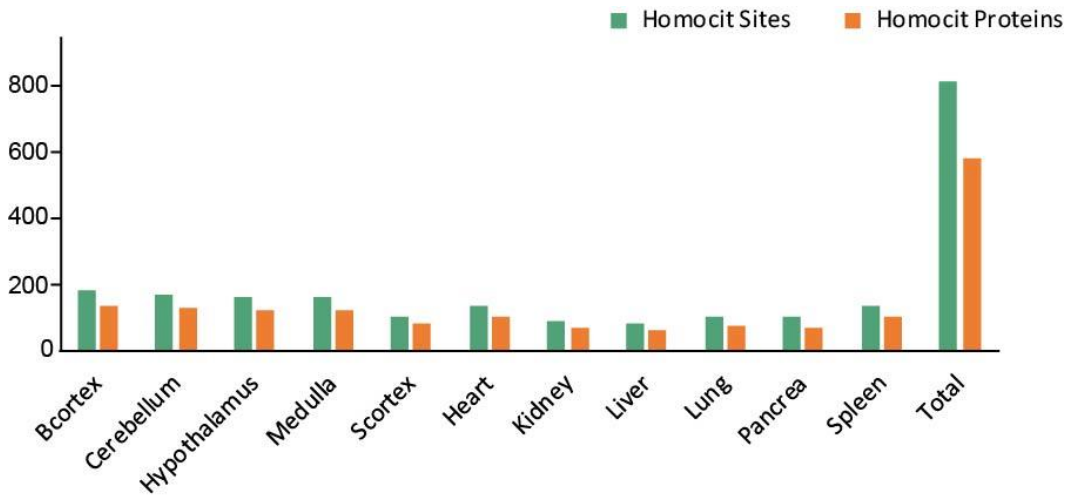


Figure 6. Number of identified, cleaved and mis-cleaved homocitrullination sites (A). Pie charts indicate the percentage of mis-cleaved homocitrullination site (B).



	Homocit sites	Homocit proteins		Homocit sites	Homocit proteins
B. cortex	185	140	Kidney	89	71
Cerebellum	176	133	Liver	86	67
Hypothalamus	168	128	Lung	102	75
Medulla	165	123	Pancreas	105	73
S. cortex	108	86	Spleen	140	107
Heart	142	105	Total	840	599

Figure 7. Number of homocitrullinated proteins and homocitrullination sites identified in mouse brain and organs.

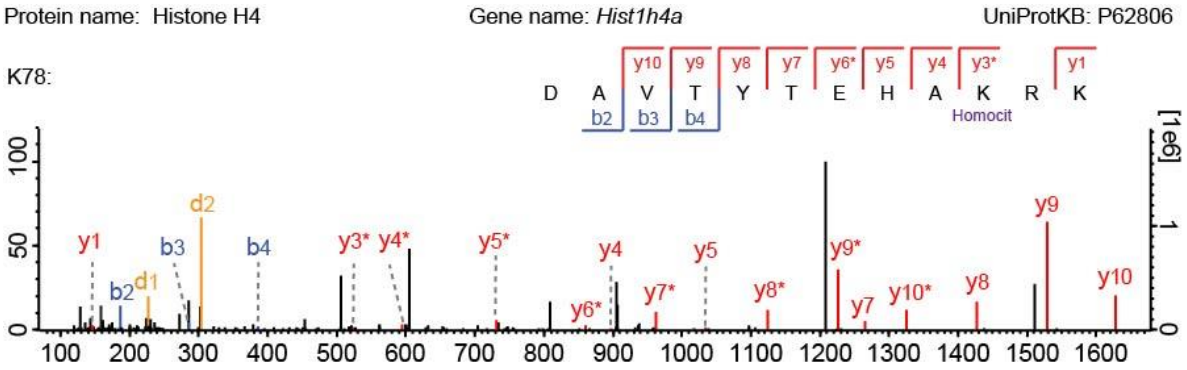


Figure 8. Tandem MS spectrum of a homocitrullinated peptide DAVTYTEHAKRK from Histone H4. Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. * denotes fragment ions resulting from the neutral loss of biotin tag.

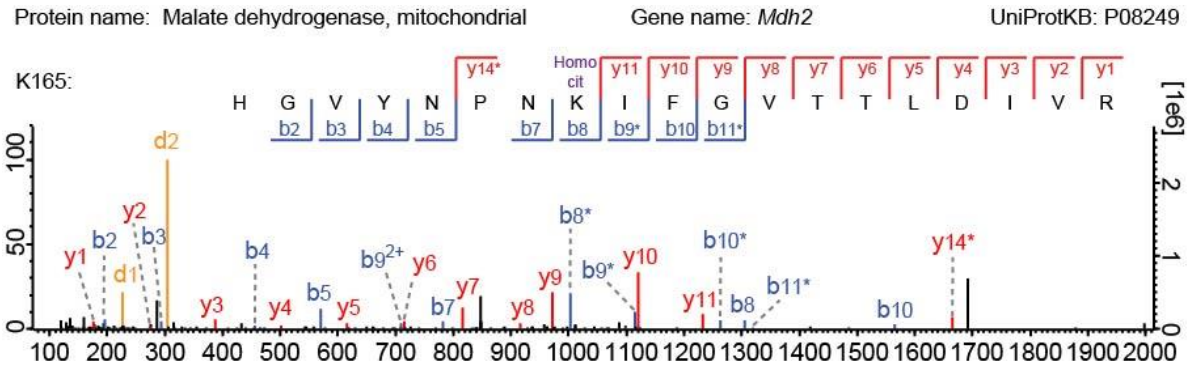
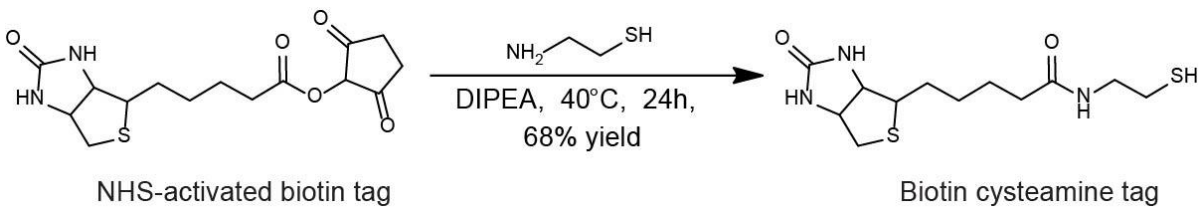


Figure 9. Tandem MS spectrum of a homocitrullinated peptide HGVYNPNKIFGVTTLDIVR from malate dehydrogenase. Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. * denotes fragment ions resulting from the neutral loss of biotin tag.

Supplemental Information



NHS-activated biotin tag

Biotin cysteamine tag

Scheme S1. Syntheses of the novel biotin thiol tag for homocitrullination analysis.

Table S1: Homocitrullination analysis of samples digested by different enzymes.

High confident homocitrullination sites and homocitrullinated peptides			
Enzyme	Trypsin gold	Trypsin/Lys-C	Lys-C
Homocit sites	100	82	75
Homocit peptides	73	65	62
Localization probability of identified homocitrullination sites			
Enzyme	Trypsin gold	Trypsin/Lys-C	Lys-C
Sites identified	125	95	81
Sites with prob (0.75-1)	100	82	75
Sites with prob (<0.75)	25	13	6

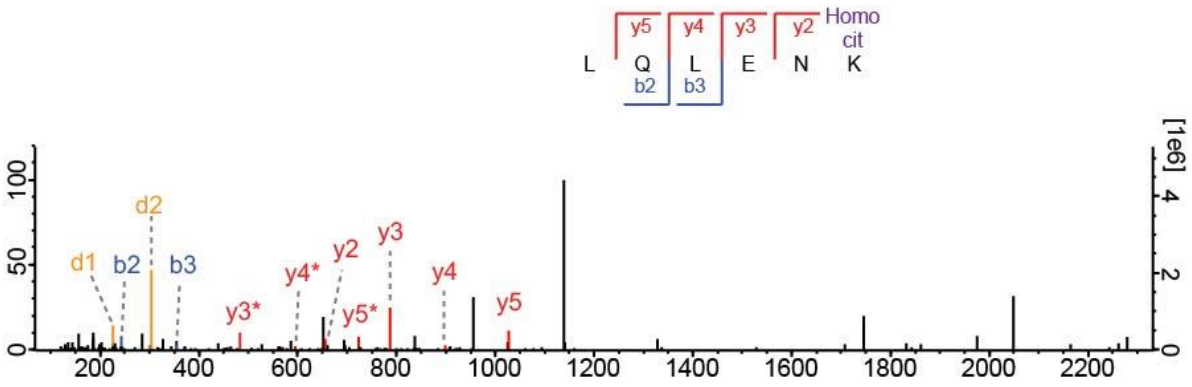


Figure S1. Tandem MS spectrum of the unique peptide LQLENK with a C-terminal homicitrulline identified from trypsin/Lys-C digested sample. Two diagnostic ions at m/z 227 (d1) and m/z 304 (d2) were observed. * denotes fragment ions resulting from the neutral loss of biotin tag.

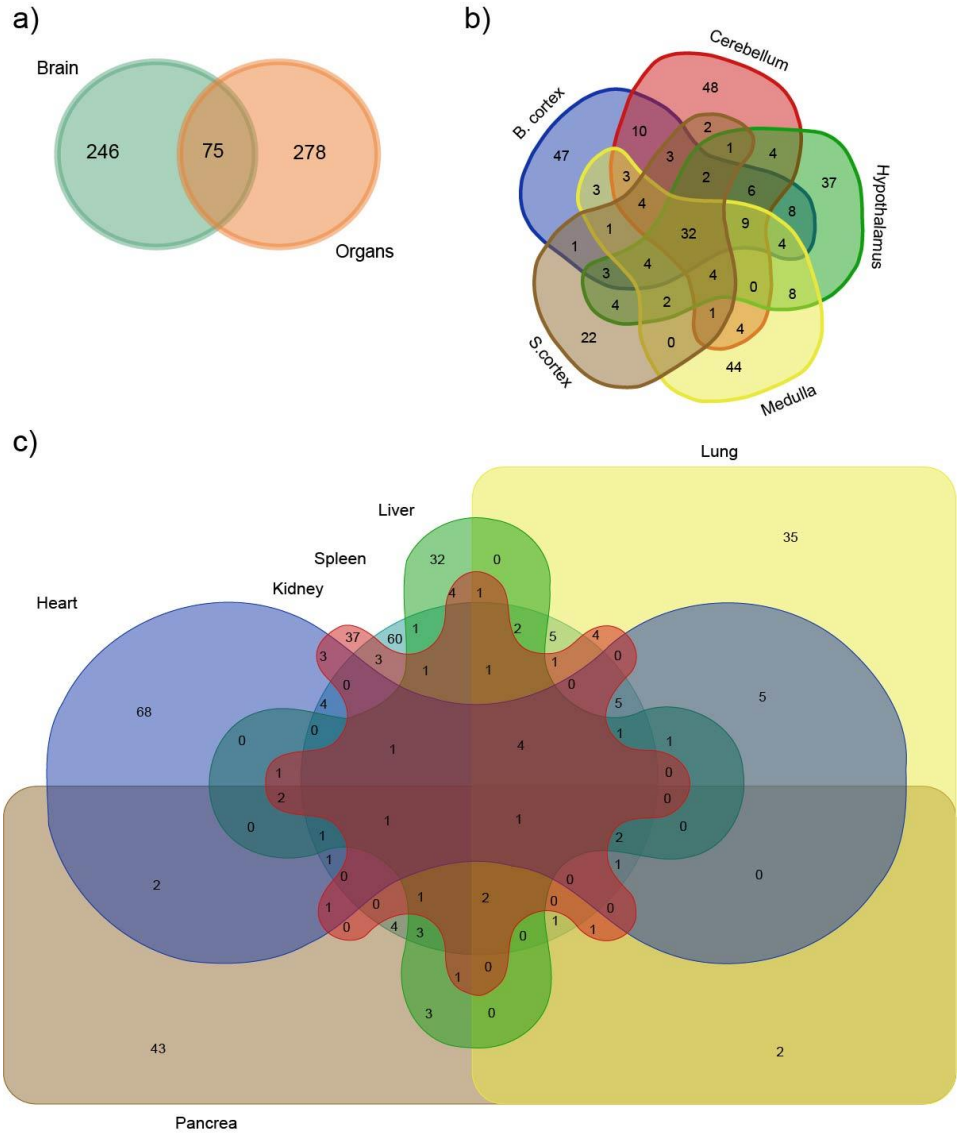


Figure S2. Mouse tissue-specific distribution patterns of homocitrullinated proteins. The overlap of homocitrullinated proteins identified from the brain and organs (a); The overlap of homocitrullinated proteins identified in five brain regions (b); The overlap of homocitrullinated proteins identified from six organs (c)

Histone	Site	Identified Peptide sequence
Histone H1.4	K119	AASGEAKPKAK
Histone H2B	K109	HAVSEGTKAVTK
	K117	LLLPGELAKHAVSEGTK
Histone H3.3	K57	RYQKSTELLIR
	K80	EIAQDFKTDLR
	K123	RVTIMPKDIQLAR
Histone H4	K78	DAVTYTEHAKRK
	K80	RKTVTAMDVVYALK
	K92	TVTAMDVVYALKR

Figure S3. Identified homocitrullination sites localized at different types of mouse histones.

Chapter 5

MS Investigation of Signaling Pathways Involved in Synergistic Effect of LPS and TPG on Enhanced IFN-β Production in Macrophages

This project is a collaboration between our lab and Prof. Judith Smith lab. Cell samples were prepared by Dr. Yiping Liu in Prof. Judith Smith lab. Protein digestion, peptide labeling and enrichment as well as MS-based analysis is performed by myself in our lab.

Abstract

Lipopolysaccharide (LPS) is a well-known endotoxin, which could stimulate the production of interferon β (IFN-β) in macrophage cells. Recently, the log-fold increase of interferon β production in macrophages subsequently treated by thapsigargin (TPG) and LPS has been reported. It has been demonstrated that the unfolded protein response in macrophages, which is in response to the TPG-induced endoplasmic reticulum stress, plays an important role in the synergistic effect of LPS and TPG on enhanced IFN-β production in macrophages. However, the mechanism underlying this phenomenon has not been clearly explored. In this study, we have performed quantitative proteomics and quantitative phosphoproteomics studies to reveal the global changes of proteins and phosphoproteins in LPS/TPG-treated macrophages. We found that proteins and phosphoproteins in LPS/TPG-treated macrophages are significantly changed in comparison to those in LPS-only and TPG-only treated groups as well as non-treated group. Signaling pathways, including splicesome, DNA replication, ribosome biogenesis and TNF signaling pathways, appeared to be associated with the synergistic effect of LPS and TPG on the enhanced IFN-β production in macrophages.

Introduction

Unfolded protein response (UPR) is an adaptive response involved in cell survival or death, which is triggered by many factors such as hypoxia,[1,2] nutrient deprivation,[3] hypoglycemia,[4,5] viral infection [6-7] or calcium gradient deregulation that disrupt endoplasmic reticulum (ER) function.[8-10] Currently, it is known that UPR has three pathways including inositol-requiring enzyme (IRE)-1,[11] protein kinase receptor–like ER kinase (PERK)[12] and activating transcription factor 6 (ATF-6).[13] IRE-1 is an endonuclease that is activated after the release of binding immunoglobulin protein (BiP), and can cleave a 26-bp intron from the X-box binding protein 1 (XBP-1) transcription factor mRNA.[14] This unusual splicing event removes a premature stop codon through frame shifting the open reading frame, thus allowing for the translation of the full-length XBP-1 transcription factor. Upon the release of BiP, PERK transiently inhibits global protein translation apart from selected transcripts (e.g., ATF-4). Meanwhile, ATF-6 leaves the ER and travels to the Golgi, where it is processed to an active form. Generally, UPR target genes aimed at resolving ER stress include folding chaperones and proteins that aid in ER associated protein degradation. If these and other adaptations fail, the UPR results in apoptosis. [10]

Lipopolysaccharide (LPS) is a component of the cell wall of Gram-negative bacteria, which can stimulate immune cells, such as macrophages via toll-like receptor 2 (TLR2) and TLR4.[15] This stimulation results in the generation of various proinflammatory cytokines, such as type I interferon, tumor necrosis factor α (TNFα), interleukin 1 (IL-1) and IL-6.[16] Our previous studies have shown that thapsigargin (TPG)-pretreated macrophages undergoing an intracellular UPR can

respond to LPS treatment with greatly enhanced IFN-β production.[17] Although we have demonstrated that XBP1 splicing and the phosphorylation of interferon regulatory factor (IRF3) contributed to this synergistic effect, the comprehensive mechanism is still unclear.[18]

In this study, quantitative proteomics and quantitative phosphoproteomics were performed using our 12-plex DiLeu isobaric labeling tags, to explore the global changes of cellular proteins and signaling pathways involved in the synergistic effects of LPS and TPG on the enhanced IFN-β production in macrophages. A large number of proteins and phosphoproteins were significantly changed in LPS/TPG-treated macrophages, providing new insights and potential mechanisms underlying the synergy effect of LPS and TPG treatments.

Materials and methods

Chemicals and reagents

Methanol (MeOH), ethanol (EtOH), acetonitrile (ACN), formic acid (FA), sodium chloride (NaCl), trifluoroacetic acid (TFA) and ammonium bicarbonate were purchased from Fisher Scientific (Pittsburgh, PA). Microcon-30kDa centrifugal filter unit was purchased from EMD Millipore Sigma (Burlington, MA). Dithiothreitol was purchased from Promega (Madison, WI). Distilled water mentioned in this work was Milli-Q water from a Millipore filtration system (Bedford, MA). Tris (2-carboxyethyl) phosphine (TCEP) and triethylammonium bicarbonate buffer (TEAB) were purchased from Sigma-Aldrich (St. Louis, MO). All reagents were used without additional purification.

Cell sample preparation

The immortalized macrophage cell (iMac) generated in our lab was maintained in DMEM/high glucose with 4 mM L-glutamine, sodium pyruvate (HyClone Laboratories, Logan, UT) and supplemented with 10% FBS (HyClone Laboratories), 100 U/ml penicillin, and 100 mg/ml streptomycin. Five million of the iMac cells in 10 cm dishes were pretreated in 4 different conditions: Group 1: 0.05 mM thapsigargin (TPG) for 24 h, Group 2: Salmonella enteritidis LPS (Sigma-Aldrich) 100 ng/ml for 24h. Group 3: 0.05 mM TPG for 1 hr, then adding 100ng/ml LPS for 24 h. Group 4: no treatment as wild type (WT). Each group has three biological samples. After the treatment, all the cells were scrapped in to 15ml tube and washed with PBS buffer for 3 times. Finally, 12 cell samples were stored at -80°C for further analysis.

Protein extraction and digestion

30µl cell pellets was mixed with 90µl extraction buffer containing 4 % SDS/50mM Tris base and incubated at 95°C for 10 min. The cell lysate was then sonicated for 15 min. Protein concentration was determined by BCA assay (Thermo Fisher Scientific, Waltham, MA). Protein extract was treated with 10 mM dithiothreitol (DTT) for 30 min at room temperature and further alkylated with 50mM iodoacetic acid (IAA) for 30 min in the dark. SDS was removed by the traditional acetone precipitation method. Cold acetone (-20°C) was added to the cell lysate to the final concentration of 80% (v/v). The sample was stored at -20°C overnight and then centrifugated at 16,000 × g for 15 min. The supernatant was removed, and the pellet was incubated with 600 µl 80% acetone/water (v/v, -20°C) at -20°C for the additional 2 h. The supernatant was completely removed after centrifugation. The protein pellet was air-dried in room temperature and

reconstituted in 8 M urea. Lys-C/trypsin (1:100, Promega) was added for the first-round digestion at 37°C for 4 h. Then, the sample solution was diluted with 50mM Tris buffer and trypsin (1:100, Promega) was added for the second-round digestion at 37°C for 14 h. Trypsin digestion was quenched with 1% TFA and tryptic peptides were desalted with Sep-Pak C_{18} cartridge (Waters). Peptide samples were finally lyophilized and stored at -80°C until use.

12-plex DiLeu isobaric tag labeling

Peptide assay (Thermo Fisher Scientific, Waltham, MA) was performed to determine the peptide concentration of each sample. Twelve samples were labeled by 12-plex DiLeu isobaric tags as the following: WT group (115a, 115b and 117c), TPG-treated group (116c, 117b and 118c), LPS-treated group (116a, 116b and 118b) and LPS/TPG-treated group (117a, 118a and 118d).

Three hundred microgram peptides from each sample were labeled. DiLeu isobaric tags were suspended in anhydrous DMF and combined with DMTMM and NMM at 0.6× molar ratios to DiLeu. The activation of DiLeu isobaric tags was performed at room temperature for 60 min. Peptides were labeled by adding the activated DiLeu isobaric tags to peptide digest at a 10:1 ratio (w/w). The DiLeu labeling reaction was carried out at room temperature for 2 h. The reaction was quenched by adding 5% NH_2OH to the final concentration of 0.25%. After quench, peptides from different sample were combined and dried under SpeedVac.

Sample cleaning and fractionation

For proteomic analysis, 12-plex DiLeu labeled peptide mixture was cleaned with strong cation exchange (SCX) chromatography. TopTipTM packed with PolySULFOETHYL A beads (Poly LC) was prewashed with 100µl ACN and then equilibrated with 100µl loading buffer (0.2%

FA/25% ACN/water (v/v/v)) for three times. 300 µg peptide mixture was reconstituted in 100µl loading buffer and loaded onto beads. The leftover DiLeu tag was completely removed by washing the beads with 100µl loading buffer for four times. Peptides were finally eluted with 100µl 0.4M ammonium formate/25% ACN/water (v/v) for 3 times. The eluent was combined and dried for high pH reversed-phase peptide fractionation. The peptide sample was separated on the C_{18} column (beads) and fractioned into 5 fractions. Each fraction was lyophilized and stored in -80°C freezer until use.

Enrichment of phosphopeptides using Ti-IMAC technique

Titanium (IV) immobilized metal affinity chromatography (Ti(IV)-IMAC) adsorbents were prepared according to our previous protocol.[19] Enrichment of 12-plex DiLeu labeled peptide mixture was performed as the following procedure. First, the sample was reconstituted in 1 ml 40% ACN/3% trifluoroacetic acid (TFA) (v/v) loading buffer and then incubated with 20 mg Ti(IV)-IMAC adsorbents for 30 min. Then, the sample was centrifuged at 20,000 x g for 5 min and the supernatants were discarded. The adsorbents were washed sequentially by adding 1 ml 50% ACN /6% TFA (v/v)/200 mM NaCl and 1 ml 30% ACN /0.1% TFA (v/v) to remove the non-specific adsorbed peptides. Finally, the adsorbents were incubated with 200µl 10% NH_4OH (v/v) for 15 min to elute the captured phosphopeptides and then centrifuged at 20,000 x g for 5 min. The supernatant was transferred in a new centrifuge tube and lyophilized to dry. Samples were reserved in -80°C freezer and re-dissolved in 0.1% formic acid (FA) for LC-MS/MS analysis.

Mass spectrometry analysis

The 12-plex DiLeu labeled peptides were analyzed on the Q Exactive™ HF Hybrid Quadrupole-Orbitrap™ Mass Spectrometer coupled to a Dionex UPLC system (Thermo Fisher Scientific, San Jose, CA). The mobile phase was composed of 0.1% formic acid in water (A) and 0.1% formic acid in ACN (B). The flow rate was set as 300nl/min. Sample was loaded onto a customized C_{18} column filled with 1.7μm Ethylene Bridged Hybrid packing materials (130 Å, Waters) and separated with the following gradient: 3% B for the first 16 min upon sample trapping; 3%-30% B for 18.3-120 min; 30%-75% B for 120-120.5 min; 75% B for 120.5-130 min; 75%-95% B for 130.0-130.5 min; 95% B for 130.5-140.0 min; 95%-3% B for 140.0-140.5 min and equilibrated at 3% B for 15 min. MS data was acquired in the positive ion mode with the spray voltage of 2 kV. MS^1 spectra were collected from m/z 300-1500 by orbitrap at a resolution of 60,000. Automatic gain control (AGC) target and maximum injection time was set to $1e^6$ and 100 ms, respectively. Ions with the charge of 2-6 were included for screening. The fifteen precursors were selected in an order of intensity and isolated in an isolation window of 1.4 *m/z* for HCD fragmentation using a normalized collisional energy of 27%. Tandem MS spectra were collected at a resolution of 60,000. Automatic gain control (AGC) target and maximum injection time was set to $2e^5$ and 200 ms, respectively. The fixed first mass of MS^2 spectra was set to *m/z* 105.

The 12-plex DiLeu labeled phosphopeptides were analyzed on the Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer coupled to a Dionex UPLC system (Thermo Fisher Scientific, San Jose, CA). The mobile phase was composed of 0.1% formic acid in water (A) and 0.1% formic acid in ACN (B). The flow rate was set as 300nl/min. Phosphopeptides were seperated on a customized C_{18} column filled with 1.7μm Ethylene Bridged Hybrid packing materials (130

Å, Waters) using the following gradient: 3% B for the first 18.3 min upon sample trapping; 3%-30% B for 18.3-120 min; 30%-75% B for 120-120.5 min; 75% B for 120.5-130 min; 75%-95% B for 130.0-130.5 min; 95% B for 130.5-140.0 min; 95%-3% B for 140.0-140.5 min and equilibrated at 3% B for 15 min. MS data was acquired in the positive ion mode with the spray voltage of 2 kV. MS^1 spectra were collected from m/z 300-1500 by orbitrap at a resolution of 120,000 at m/z 200. RF lens, automatic gain control (AGC) target and maximum injection time was set to 30%, $5e^5$ and 50 ms, respectively. Ions with the charge of 2-8 were included for screening. Dynamic exclusion was set at 20 s with a 10 ppm tolerance. The top 20 precursors with highest intensities were isolated with an isolation window of 1 m/z for stepped HCD fragmentation using collisional energy of 22%, 30% and 38%. The MS^2 spectra was acquired in the Orbitrap with a resolution of 60,000 and a start mass of *m/z* 110. AGC target and the maximum injection time was set to $1e^5$ and 150 ms, respectively.

Data Analysis

Data analysis of proteomics was performed using Proteome Discoverer 2.1 (Thermo Scientific). Raw data was searched against the mouse UniProt FASTA database (Sep 26, 2019) using the Sequest HT algorithm. b/y ions were enabled for searching data. Trypsin was defined as the enzyme used and a maximum of two missed-cleavages was allowed. A precursor tolerance of 50 ppm and a fragment ion tolerance of 0.02 Da were allowed. Carbamidomethylation and DiLeu labels on peptide N-terminus and lysine residues were set as static modifications, while dynamic modification consisted of oxidation of methionine residue for all search. Identifications were validated with 1% PSM and 1% protein FDR using percolator. Quantitative analysis was

conducted in Proteome Discoverer with a reporter ion integration tolerance of 20 ppm. Reporter ion intensities and their ratio values were exported to Excel for further statistical analysis with student t-test.

Maxquant was applied to quantitative phosphoproteomics analysis. The raw data was searched against the mouse UniProt FASTA database (Sep 26, 2019) using the Andromeda search engine. Trypsin was used as the enzyme with maximum of two missed-cleavages. Carbamidomethylation and DiLeu labels on peptide N-terminus and lysine residues were set as static modifications, while dynamic modification consisted of oxidation of methionine residues and the phosphorylation on serine, threonine and tyrosine residues. Proteins and peptides were identified with a target-decoy approach in revert mode and false-discovery rate (FDR) at the protein or peptide level was set to 0.01. Peptides that were found as reverse or potential contaminant hits were filtered out. Phosphorylation site localization probability threshold was set to 0.75. Gene ontology (GO) enrichment analysis of identified citrullinated proteins was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8. KEGG pathway analysis was performed using String (11.0).

Results and Discussion

The synergistic effect of LPS and TPG on the enhanced IFN-β production in macrophages was firstly discovered and investigated by our collaborators, Smith and cowrokers.[14, 17] As shown in **Figure 1**, LPS stimulation results in the production of IFN-β in macrophages. However, log-fold increased production of IFN-β could be observed in macrophages subsequently treated by

TPG and LPS. It has been demonstrated that the unfolded protein response (UPR) in macrophages, which is in response to the TPG-induced ER stress, contributed to the elevated expression of IFN-β, while other potential mechanisms have not been explored. Here, we aim to systematically profile the changes of proteins and phosphoproteins in macrophages undergoing different treatments for the discovery of important signaling pathways involved in the accumulation of IFN-β production in LPS/TPG-treated macrophages.

Quantitative proteomics of macrophages treated in four different conditions

The quantitative proteomics was performed using the workflow depicted in **Figure 2**. In total, 5031 proteins were identified and quantitated from four groups of mouse macrophage cells (WT, LPS-treated, TPG-treated and LPS/TPG-treated).

We firstly studied the protein changes in three treated groups in comparison to the WT group. In general, 61 proteins in LPS-treated group, 61 proteins in TPG-treated group and 187 proteins in LPS/TPG-treated group showed more than 2-fold change (FC) with high confidence ($P<0.05$) (**Figure 3**). As shown in **Figure 3d**, it is no surprise that the significantly changed proteins identified in TPG-treated group and LPS-treated group showed poor overlap, which could be easily explained by the different stimulation mechanism of TPG and LPS. The changed proteins in LPS/TPG-treated group has reasonable overlap with both TPG-treated group and LPS-treated group. Furthermore, 107 significantly changed proteins were only observed in the LPS/TPG-treated group. An example identified peptide, which was significantly up-regulated in LPS/TPG-treated group, is shown in **Figure 4**. A large number of fragment ions enabled high-confident identification of this unique peptide, while the intensities of DiLeu reporter ions in the low mass

region represented the relative intensities of this peptide in different samples, allowing relative quantitation. Gene Ontology (GO) enrichment analysis was performed to study the biological processes undergoing significant changes in three treatment groups (**Figure S1**).

We then investigated the changed proteins in LPS/TPG-treated group *versus* LPS-treated group, which is more important to explain synergistic effect of LPS and TPG. 39 proteins that were identified with high confidence exhibited two-fold changes in LPS/TPG-treated group compared to those in LPS-only treated group (**Figure 5**). Functional analysis indicated that these significantly changed proteins in LPS/TPG-treated samples were mainly involved in the response to stress, immune response and immune system process (**Figure S2**).

Quantitative phosphoproteomics of macrophages treated in four different conditions

Protein phosphorylation plays an important role in regulating membrane transport,[20, 21] enzyme activity [22, 23] and signaling networks. Therefore, quantitative phosphoproteomics was performed to reveal the signaling pathways associated with the synergistic effect of LPS and TPG on the enhanced IFN-β production in macrophages.

By using the Ti-IMAC technique, 12-plex DiLeu labeled phosphopeptides were enriched for MS-based quantitative analysis (**Figure 6**). We totally identified 5216 phosphorylation sites with high confidence and quantified 1874 phosphoproteins from four groups of mouse macrophage. Four different groups could be well clustered, indicating the good reproducibility between replicates (**Figure 7**).

The change of phosphoproteins in three treated groups are summarized in **Figure 8**. Compared to WT group, 42 phosphoproteins, 544 phosphoproteins and 774 phosphoproteins were

significantly down-regulated in LPS-treated group, TPG-treated group and LPS/TPG-treated group, respectively. Meanwhile, the significant up-regulation of 50 phosphoproteins, 59 phosphoproteins and 114 phosphoproteins in LPS-treated group, TPG-treated group and LPS/TPG-treated group, respectively, were also explored. The overlap of significantly changed phosphorylation sites and phosphoproteins between three treated groups was shown in **Figure S1**. Similar with the quantitative proteomics, the significantly changed phosphorylation sites and phosphoproteins in TPG-treated group and LPS-treated group were poorly overlapped. Furthermore, the high degree overlap between the TPG-treated group and LPS/TPG-treated group, especially the down-regulated phosphoproteins, suggested that the global phosphoproteomics changes in the LPS/TPG-treated group was more similar with that of TPG-treated group than LPS-treated group.

We further compared the LPS/TPG-treated group with the LPS-treated group. 462 phosphoproteins were significantly down-regulated, while 37 significantly up-regulated phosphoproteins were detected (**Figure 9**). These significantly changed proteins were subjected to GO analysis and KEGG pathway analysis. The significantly changed phosphoproteins are mainly related to processes regulating transcription and translation (**Figure 10**). Several KEGG pathways, such as splicesome, DNA replication and ribosome biogenesis, were down-regulated, while TNF signaling pathways and necroptosis were up-regulated (**Figure 11**).

Conclusions

In this study, we have conducted quantitative proteomics and phosphoproteomics using our 12-plex DiLeu isobaric tag labeling technique, to unravel the global changes of proteins and phosphoproteins in macrophages undergoing four different treatment conditions. We found that the global changes of proteins and phosphoprotein in LPS/TPG-treated macrophages are more pronounced than those in LPS-treated group and TPG-treated group. Phosphoproteomic changes in LPS/TPG-treated macrophages was very similar with that in the TPG-treated group, but not in the LPS-treated group. Multiple signaling pathways were involved in the synergistic effect of LPS and TPG on the enhanced IFN-β production in macrophages.

Acknowledgments

This research is supported in part by the National Institutes of Health through grants-R21 AG060242, U01CA231081, RF1AG052324, and R01 DK071801. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531 to LL) and the University of Wisconsin-Madison, Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation. LL acknowledges a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

References

(1) Wouters, B. G.; Koritzinsky, M., Hypoxia signalling through mTOR and the unfolded protein response in cancer. *Nat Rev Cancer* **2008,** 8, (11), 851-64.

(2) Feldman, D. E.; Chauhan, V.; Koong, A. C., The unfolded protein response: a novel component of the hypoxic stress response in tumors. *Mol Cancer Res* **2005,** 3, (11), 597-605.

(3) Chang, S. H.; Barbosa-Tessmann, I.; Chen, C.; Kilberg, M. S.; Agarwal, A., Glucose deprivation induces heme oxygenase-1 gene expression by a pathway independent of the unfolded protein response. *J Biol Chem* **2002**, 277, (3), 1933-40.

(4) Yu, D. H.; Macdonald, J.; Liu, G.; Lee, A. S.; Ly, M.; Davis, T.; Ke, N.; Zhou, D.; Wong-Staal, F.; Li, Q. X., Pyrvinium targets the unfolded protein response to hypoglycemia and its anti-tumor activity is enhanced by combination therapy. *PLoS One* **2008**, 3, (12), e3951.

(5) Kim, J. L.; La Gamma, E. F.; Estabrook, T.; Kudrick, N.; Nankova, B. B., Whole genome expression profiling associates activation of unfolded protein response with impaired production and release of epinephrine after recurrent hypoglycemia. *PLoS One* **2017**, 12, (2), e0172789.

(6) Medigeshi, G. R.; Lancaster, A. M.; Hirsch, A. J.; Briese, T.; Lipkin, W. I.; Defilippis, V.; Fruh, K.; Mason, P. W.; Nikolich-Zugich, J.; Nelson, J. A., West Nile virus infection activates the unfolded protein response, leading to CHOP induction and apoptosis. *J Virol* **2007**, 81, (20), 10849-60.

(7) Chan, S. W.; Egan, P. A., Hepatitis C virus envelope proteins regulate CHOP via induction of the unfolded protein response. *FASEB J* **2005**, 19, (11), 1510-2.

(8) Hoyer-Hansen, M.; Jaattela, M., Connecting endoplasmic reticulum stress to autophagy by unfolded protein response and calcium. *Cell Death Differ* **2007**, 14, (9), 1576-82.

(9) Hammadi, M.; Oulidi, A.; Gackiere, F.; Katsogiannou, M.; Slomianny, C.; Roudbaraki, M.; Dewailly, E.; Delcourt, P.; Lepage, G.; Lotteau, S.; Ducreux, S.; Prevarskaya, N.; Van Coppenolle, F., Modulation of ER stress and apoptosis by endoplasmic reticulum calcium leak via translocon during unfolded protein response: involvement of GRP78. *FASEB J* **2013**, 27, (4), 1600-9.

(10) Schroder, M.; Kaufman, R. J., ER stress and the unfolded protein response. *Mutat Res* **2005**, 569, (1-2), 29-63.

(11) Deng, Y.; Srivastava, R.; Quilichini, T. D.; Dong, H.; Bao, Y.; Horner, H. T.; Howell, S. H., IRE1, a component of the unfolded protein response signaling pathway, protects pollen development in Arabidopsis from heat stress. *Plant J* **2016**, 88, (2), 193-204.

(12) Harding, H. P.; Zhang, Y.; Bertolotti, A.; Zeng, H.; Ron, D., Perk is essential for translational regulation and cell survival during the unfolded protein response. *Mol Cell* **2000**, 5, (5), 897-904.

(13) Yoshida, H.; Okada, T.; Haze, K.; Yanagi, H.; Yura, T.; Negishi, M.; Mori, K., ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Mol Cell Biol* **2000**, 20, (18), 6755-67.

(14) Zeng, L.; Liu, Y. P.; Sha, H.; Chen, H.; Qi, L.; Smith, J. A., XBP-1 couples endoplasmic reticulum stress to augmented IFN-beta induction via a cis-acting enhancer in macrophages. *J Immunol* **2010**, 185, (4), 2324-30.

(15) Beg, A. A., Endogenous ligands of Toll-like receptors: implications for regulating inflammatory and immune responses. *Trends Immunol* **2002**, 23, (11), 509-12.

(16) Gessani, S.; Testa, U.; Varano, B.; Di Marzio, P.; Borghi, P.; Conti, L.; Barberi, T.; Tritarelli, E.; Martucci, R.; Seripa, D.; et al., Enhanced production of LPS-induced cytokines during differentiation of human monocytes to macrophages. Role of LPS receptors. *J Immunol* **1993**, 151, (7), 3758-66.

(17) Smith, J. A.; Turner, M. J.; DeLay, M. L.; Klenk, E. I.; Sowders, D. P.; Colbert, R. A., Endoplasmic reticulum stress and the unfolded protein response are linked to synergistic IFN-beta induction via X-box binding protein 1. *Eur J Immunol* **2008**, 38, (5), 1194-203.

(18) Liu, Y. P.; Zeng, L.; Tian, A.; Bomkamp, A.; Rivera, D.; Gutman, D.; Barber, G. N.; Olson, J. K.; Smith, J. A., Endoplasmic reticulum stress regulates the innate immunity critical transcription factor IRF3. *J Immunol* **2012**, 189, (9), 4630-9.

(19) Zhou, H.; Ye, M.; Dong, J.; Corradini, E.; Cristobal, A.; Heck, A. J.; Zou, H.; Mohammed, S., Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nat Protoc* **2013**, 8, (3), 461-80.

(20) MacDonald, J. A.; Storey, K. B., Regulation of ground squirrel Na+K+-ATPase activity by reversible phosphorylation during hibernation. *Biochem Biophys Res Commun* **1999**, 254, (2), 424-9.

(21) Lowndes, J. M.; Hokin-Neaverson, M.; Bertics, P. J., Kinetics of phosphorylation of Na+/K(+)-ATPase by protein kinase C. *Biochim Biophys Acta* **1990**, 1052, (1), 143-51.

(22) Tetlow, I. J.; Wait, R.; Lu, Z.; Akkasaeng, R.; Bowsher, C. G.; Esposito, S.; Kosar-Hashemi, B.; Morell, M. K.; Emes, M. J., Protein phosphorylation in amyloplasts regulates starch branching enzyme activity and protein-protein interactions. *Plant Cell* **2004**, 16, (3), 694-708.

(23) Bobrovskaya, L.; Dunkley, P. R.; Dickson, P. W., Phosphorylation of Ser19 increases both Ser40 phosphorylation and enzyme activity of tyrosine hydroxylase in intact cells. *J Neurochem* **2004**, 90, (4), 857-64.

(24) Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M., Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **2006**, 127, (3), 635-48.

(25) de la Fuente van Bentem, S.; Mentzen, W. I.; de la Fuente, A.; Hirt, H., Towards functional phosphoproteomics by mapping differential phosphorylation events in signaling networks. *Proteomics* **2008**, 8, (21), 4453-65.

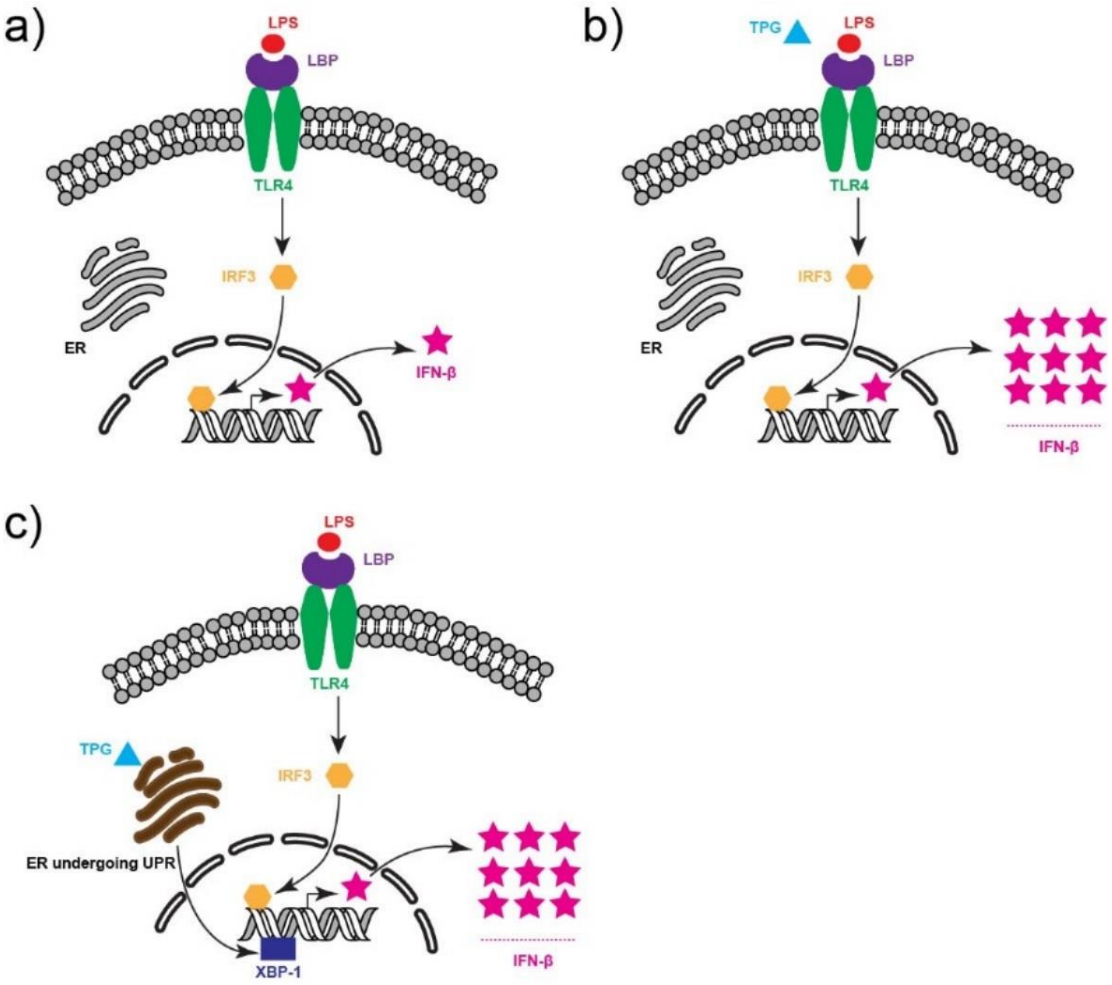


Figure 1. The synergistic effect of LPS and TPG on the enhanced IFN-β production in macrophages. a) LPS-induced production of IFN-β in macrophages; b) Log-fold increased production of IFN-β in TPG-pretreated macrophages induced by LPS; c) XBP-1 couples ER stress to augmented IFN-β induction via a cis-acting enhancer in macrophages. LBP: Lipopolysaccharide binding protein; TLR4: Toll Like Receptor 4; IRF3: Interferon Regulatory Factor 3; UPR: Unfolded protein response; XBP-1: X-box binding protein 1.

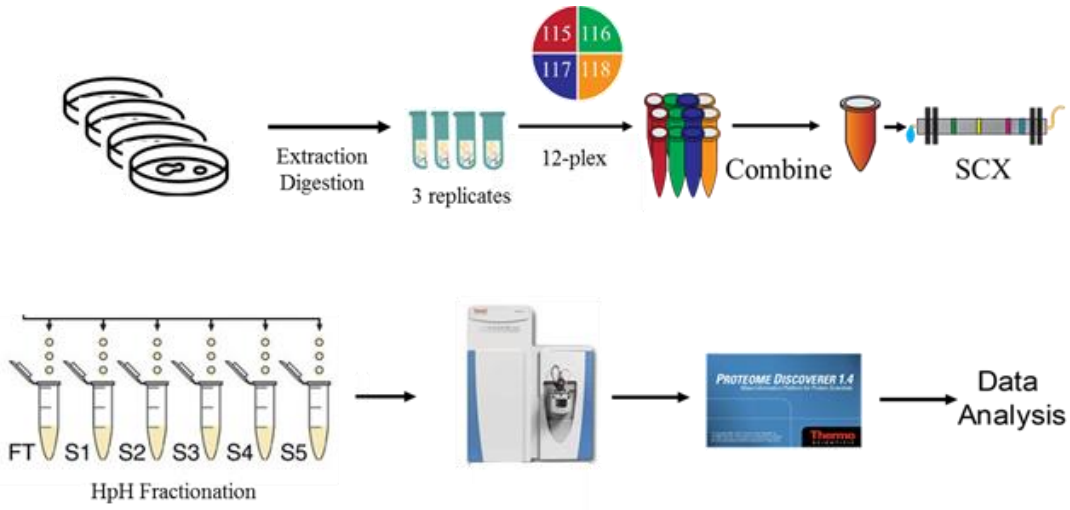


Figure 2. Workflow for quantitative proteomic analysis.

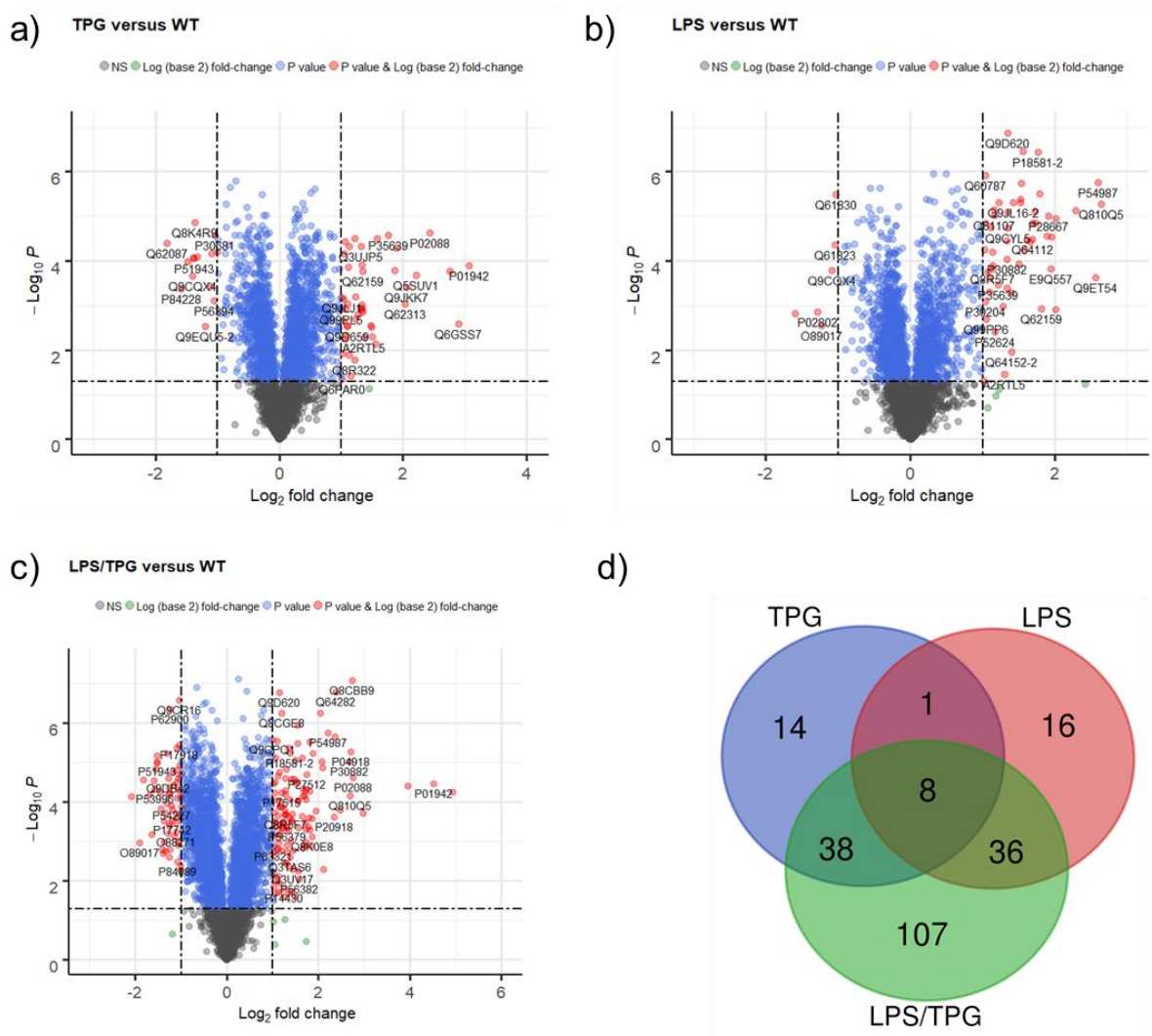


Figure 3. Volcano Plots showing the significantly regulated proteins in TPG-treated samples (a), LPS-treated samples (b) and LPS/TPG-treated samples (c) as well as the Venn diagram showing the overlap of significantly changed proteins in three differentially treated samples (d).

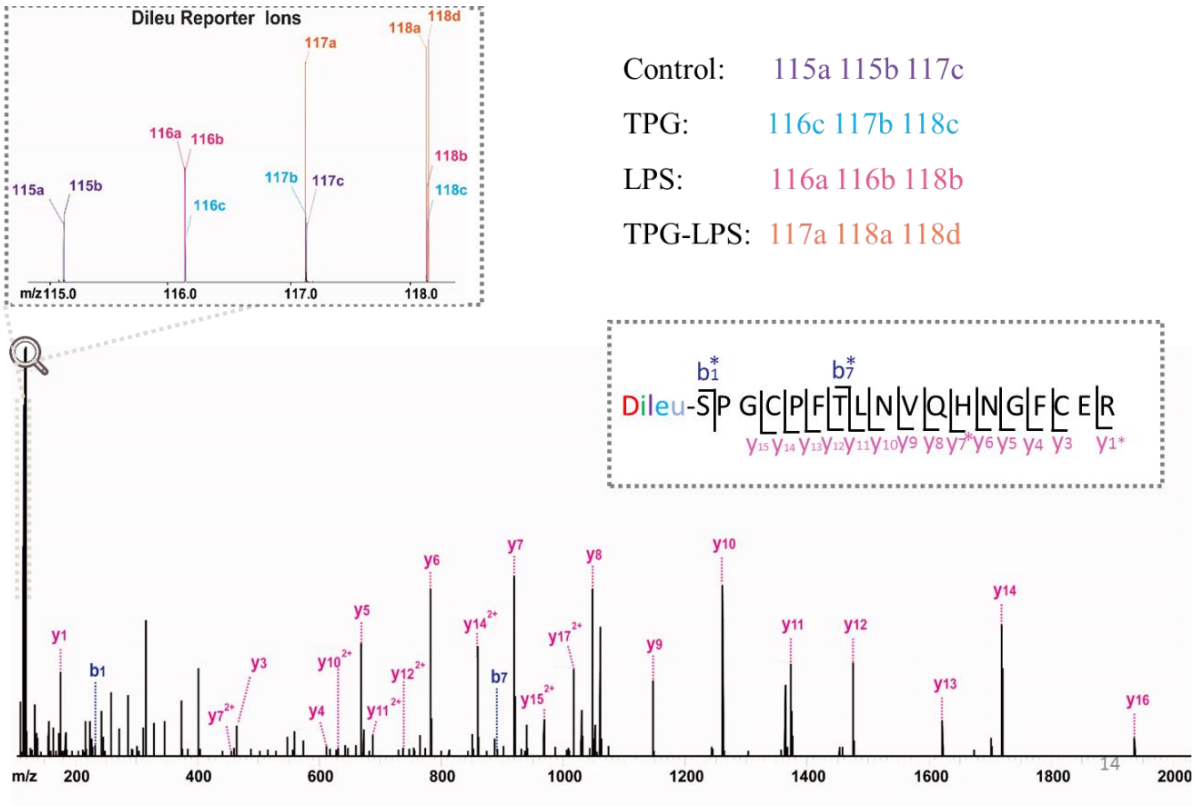


Figure 4. A representative tandem MS spectrum of 12-plex DiLeu labeled peptide significantly up-regulated in LPS/TPG-treated samples. The b- and y- product ions represent the backbone fragmentation of the peptide for identification. The intensities of reporter ions are used for relative quantification.

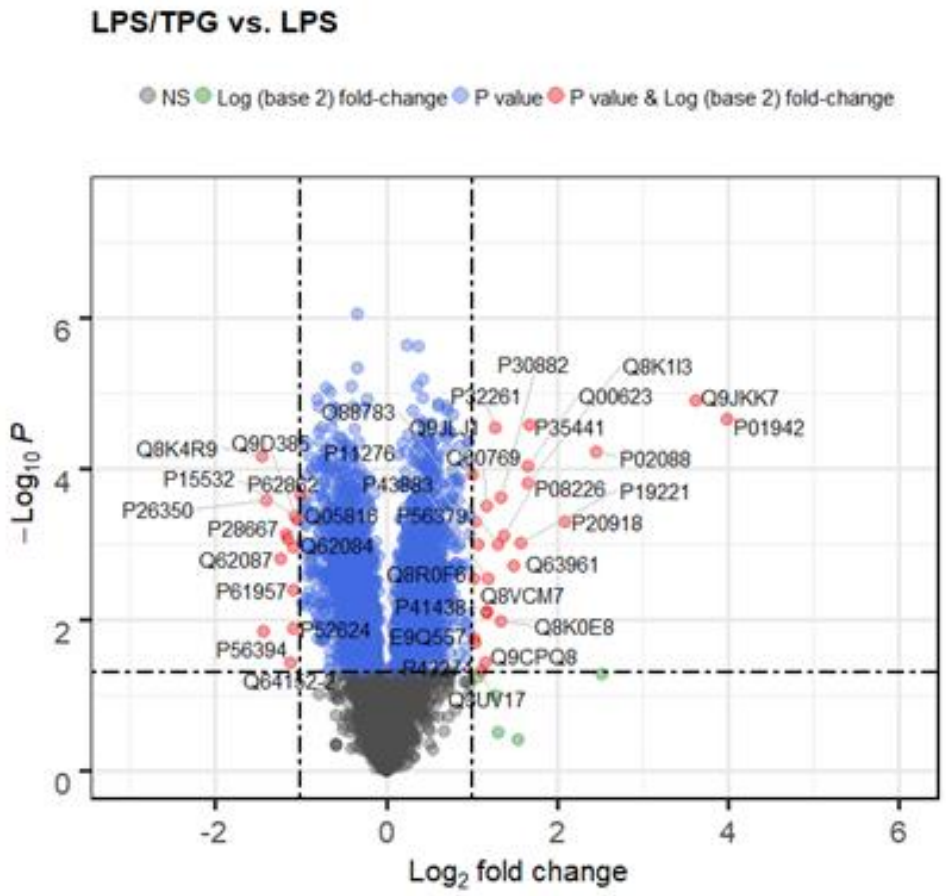


Figure 5. Volcano plots showing the significantly changed proteins in LPS/TPG-treated samples versus LPS-treated samples.

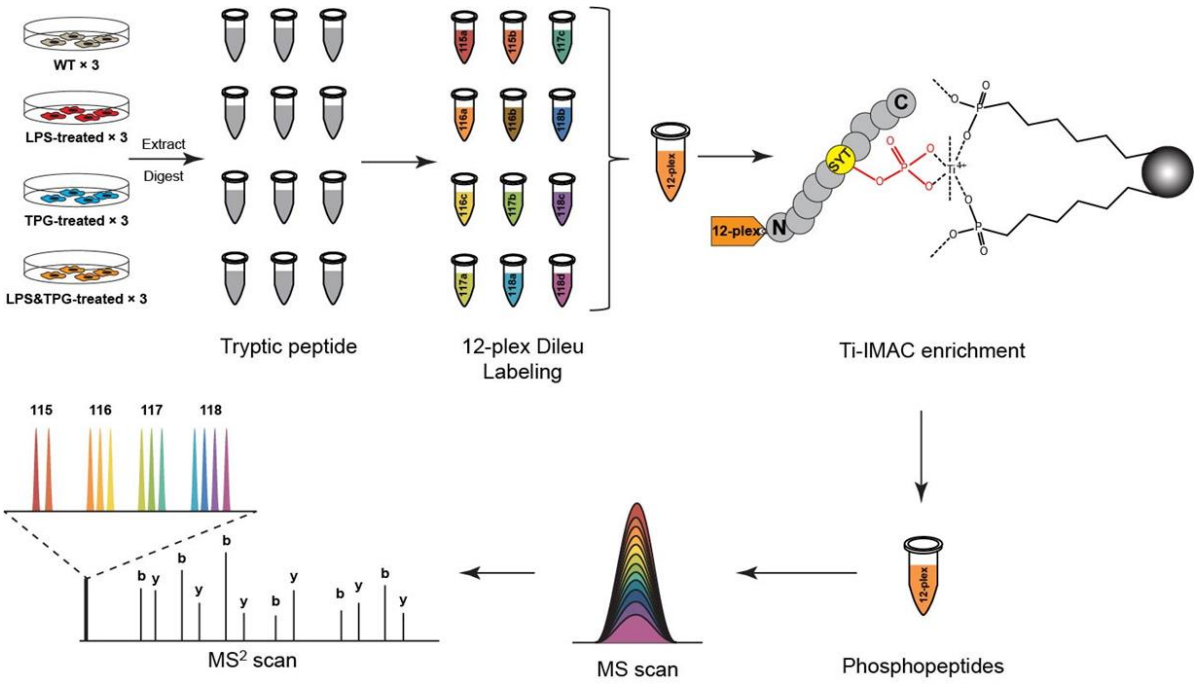


Figure 6. Workflow for quantitative phosphoproteomic analysis.

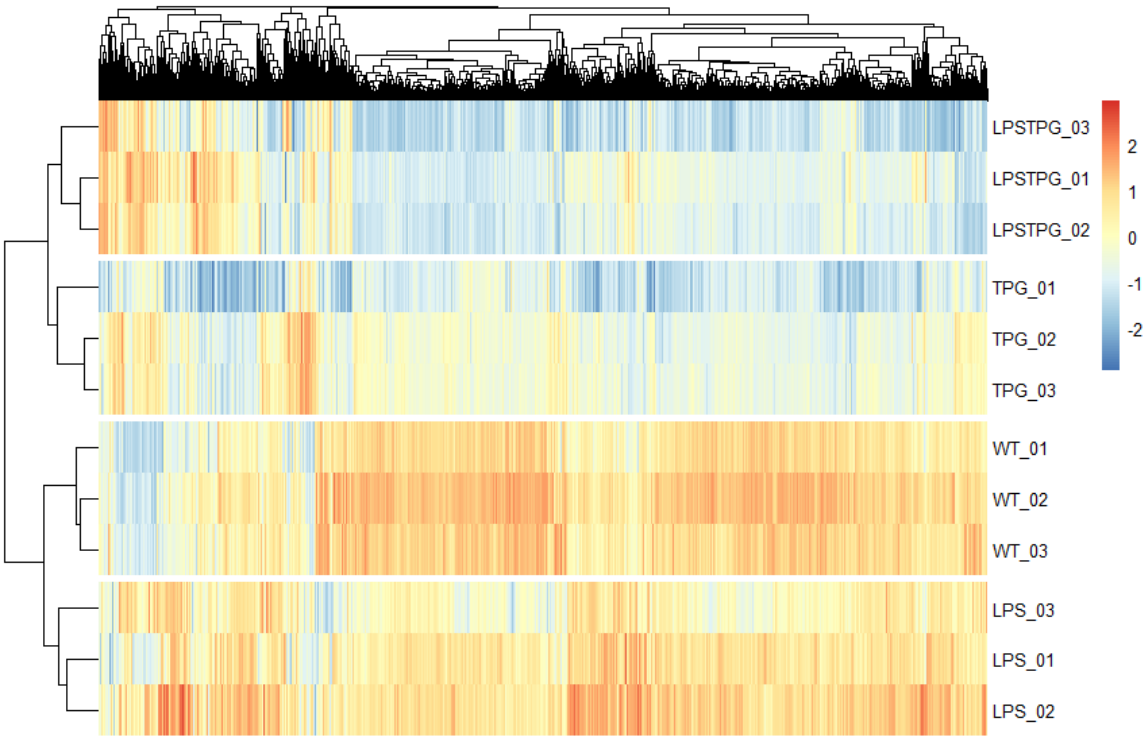
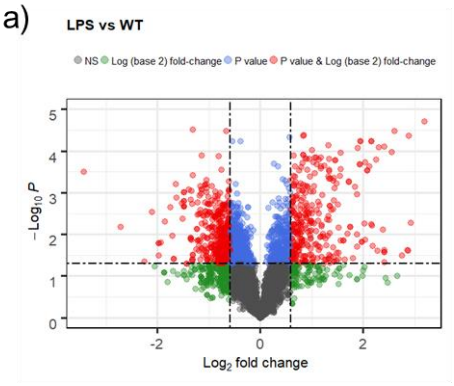
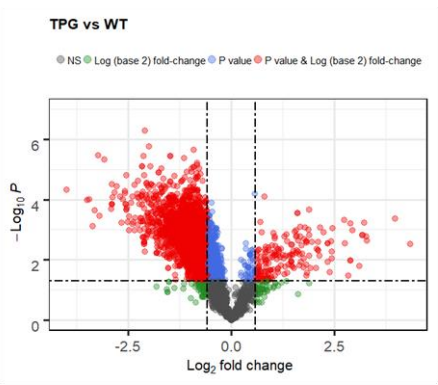


Figure 7. Heat map showing the good reproducibility between replicates of four differentially treated groups.

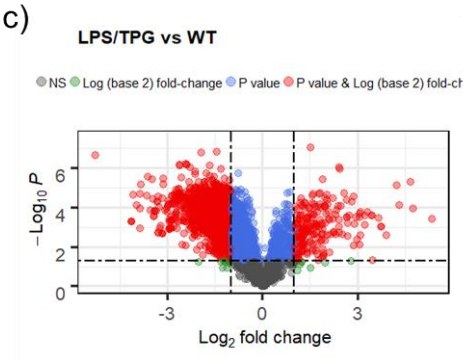
a)



b)



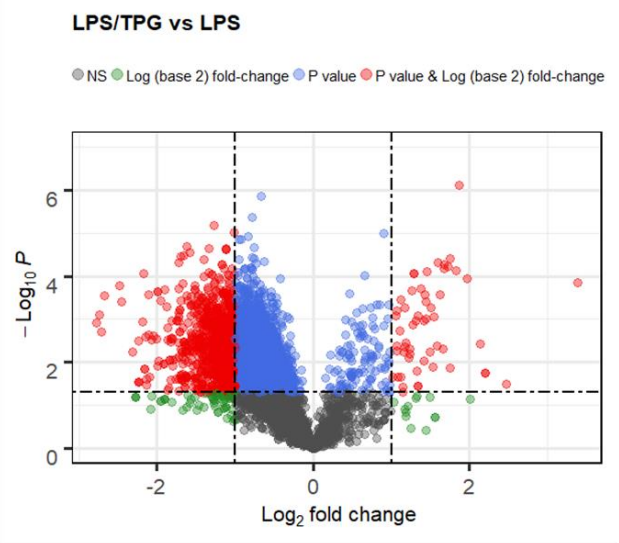
c)



d)

	Down-regulated P-sites/P-proteins (FC>2, P<0.05)	Up-regulated P-sites/P-proteins (FC>2, P<0.05)
LPS/WT	64/42	88/50
TPG/WT	1046/544	108/59
LPSTPG/WT	1739/774	205/114

Figure 8. Volcano plots showing the significantly regulated phosphoproteins in LPS-treated samples (a), TPG-treated samples (b) and LPS/TPG-treated samples (c) as well as the number of significantly up-regulated and down-regulated phosphorylation sites and phosphoproteins in three differentially treated samples (d).



	Down-regulated phosphorylation sites (FC>2, P<0.05)	Down-regulated phosphoproteins (FC>2, P<0.05)	Up-regulated phosphorylation sites (FC>2, P<0.05)	Up-regulated phosphoproteins (FC>2, P<0.05)
LPS/TPG versus LPS	964	462	55	37

Figure 9. Number of significantly changed phosphoproteins in LPS/TPG-treated samples versus LPS-treated samples.

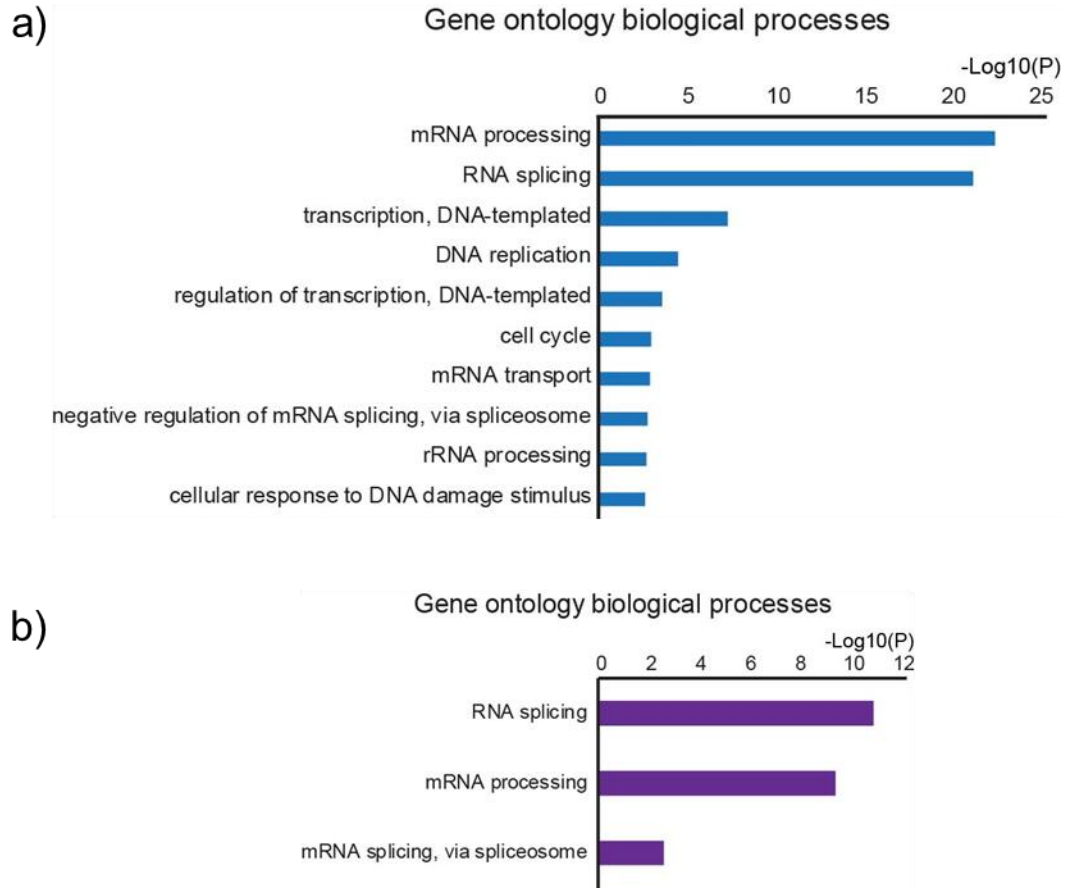


Figure 10. GO analysis indicating the down-regulated and up-regulated biological processes in the LPS/TPG-treated macrophages.

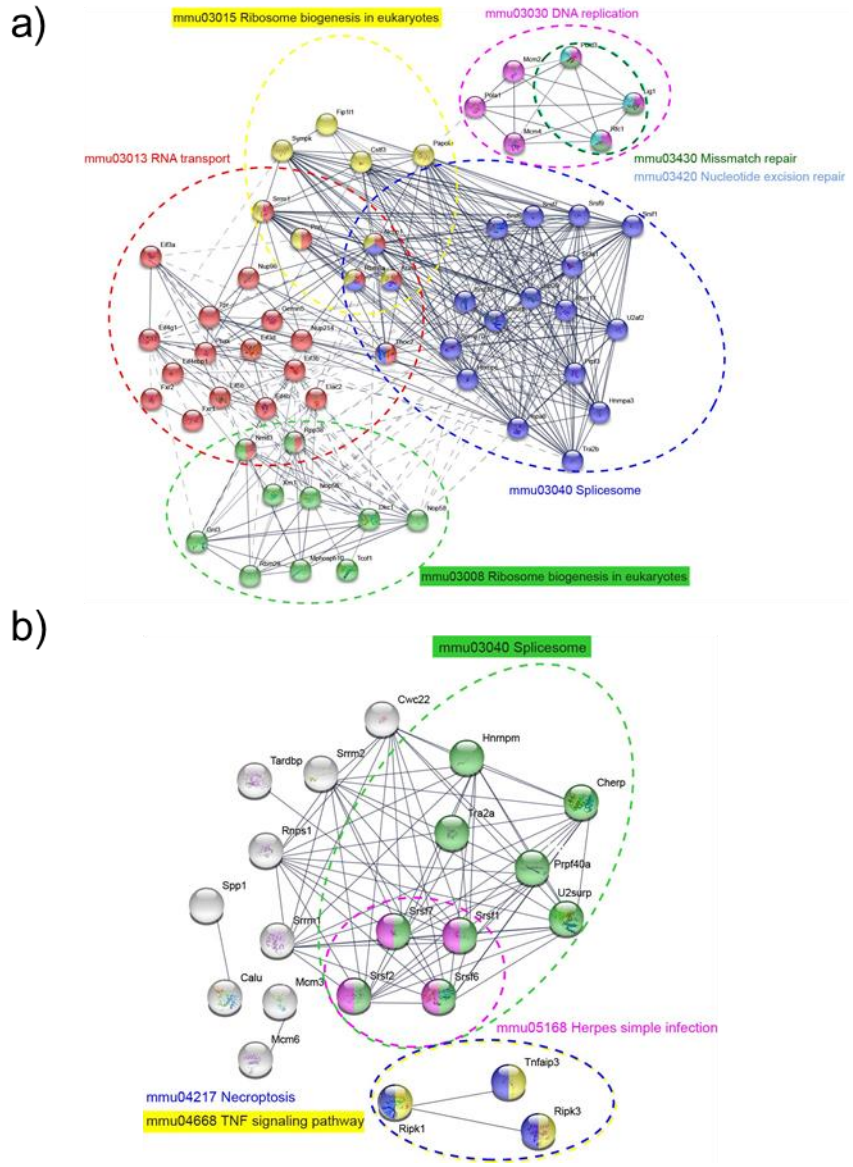


Figure 11. Down-regulated (a) and up-regulated KEGG pathways (b) in the LPS/TPG-treated macrophages.

Supplemental Information

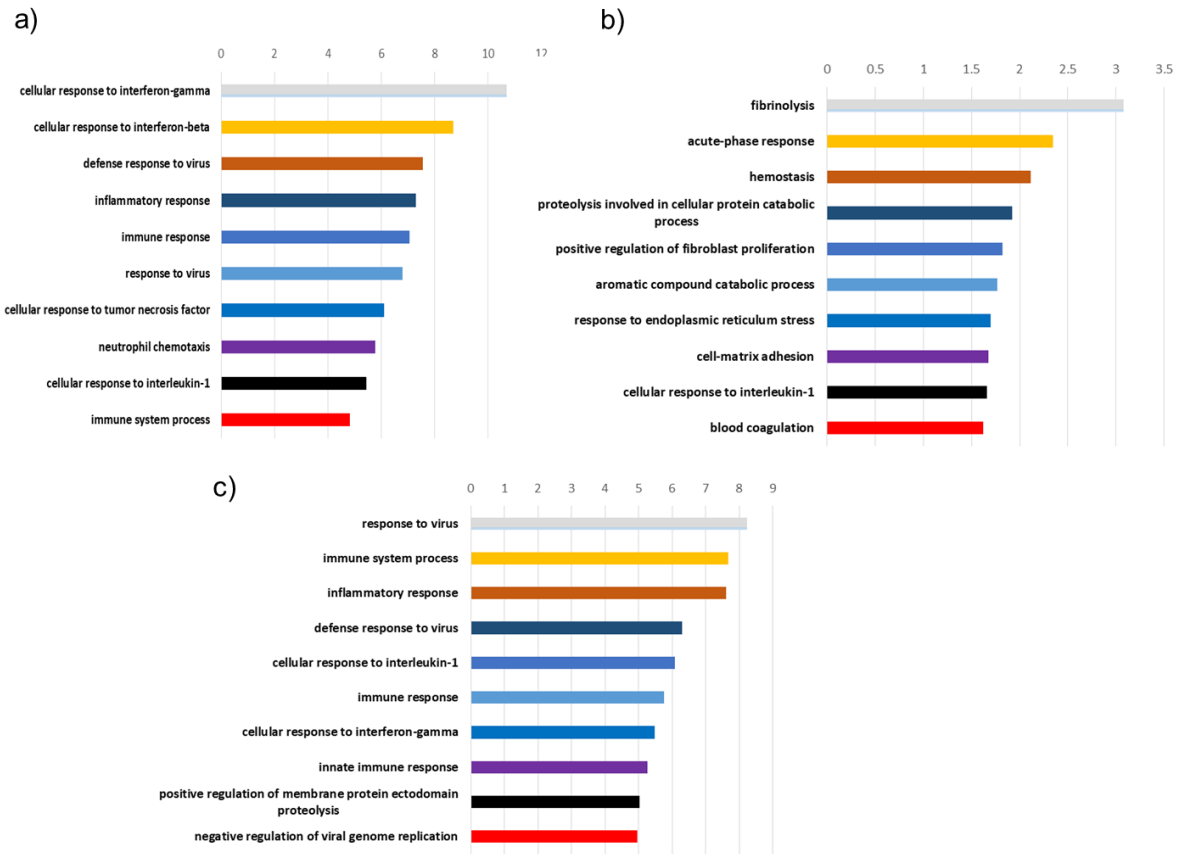


Figure S1. GO biological processes of significantly changed proteins in TPG-treated samples (a), LPS-treated samples (b) and LPS/TPG-treated samples (c).

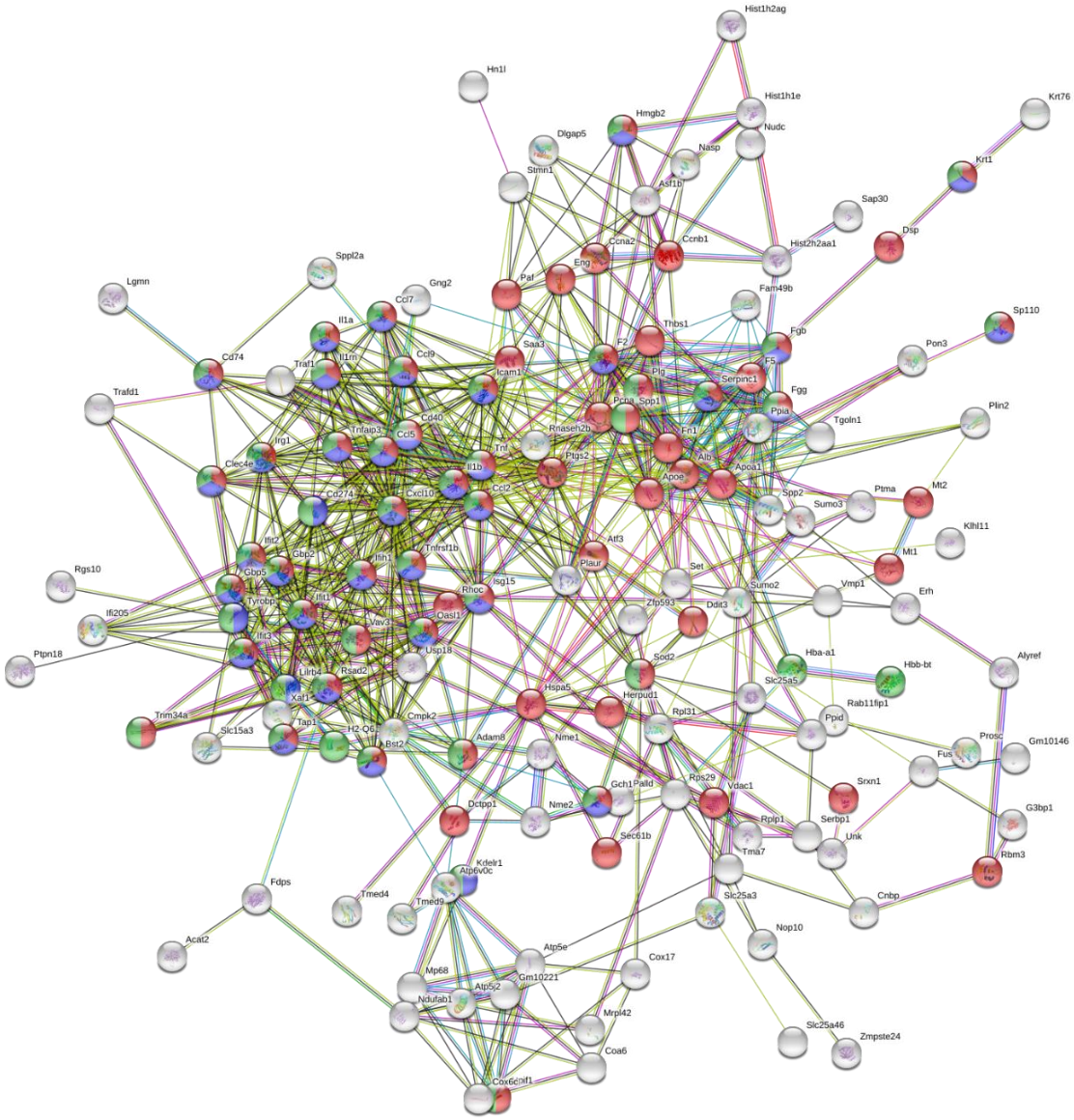


Figure S2. Functional analysis indicated that significantly changed proteins in LPS/TPG-treated samples versus WT samples were involved in response to stress (red), immune response (blue) and immune system process (green).

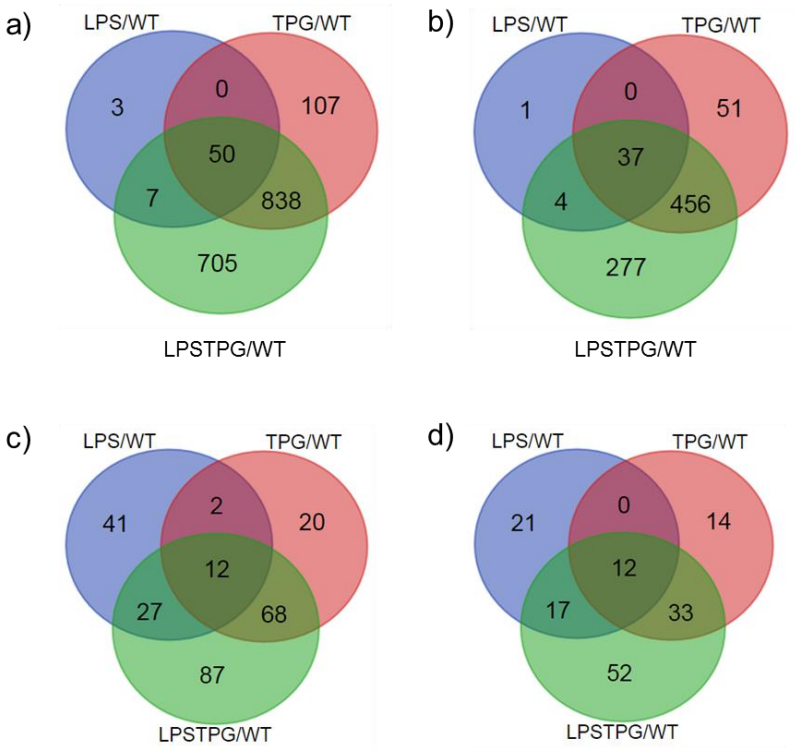


Figure S3. Number of down-regulated phosphorylation sites (a) and phosphorylated proteins (b), and up-regulated phosphorylation sites (c) and phosphorylated proteins (d) in three differentially treated samples.

Chapter 6

Mass Spectrometry Imaging of N-glycans From Formalin-fixed Paraffin-embedded Tissue Sections Using a Novel Subatmospheric Pressure Ionization Source

Adapted from: Yatao Shi, Zihui Li, Mildred A Felder, Qinying Yu, Xudong Shi, Yajing Peng, Qinjingwen Cao, Bin Wang, Luigi Puglielli, Manish S Patankar, Lingjun Li. High-resolution mass spectrometry imaging of N-glycans from formalin-fixed paraffin-embedded tissue sections using a novel sub-atmospheric pressure ionization source. *Anal Chem*. **2019**, 91, 20, 12942-12947.

Abstract

N-linked glycosylation, featuring various glycoforms, is one of the most common and complex protein post-translational modifications (PTMs) controlling protein structures and biological functions. It has been revealed that abnormal changes of protein N-glycosylation patterns are associated with many diseases. Hence, unraveling the disease-related alteration of glycosylation, especially the glycoforms, is crucial and beneficial to improve our understanding about the pathogenic mechanisms of various diseases. In past decades, given the capability of in-situ mapping of biomolecules and their region-specific localizations, matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) has been widely applied to the discovery of potential biomarkers for many diseases. In this study, we coupled a novel subatmospheric pressure (SubAP)/MALDI source with a Q Exactive HF hybrid quadrupole-orbitrap mass spectrometer for in-situ imaging of N-linked glycans from formalin-fixed paraffin-embedded (FFPE) tissue sections. The utility of this new platform for N-glycan imaging analysis was demonstrated with a variety of FFPE tissue sections. A total of 55 N-glycans were successfully characterized and visualized from a FFPE mouse brain section. Furthermore, 29 N-glycans with different spatial distribution patterns could be identified from a FFPE mouse ovarian cancer tissue section. High-mannose N-glycans exhibited elevated expression levels in the tumor region, indicating the potential association of this type of N-glycans with tumor progression.

Introduction

N-linked glycosylation is a type of common post-translational modification (PTM), which plays an important role in various physiological processes. Prior studies revealed that aberrant glycosylation was associated with many diseases including Alzheimer's disease,[1] cancers[2,3] and cardiovascular diseases.[4] Although electrospray ionization mass spectrometry (ESI-MS)-based glycomic[5,6] and glycoproteomic analyses[7-10] have been applied to investigate N-glycan expression changes between healthy and disease specimens, the homogenization of the tissue samples conducted in the vast majority of previous studies inevitably result in the loss of spatial information regarding N-glycan localization. The region-specific N-glycan distribution patterns, especially in the disease area, could be critical in discovering N-glycan biomarkers and revealing disease mechanisms.

Complementary to ESI-MS technique, matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) technique has been introduced for in-situ biomolecular analysis due to its capability of retaining and directly visualizing biomolecules in tissue sections.[11-13] Depending on the source pressure, MALDI-MS can be divided into three types: vacuum MALDI-MS, atmospheric MALDI-MS (AP/MALDI-MS) and subatmospheric pressure MALDI-MS (SubAP/MALDI-MS). Vacuum MALDI-MS is the one that commonly used, and generally has higher sensitivity than that of AP/MALDI-MS, enabled by the low pressure (less than 1.33 Pascal (Pa)) in the source.[14] However, to gain the high vacuum in the source, it typically requires a dedicated MALDI MS system to perform these in-situ profiling and imaging experiments. Different from the traditional vacuum MALDI source, the AP/MALDI source and SubAP/MALDI

source are independent of mass spectrometer and can be switched back and forth with an ESI source on an MS instrument platform, providing greater instrumental versatility and flexibility to allow both ESI-based analysis and MALDI-based analysis on a single mass spectrometer. Furthermore, the SubAP/MALDI source is operated at a subatmospheric pressure (133 Pa-1330 Pa) and offers better sensitivity than AP/MALDI source, making it a more attractive alternative to the traditional vacuum MALDI source.[15]

In past decades, MALDI-MS has been successfully applied to map the spatial distributions of various biomolecules on tissue sections, such as lipids,[16,17] small metabolites[18,19] and neuropeptides.[20] Recently, the scope of the vacuum MALDI-MS, has been successfully extended to the in-situ N-glycan imaging analysis,[21-23] while the application of SubAP/MALDI-MS to N-glycan imaging has rarely been reported, even though its sensitivity has been proven to be sufficient for many biological applications.[15] Considering the advantages of SubAP/MALDI-MS, it could be beneficial to systematically evaluate the potential of SubAP/MALDI-MS, as a complementary tool of vacuum MALDI-MS, in the imaging of N-glycans from tissue sections.

In this study, a new platform was developed by coupling the novel SubAP/MALDI source to a Q Exactive HF (QE HF) hybrid quadrupole-Orbitrap mass spectrometer. After optimization with N-glycan mixtures, this platform was successfully employed to characterize and visualize N-glycans released from formalin-fixed paraffin-embedded (FFPE) tissue sections including mouse brain and murine ovarian cancer tissue sections. Our results, for the first time, demonstrated the applicability of the SubAP/MALDI-MS platform for MS imaging (MSI) of N-glycans with high spatial resolution and mass accuracy.

Materials and methods

Chemicals and reagents

Methanol (MeOH), ethanol (EtOH), acetonitrile (ACN), formic acid (FA), sodium chloride (NaCl), trifluoroacetic acid (TFA), citric acid and ammonium bicarbonate were purchased from Fisher Scientific (Pittsburgh, PA). Xylene and microcon-30kDa centrifugal filter unit were purchased from EMD Millipore Sigma (Burlington, MA). Dithiothreitol and PNGase F were purchased from Promega (Madison, WI). Distilled water mentioned in this work was Milli-Q water from a Millipore filtration system (Bedford, MA). α-cyano-4-hydroxycinnamic acid (CHCA), Tris (2-carboxyethyl) phosphine (TCEP), triethylammonium bicarbonate buffer (TEAB) and bovine thyroglobulin (BTG) were purchased from Sigma-Aldrich (St. Louis, MO). 2,5-dihydroxybenzoic acid (DHB) was purchased from Acros Organics (Morris Plains, NJ). All reagents were used without additional purification. Microscope glass slides were purchased from VWR international, LLC (Radnor, PA). Indium tin oxide (ITO)-coated glass slides were purchased from Delta technologies (Loveland, CO).

Preparation of FFPE tissue sections

Mouse studies were approved by the Animal Care and Use Committee of the University of Wisconsin-Madison. For the FFPE mouse brain section, mouse was anesthetized with carbon dioxide. The brain was collected immediately after transcardial PBS perfusion, fixed overnight in 10% neutral buffered formalin, and paraffin embedded using standard techniques. 6 µm coronal tissue sections were prepared using a microtome (HistoCore MULTICUT, Leica Biosystems).

For the development of ovarian tumor in a mouse model, syngeneic murine ovarian cancer cells (Mouse Ovarian Surface Epithelial Cells, known as MOSEC[24]) were suspended in phosphate buffered saline and injected into the peritoneal cavity of 8 to 10-week-old female C57BL/6 mice, $10x10^6$ cells/animal. The MOSEC were a gift from Dr. Kathy Roby, Kansas University Medical Center, Kansas City, KS. The mice were housed under controlled conditions in the vivarium and monitored over a period of 2-3 months for the development of ascites as determined by behavioral changes/indications of discomfort, weight gain and presence of a distended belly, and body condition evaluation. Accumulation of ascites typically occurred 10-12 weeks post implantation of the cancer cells. Animals showing significant accumulation of ascites and with worsening body condition were sacrificed and tumors from the peritoneal cavity were obtained during necropsy. The tumors were immediately stored in formalin and embedded in paraffin the following day. 6 μm tissue section was cut from the paraffin block, adhered to the ITO-coated slide, and used for mass spectrometry-based N-glycan imaging.

Preparation of N-glycans released from bovine thyroglobulin

N-glycans were released from glycoprotein standard-bovine thyroglobulin (BTG) with slightly modified filter-aided N-Glycan separation (FANGS) strategy.[25] Glycoproteins were dissolved in water to have a concentration of 2 μg μL^{-1} and mixed with 5 μL of 0.5 M TCEP. Heat-denaturation was performed by switching sample tubes between 100°C water bath and room temperature for four cycles of 15 seconds each. The mixture was then loaded onto a 30 kDa molecular weight cutoff (MWCO) filter and buffer exchanged with 100 μL 0.5 M triethylammonium bicarbonate buffer (TEAB) by centrifugation at 14000g for three cycles (15

minutes for each) to remove contaminants with low molecular weight. 4 μL of PNGase F in 96 μL of 0.5 M TEAB buffer was added to each filter and incubated at 37°C overnight. N-glycans were separated with deglycosylated proteins and eluted by centrifugation at 14000g for 10 minutes. The filter was then washed with 100 μL 0.5 M TEAB buffer to ensure complete elution of N-glycans. Two fractions were combined and dried *in vacuo*. The released N-glycans were analyzed using the SubAP/MALDI QE HF mass spectrometer.

Pretreatment of FFPE tissue section for SubAP/MALDI N-glycan imaging

SubAP/MALDI imaging of N-glycans from FFPE tissue section was performed using workflow illustrated in **Figure S1**. 6μm FFPE tissue sections were placed on a hot plate and heated at 65°C for 20 min. After cooling down, FFPE tissue sections were consecutively washed twice with xylene, ethanol, 95% ethanol and 70% ethanol to remove paraffin. Following that, the antigen retrieval process was performed by boiling tissue sections in 20 mM freshly prepared citric acid buffer for 1 hour. Enzyme and matrix application were performed by a robotic TM sprayer system (HTX Technologies, Carrobo, NC). For enzyme deposition, 20 μl PNGase F dissolved in 330 μl 50 mM ammonium bicarbonate solution was sprayed at a flowrate of 0.02 ml min^{-1}, and in total 16 passes were performed. The nozzle temperature was set to 35 °C with a moving velocity of 800 mm min^{-1}. Then, FFPE tissue sections were incubated in a humidity chamber at 37°C for 12 hours. CHCA dissolved in ACN: H_2O: TFA (v: v: v, 50: 50: 0.1) solution at a concentration of 7 mg ml^{-1} was used as the matrix for N-glycan imaging. 24 passes of matrix spraying were performed at a flow rate of 0.05 ml min^{-1}, and 30 seconds drying time was set between each pass.

The nozzle temperature was set to 80°C with a moving velocity of 800 mm min^{-1}. After matrix spraying, tissue slides were dried in a vacuum chamber for 30 min and stored at -20°C until use.

Mass spectrometry

The QE HF mass spectrometer is coupled with a novel SubAP/MALDI (ng) ion source equipped with a 355 nm Nd:YAG laser (**Figure 1**). The laser spot size is 10 μm and the maximum output frequency reaches 10 kHz. Spiral plate motion was used for profiling experiments, while constant speed raster mode was employed for the imaging of tissue sections. A microscopic slide in the sample holder is approximately 2 mm away from the MS inlet capillary. Enabled by two novel design features, the SubAP/MALDI source has better sensitivity than the regular AP/MALDI source. Firstly, it operates at pressures from 1 Torr to 10 Torr instead of regular atmospheric pressure. Secondly, an ion funnel interface is designed to improve the ion collection and transfer into the mass spectrometer. The source is controlled by the Target software (MassTech Inc., Columbia, MD), and the ImageQuest software (Thermo Fisher Scientific, Waltham, MA) is applied to construct the images of target analytes using the XY coordinates of laser spots along with the acquired MS data file. For N-glycan imaging experiments, laser energy of 20 % and repetition rate of 200 Hz were used. Ion funnel parameters were set as below: Voltage (V) values for V1 through V7 were set as 0, 0, 5, 10, 170, 200 and 300, respectively; RF amplitude was set as 250 V; QE-HF mass spectrometer was operated at positive full MS scan mode with the mass range of m/z 1100-3000. The capillary temperature and S-lens RF level were set to 200 °C and 50, respectively. Automatic gain control (AGC) of 5e6 was used to ensure that the maximum ion

injection time of 300 ms could be reached at each pixel. Resolution was sample-dependent and will be discussed in Results and discussion section.

Histology staining

H&E staining was performed as described previously.[26] Tissue sections were deparaffinized through a xylene and graded ethanol series, then rinsed in hematoxylin, in 1% acid alcohol, and then in eosin. After rinsing in running water to remove excessive staining, slides were dehydrated in graded ethanol and xylene.

Data processing

MS spectra were processed by Xcalibur (Thermo Scientific, Bremen, Germany), and observed N-glycans were annotated by using the GlycoWorkbench[27] software with less than 10 ppm mass tolerance. Signal intensities of N-glycans were normalized to total ion chromatogram (TIC) and ImageQuest (Thermo Scientific, Bremen, Germany) was used to construct N-glycan images with the mass tolerance window of 5 ppm. N-glycan compositions were tentatively identified by searching against UniCarbKB database.

Results and Discussion

Optimization of operating pressure

SubAP/MALDI source was operated at subatmospheric pressure normally ranging from 133 Pa to1330 Pa. Given that the ion source pressure is critical to ion transport efficiency, it is necessary to examine the impact of source pressure on the N-glycan signal intensities and select an optimal pressure for N-glycan imaging analysis. When coupled to QE HF mass spectrometer, a minimum source pressure of 386 Pa was required to maintain the stability of the entire platform.

Hence, the pressure optimization was initiated from 400 Pa using N-glycans released from bovine thyroglobulin (A representative MS spectrum is shown in **Figure S2**). N-glycan mixture was mixed with CHCA, and further spotted onto sample plate for analysis. Ion abundance of 4 dominant N-glycans were normalized to TIC chromatogram, and used for pressure optimization. As shown in **Figure 2a-c**, TIC-normalized ion abundance and ion abundance of N-glycans significantly dropped following the increase of operating pressure, while no apparent change of TIC intensities was observed. This result suggested that relatively lower pressure could improve the sensitivity of SubAP/MALDI measurements probably due to enhanced transportation of N-glycans. Therefore, to balance the stability and sensitivity of the platform, operating pressure of 400 Pa was selected for N-glycan analysis.

Matrix optimization

DHB and CHCA are two common matrices used for MALDI imaging of small biomolecules. In general, high laser energy is required to ionize analytes co-crystalized with DHB matrix, while lower laser energy is needed for the ionization of CHCA-coated analytes. The performance of these two matrices in N-glycan imaging on the SubAP/MALDI-MS platform was studied. When CHCA was used, increasing laser energy improved TIC intensities (**Figure 2f**), though the optimal N-glycan ion abundance, either absolute or normalized, was obtained at the laser energy of 6.8 micro Joules per square centimeter ($\mu J/cm^2$) **(Figure 2d-e**). For DHB, higher TIC-normalized ion abundance of N-glycans were observed due to the much lower TIC intensities in comparison to CHCA (**Figure 2g**). Ion abundance and TIC augmented following the increase of laser energy, while higher energy was required to get similar results as that of CHCA (**Figure**

2h-i). Thus, to get longer laser lifespan and higher glycan signal intensities, matrix of CHCA and laser energy of 6.8μJ/cm^2 were used for SubAP/MALDI-based N-glycan imaging.

High-resolution imaging of N-glycans from FFPE mouse brain sections

Recent N-glycoproteomic studies revealed that N-glycoproteins with heterogeneous N-glycans were highly expressed in mouse brain.[28-31] Therefore, to test the applicability of the novel SubAP/MALDI-MS platform for high resolution MS imaging acquisition, a 6 μm-thick FFPE mouse brain coronal section was prepared and treated by PNGase F to release N-glycans for imaging analysis with a pixel size of 25μm. A total of 55 N-glycans including high-mannose N-glycans, fucosylated N-glycans and sialylated glycans were detected and annotated by the accurate mass matching (**Figure S3**, **Table S1**). The new SubAP/MALDI-MS platform detected more N-glycans in comparison to previous N-glycan imaging studies of either frozen or FFPE mouse brain section using vacuum MALDI-MS platforms (**Figure 3a**, **Table S2**).[32,33] Also, as shown in **Figure 3b**, the new SubAP/MALDI source is capable of detecting most N-glycans identified by vacuum MALDI-MS. Furthermore, by matching with the optical image of the mouse brain section (**Figure 3c**), N-glycans with different distribution patterns on the brain tissue section could be revealed and representative N-glycans were shown in **Figures 3d**-**3h**. These results manifested that the new instrument platform had adequate sensitivity for the detection and visualization of N-glycans from FFPE tissue sections, and therefore could be employed as an alternative to vacuum MALDI-MS platform for *in-situ* N-glycan imaging analysis.

Mapping of N-glycan distribution on FFPE mouse ovarian cancer tissue section

N-glycosylation is a protein PTM that plays an important role in many devastating diseases. Accumulating evidence have indicated that aberrant alteration of N-glycosylation can occur during the onset and progression of ovarian cancer.[22,34-36] In particular, high-mannose N-glycans, which are synthesized in the endoplasmic reticulum (ER) and are crucial for proper protein folding, have been shown to be dominantly expressed in the cancer area.[22] Herein, a mouse model with ovarian cancer was established and a 6 μm FFPE ovarian tissue section containing tumor region was prepared. Following that, the new SubAP/MALDI-MS platform was employed to map the distribution patterns of N-glycans in the tissue section. Considering that the mouse ovarian tissue section was bigger, but less heterogeneous than the mouse brain section, a pixel size of 50 μm was selected without significant compromise on the image quality. In total, 29 N-glycans, including complex N-glycans and high-mannose N-glycans, were successfully detected and annotated by the accurate mass matching (**Figure 4**, **Table S3**). More homogenous distributions of complex N-glycans on tissue section were observed and shown in **Figures 5a-h**. In contrast, high-mannose N-glycans, from $Hex_4HexNAc_2$ to $Hex_{10}HexNAc_2$, were highly expressed in the cancer region though no apparent accumulation was observed for $Hex_3HexNAc_2$ as the simplest N-glycan (**Figures 5i-p**). This observation is consistent with the previous finding that increasing N-glycan branches could lead to structural and functional changes of N-glycoproteins, contributing to the progression of cancers.[3,37] Overall, our MSI results revealed that alteration of high-mannose N-glycans was involved in mouse ovarian cancer development, which was consistent with the findings of previous N-glycomic profiling studies,[22,38,39] and further demonstrating the capability of the new platform for *in-situ* N-glycan imaging analysis.

Conclusions

A novel MS-based imaging platform was developed by integrating a new SubAP/MALDI source with a Q Exactive HF orbitrap mass spectrometer for the in-situ N-glycan imaging analysis with high resolution and high mass accuracy for the first time. Several parameters, such as the operation pressure of the SubAP/MALDI source, the type of matrix and laser energy, were optimized with a mixture of N-glycans released from a glycoprotein. After optimization, the new platform was successfully applied to map distribution patterns of N-glycans from FFPE mouse brain and ovarian cancer tissue sections. More N-glycans can be identified on brain section in comparison to that of commonly used vacuum MALDI-MS platforms, and high mannose N-glycans show accumulation in the tumor region. The new platform was demonstrated to have improved sensitivity for N-glycan profiling analysis from biological tissue sections. For future directions, by simply switching the PNGase F to trypsin, it is entirely possible to extend the application of the new platform from N-glycan imaging to peptide imaging, and even in-situ glycopeptide identification enabled by the powerful Q Exactive HF orbitrap mass spectrometer or similar high performance MS instrument platforms.

Acknowledgments

The authors would like to thank Dr. Eugene Moskovets and Dr. Vladimir Doroshenko from MassTech, Inc. for access to the SubAP MALDI source and technical support. The research was supported in part by NIH R01 DK071801, R56MH110215, RF1AG052324, and U01CA231081. The MALDI Orbitrap instrument was purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531). LL acknowledges a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

References

(1) Liang, H. C.; Russell, C.; Mitra, V.; Chung, R.; Hye, A.; Bazenet, C.; Lovestone, S.; Pike, I.; Ward, M. *J. Proteome. Res.* **2015**, *14*, 5063-5076.

(2) Lau, K. S.; Dennis, J. W. *Glycobiology* **2008**, *18*, 750-760.

(3) de Leoz, M. L.; Young, L. J.; An, H. J.; Kronewitter, S. R.; Kim, J.; Miyamoto, S.; Borowsky, A. D.; Chew, H. K.; Lebrilla, C. B. *Mol. Cell. Proteomics* **2011**, *10*, M110 002717.

(4) Yang, S.; Mishra, S.; Chen, L.; Zhou, J. Y.; Chan, D. W.; Chatterjee, S.; Zhang, H. *Anal. Chem.* **2015**, *87*, 9671-9678.

(5) Park, D.; Brune, K. A.; Mitra, A.; Marusina, A. I.; Maverakis, E.; Lebrilla, C. B. *Mol. Cell. Proteomics* **2015**, *14*, 2910-2921.

(6) Yang, S.; Hoti, N.; Yang, W.; Liu, Y.; Chen, L.; Li, S.; Zhang, H. *Clin Proteom* **2017**, *14*, 3.

(7) Yu, Q.; Wang, B.; Chen, Z.; Urabe, G.; Glover, M. S.; Shi, X.; Guo, L. W.; Kent, K. C.; Li, L. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 1751-1764.

(8) Sun, S.; Shah, P.; Eshghi, S. T.; Yang, W.; Trikannad, N.; Yang, S.; Chen, L.; Aiyetan, P.; Hoti, N.; Zhang, Z.; Chan, D. W.; Zhang, H. *Nat. Biotechnol.* **2016**, *34*, 84-88.

(9) Chen, Z.; Yu, Q.; Hao, L.; Liu, F.; Johnson, J.; Tian, Z.; Kao, W. J.; Xu, W.; Li, L. *Analyst* **2018**, *143*, 2508-2519.

(10) Hong, Q.; Ruhaak, L. R.; Stroble, C.; Parker, E.; Huang, J.; Maverakis, E.; Lebrilla, C. B. *J. Proteome. Res.* **2015**, *14*, 5179-5192.

(11) Jones, E. E.; Dworski, S.; Canals, D.; Casas, J.; Fabrias, G.; Schoenling, D.; Levade, T.; Denlinger, C.; Hannun, Y. A.; Medin, J. A.; Drake, R. R. *Anal. Chem.* **2014**, *86*, 8303-8311.

(12) Lukowski, J. K.; Weaver, E. M.; Hummon, A. B. *Anal. Chem.* **2017**, *89*, 8453-8458.

(13) Chen, B.; OuYang, C.; Tian, Z.; Xu, M.; Li, L. *Anal. Chim. Acta* **2018**, *1007*, 16-25.

(14) Keller, C.; Maeda, J.; Jayaraman, D.; Chakraborty, S.; Sussman, M. R.; Harris, J. M.; Ane, J. M.; Li, L. *Front Plant Sci*. **2018**, *9*, 1238.

(15) Li, G.; Cao, Q.; Liu, Y.; DeLaney, K.; Tian, Z.; Moskovets, E.; Li, L. *Rapid Commun Mass Spectrom.* **2019**, *33*, 327-335.

(16) Shi, Y.; Johnson, J.; Wang, B.; Chen, B.; Fisher, G. L.; Urabe, G.; Shi, X.; Kent, K. C.; Guo, L. W.; Li, L. *J. Proteome. Res.* **2019**, *18*, 1669-1678.

(17) Cerruti, C. D.; Benabdellah, F.; Laprevote, O.; Touboul, D.; Brunelle, A. *Anal. Chem.* **2012**, *84*, 2164-2171.

(18) Zhang, J.; Lamotte, L.; Dodds, E. D.; Lebrilla, C. B. *Anal. Chem.* **2005**, *77*, 4429-4438.

(19) Cornett, D. S.; Frappier, S. L.; Caprioli, R. M. *Anal. Chem.* **2008**, *80*, 5648-5653.

(20) Ye, H.; Hui, L.; Kellersberger, K.; Li, L. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 134-147.

(21) Scott, D. A.; Casadonte, R.; Cardinali, B.; Spruill, L.; Mehta, A. S.; Carli, F.; Simone, N.; Kriegsmann, M.; Del Mastro, L.; Kriegsmann, J.; Drake, R. R. *Proteomics Clin Appl* **2019**, *13*, e1800014.

(22) Everest-Dass, A. V.; Briggs, M. T.; Kaur, G.; Oehler, M. K.; Hoffmann, P.; Packer, N. H. *Mol. Cell. Proteomics* **2016**, *15*, 3003-3016.

(23) Holst, S.; Heijs, B.; de Haan, N.; van Zeijl, R. J.; Briaire-de Bruijn, I. H.; van Pelt, G. W.; Mehta, A. S.; Angel, P. M.; Mesker, W. E.; Tollenaar, R. A.; Drake, R. R.; Bovee, J. V.; McDonnell, L. A.; Wuhrer, M. *Anal. Chem.* **2016**, *88*, 5904-5913.

(24) Roby, K. F.; Taylor, C. C.; Sweetwood, J. P.; Cheng, Y.; Pace, J. L.; Tawfik, O.; Persons, D. L.; Smith, P. G.; Terranova, P. F. *Carcinogenesis* **2000**, *21*, 585-591.

(25) Abdul Rahman, S.; Bergstrom, E.; Watson, C. J.; Wilson, K. M.; Ashford, D. A.; Thomas, J. R.; Ungar, D.; Thomas-Oates, J. E. *J. Proteome. Res.* **2014**, *13*, 1167-1176.

(26) Titford, M. *Biotechnic & histochemistry : official publication of the Biological Stain Commission* **2005**, *80*, 73-78.

(27) Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S. M. *J. Proteome. Res.* **2008**, *7*, 1650-1659.

(28) Fang, P.; Wang, X. J.; Xue, Y.; Liu, M. Q.; Zeng, W. F.; Zhang, Y.; Zhang, L.; Gao, X.; Yan, G. Q.; Yao, J.; Shen, H. L.; Yang, P. Y. *Oncotarget.* **2016**, *7*, 38796-38809.

(29) Tian, Y.; Kelly-Spratt, K. S.; Kemp, C. J.; Zhang, H. *J. Proteome. Res.* **2010**, *9*, 5837-5847.

(30) Liu, J.; Wang, F.; Mao, J.; Zhang, Z.; Liu, Z.; Huang, G.; Cheng, K.; Zou, H. *Anal. Chem.* **2015**, *87*, 2054-2057.

(31) Qiao, X.; Tao, D.; Qu, Y.; Sun, L.; Gao, L.; Zhang, X.; Liang, Z.; Zhang, L.; Zhang, Y. *Proteomics* **2011**, *11*, 4274-4278.

(32) Powers, T. W.; Jones, E. E.; Betesh, L. R.; Romano, P. R.; Gao, P.; Copland, J. A.; Mehta, A. S.; Drake, R. R. *Anal. Chem.* **2013**, *85*, 9799-9806.

(33) Toghi Eshghi, S.; Yang, S.; Wang, X.; Shah, P.; Li, X.; Zhang, H. *ACS Chem. Biol.* **2014**, *9*, 2149-2156.

(34) Kui Wong, N.; Easton, R. L.; Panico, M.; Sutton-Smith, M.; Morrison, J. C.; Lattanzio, F. A.; Morris, H. R.; Clark, G. F.; Dell, A.; Patankar, M. S. *J. Biol. Chem.* **2003**, *278*, 28619-28634.

(35) Abbott, K. L.; Nairn, A. V.; Hall, E. M.; Horton, M. B.; McDonald, J. F.; Moremen, K. W.; Dinulescu, D. M.; Pierce, M. *Proteomics* **2008**, *8*, 3210-3220.

(36) Saldova, R.; Wormald, M. R.; Dwek, R. A.; Rudd, P. M. *Dis. Markers* **2008**, *25*, 219-232.

(37) Kaprio, T.; Satomaa, T.; Heiskanen, A.; Hokke, C. H.; Deelder, A. M.; Mustonen, H.; Hagstrom, J.; Carpen, O.; Saarinen, J.; Haglund, C. *Mol. Cell. Proteomics* **2015**, *14*, 277-288.

(38) Chen, H.; Deng, Z.; Huang, C.; Wu, H.; Zhao, X.; Li, Y. *Tumor biol.* **2017**, *39*, 1010428317716249.

(39) Anugraham, M.; Jacob, F.; Nixdorf, S.; Everest-Dass, A. V.; Heinzelmann-Schwarz, V.; Packer, N. H. *Mol. Cell. Proteomics* **2014**, *13*, 2213-2232.

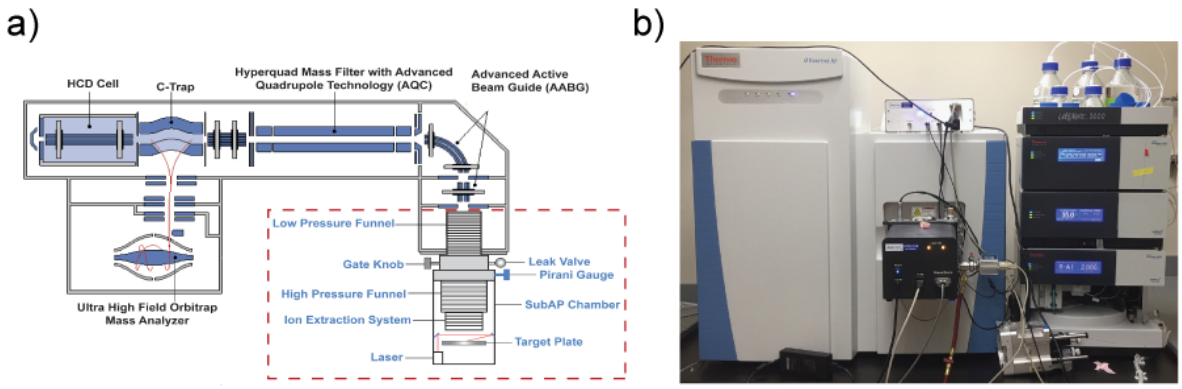


Figure 1. The new SubAP/MALDI-QEHF platform for N-glycan imaging. (a) Schematic illustration of the SubAP/MALDI-QEHF platform. (b) A snapshot showing the QEHF mass spectrometer integrated with the SubAP/MALDI source instead of an ESI source.

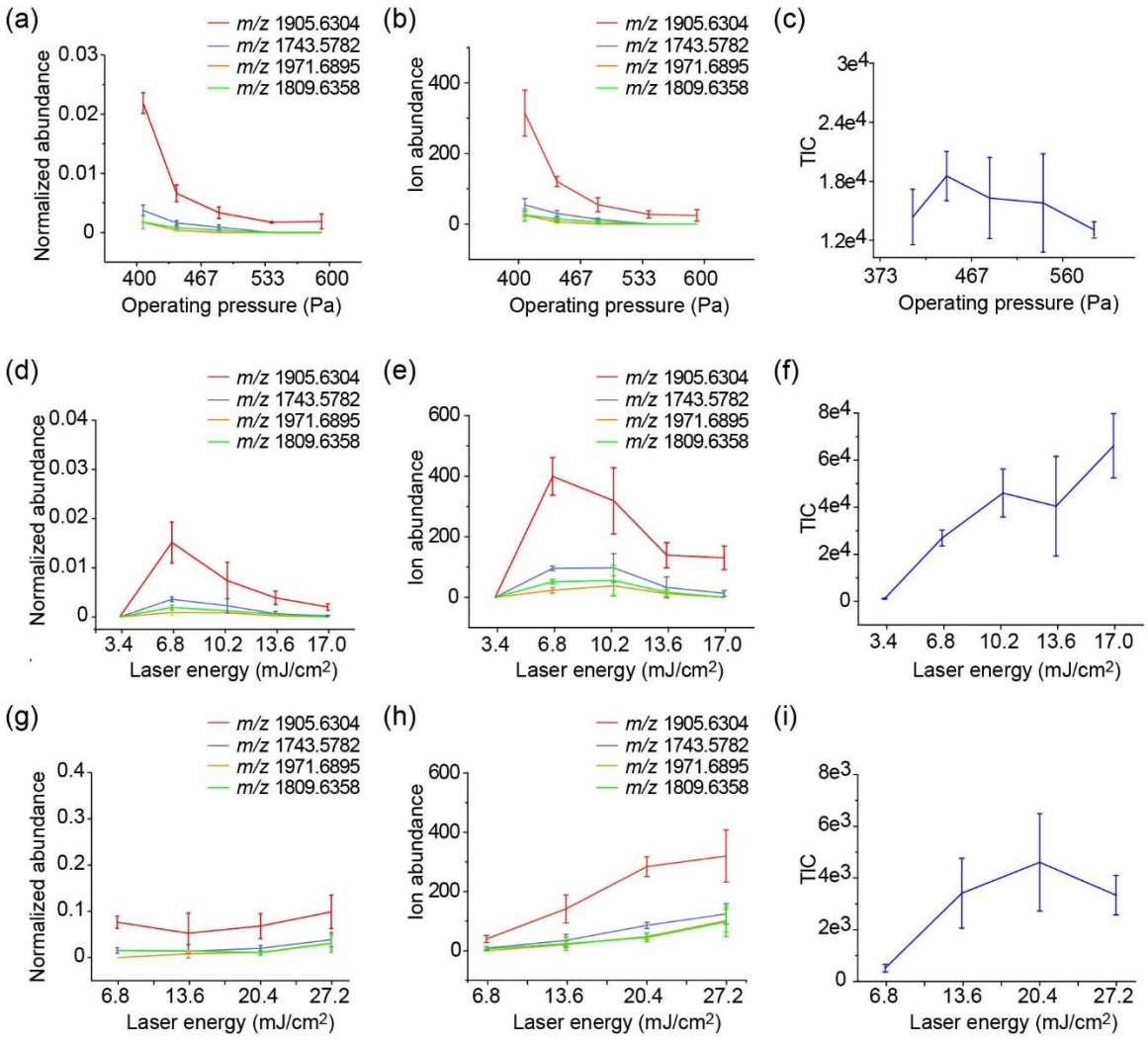


Figure 2. Parameter optimization for N-glycan analysis on SubAP/MALDI MS platform (n=3). The TIC-normalized abundance (a), ion abundance (b) and TIC (c) of four representative N-glycans detected at different operating pressure. The TIC-normalized abundance (d), ion abundance (e) and TIC (f) of four representative N-glycans detected at different laser energy. The TIC-normalized abundance (g), ion abundance (h) and TIC (i) of four representative N-glycans detected at different laser energy.

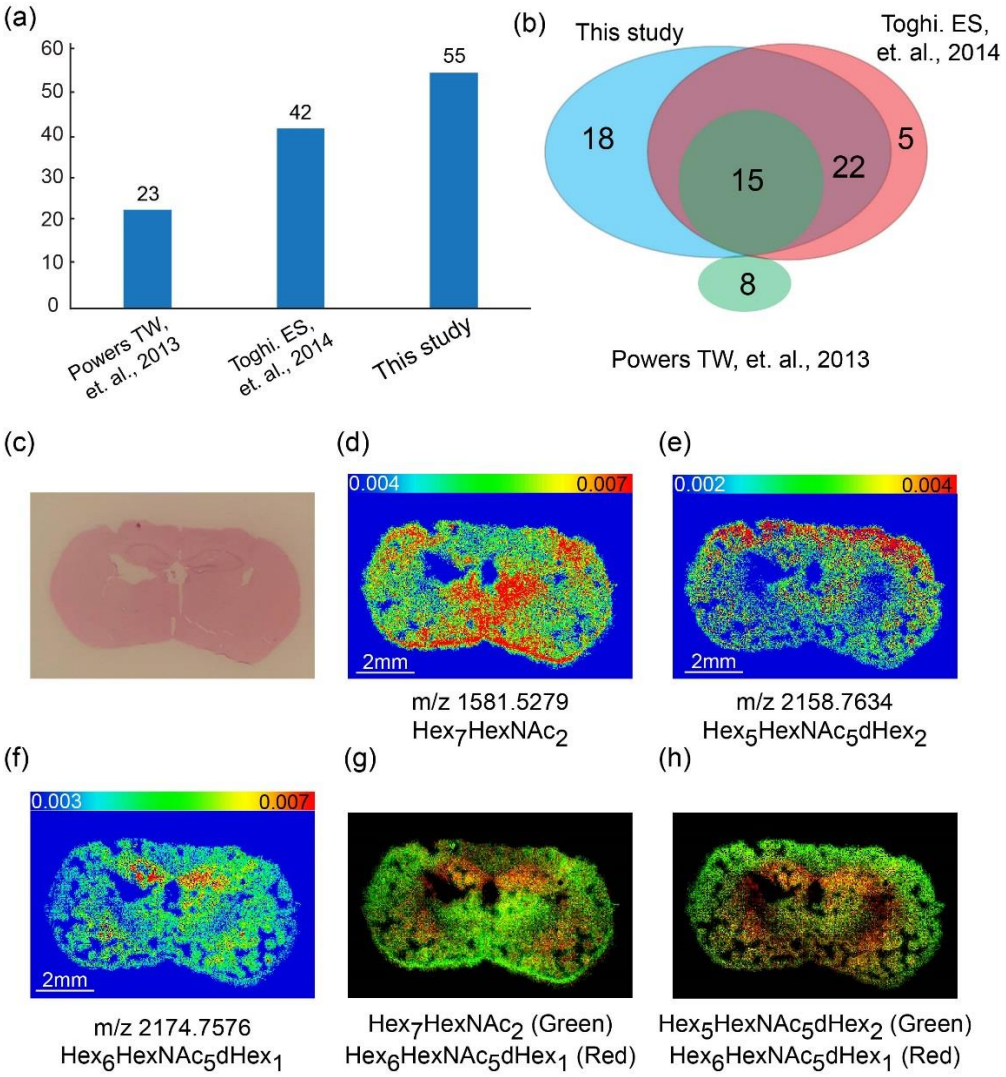


Figure 3. (a) More N-glycans were detected and imaged from FFPE mouse brain tissue section by using the novel SubAP/MALDI-MS platform. (b) The Venn diagram showing the overlap of N-glycans detected in this study with N-glycans reported in prior studies (Refs. 32 & 33) using vacuum MALDI-MS platform. (c) H&E stained FFPE mouse brain section post N-glycan imaging. (d-f) MS images of representative N-glycans detected from FFPE mouse tissue section. (g-h) Overlap of different N-glycan images clearly revealed different spatial distribution patterns of N-glycans on mouse brain section.

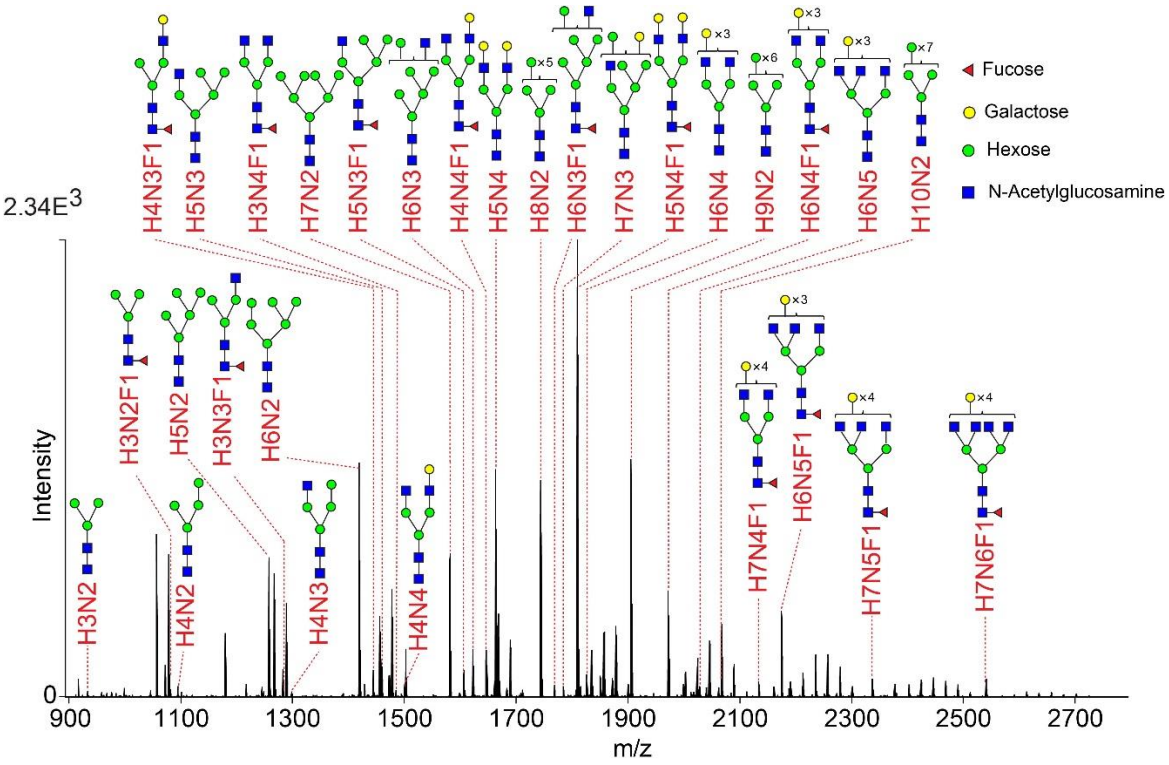


Figure 4. N-glycans detected from FFPE mouse tissue section with ovarian cancer. The annotated glycan compositions were tentatively identified by searching against UniCarbKB database. H: Hexose; N; N-Acetyl glucosamine; F: Fucose.

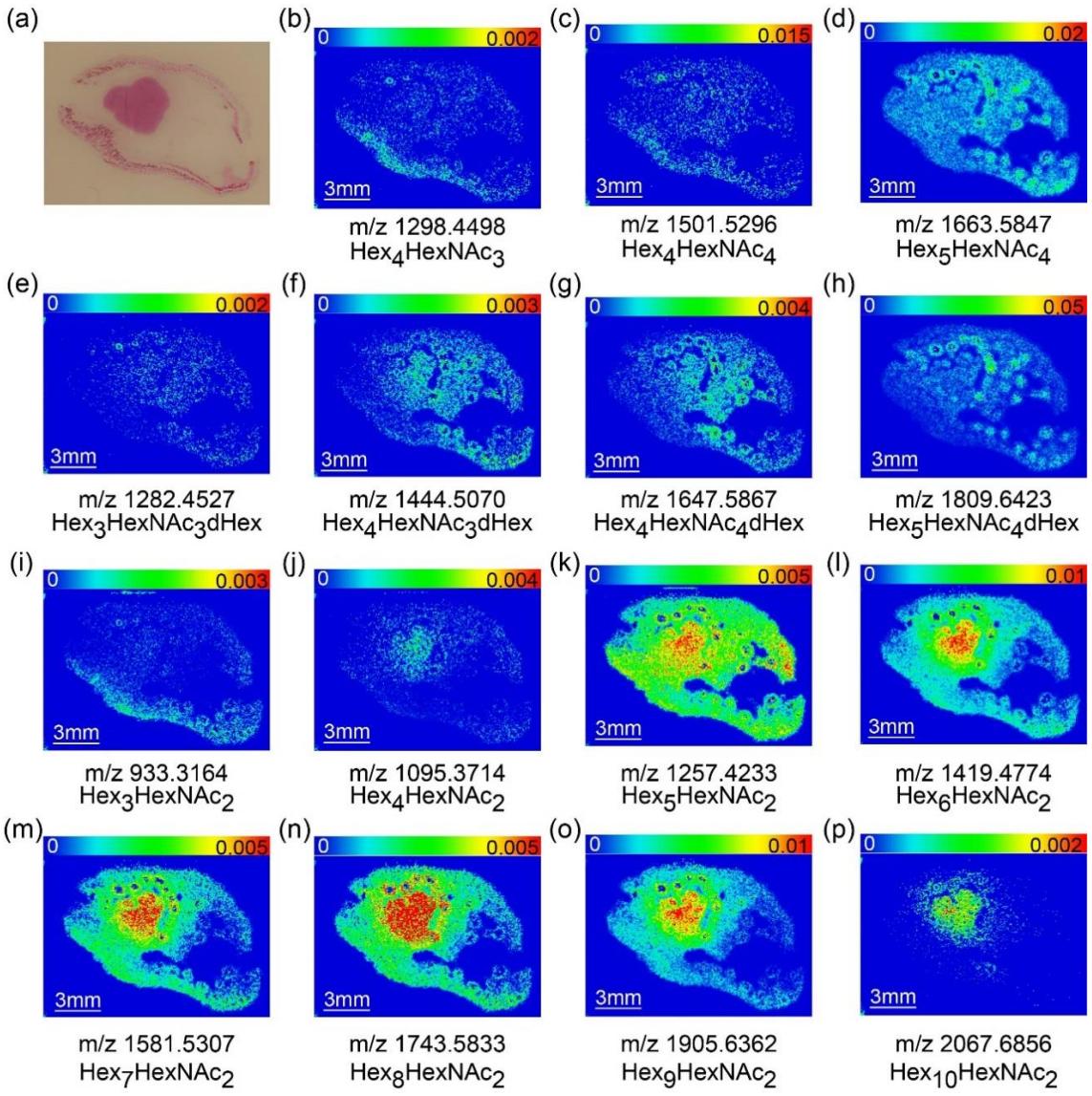


Figure 5. Images of N-glycans showing different spatial distribution patterns on FFPE mouse tissue section with ovarian cancer. (a) H&E stained FFPE mouse tissue section with ovarian cancer. (b-h) Complex N-glycans showed similar distribution in cancer area in comparison to peripheral area; (i-p) High mannose N-glycans accumulated in cancer area except $Hex_3HexNAc_2$.

Supplemental Information

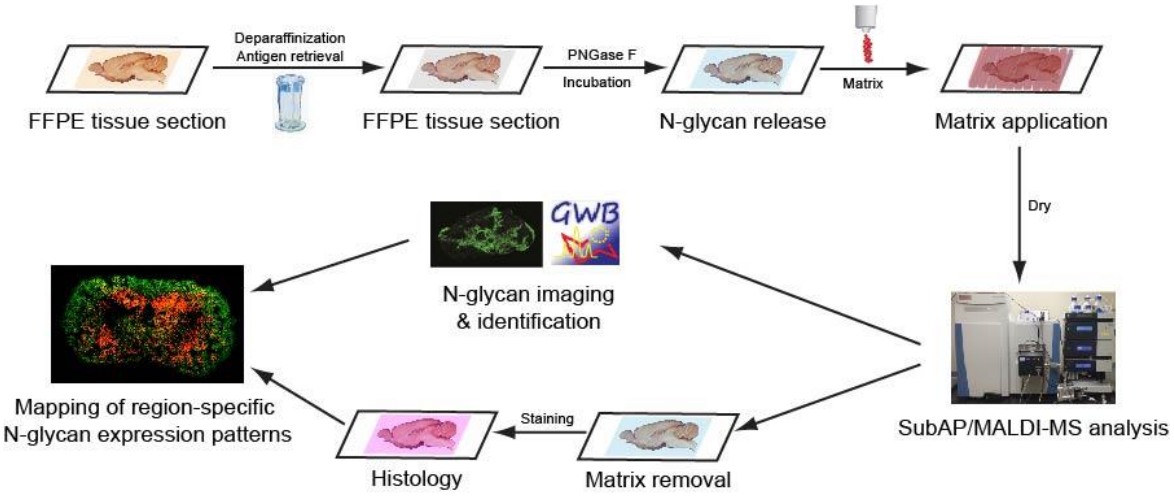


Figure S1. Typical workflow for imaging of N-glycans from FFPE tissue sections using the SubAP/MALDI-MS platform.

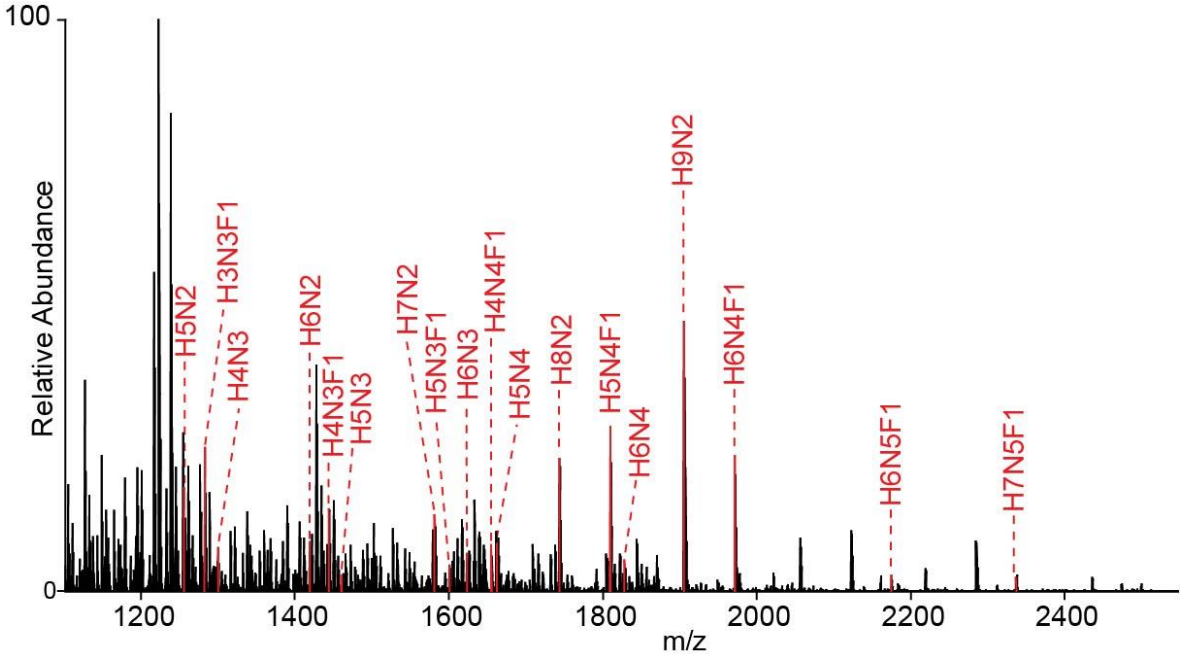


Figure S2. Representative MS spectrum of N-glycans released from bovine thyroglobulin collected with SubAP/MALDI-QEHF MS. H: Hexose; N: N-Acetylglucosamine; F: Fucose.

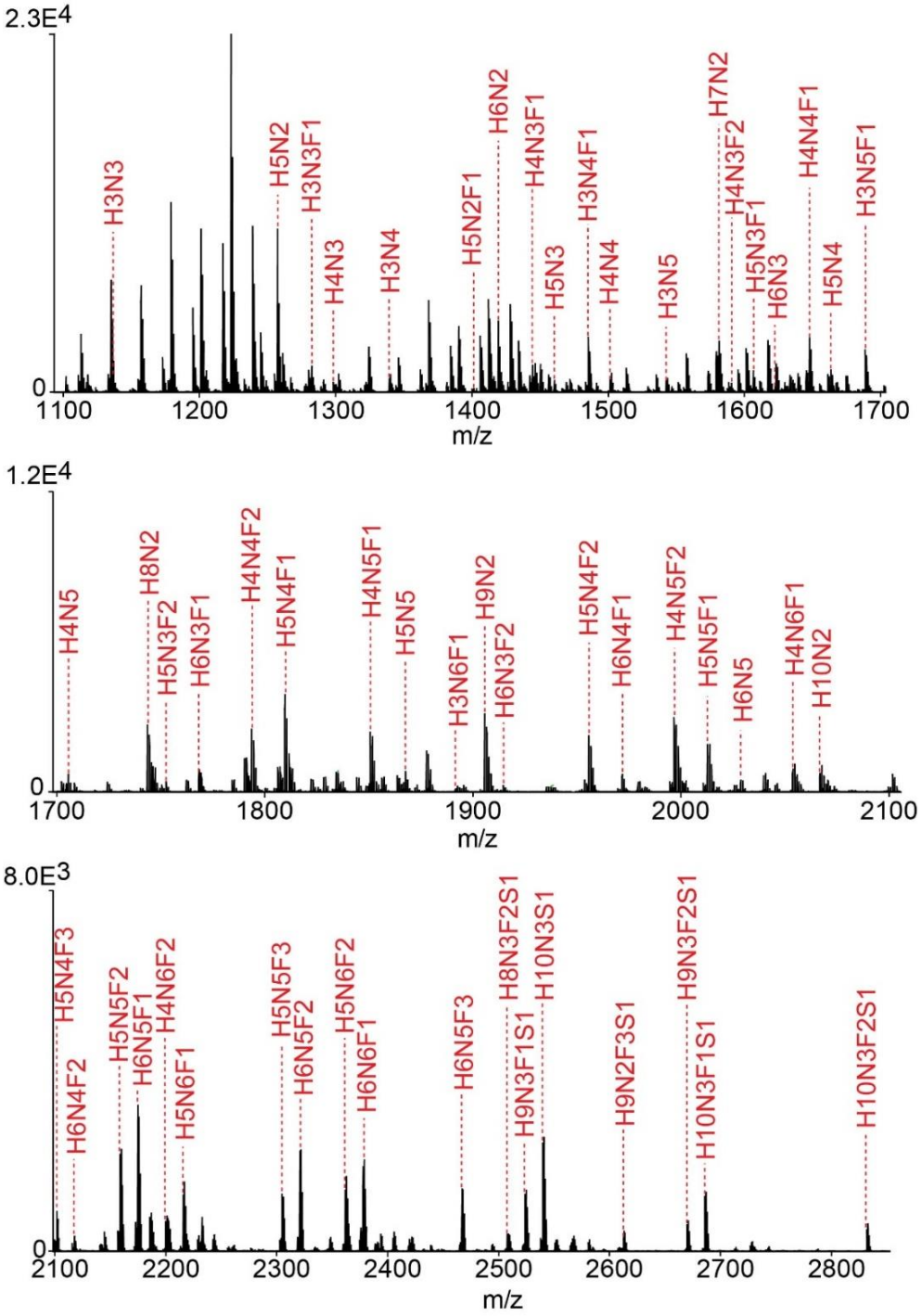


Figure S3. N-glycans detected from FFPE mouse brain section using SubAP/MALDI-MS platform. H: Hexose; N: N-Acetylglucosamine; F: Fucose; S: Sialic acid.

Table S1. N-glycans detected from FFPE mouse brain section using SubAP/MALDI-MS platform.

m/z (+Na)	Identified N-glycans	Mass deviation (ppm)	m/z (+Na)	Identified N-glycans	Mass deviation (ppm)
1136.3952	$Hex_3HexNAc_3$	1	1905.6299	$Hex_9HexNAc_2$	2
1257.4232	$Hex_5HexNAc_2$	0.5	1914.6621	$Hex_6HexNAc_3dHex_2$	4
1282.4527	$Hex_3HexNAc_3dHex_1$	1	1955.6928	$Hex_5HexNAc_4dHex_2$	2
1298.4488	$Hex_4HexNAc_3$	0.3	1971.6812	$Hex_6HexNAc_4dHex_1$	5.5
1339.4741	$Hex_3HexNAc_4$	1	1996.7157	$Hex_4HexNAc_5dHex_2$	4
1403.4740	$Hex_5HexNAc_2dHex_1$	4	2012.7104	$Hex_5HexNAc_5dHex_1$	4
1419.4784	$Hex_6HexNAc_2$	2	2028.6978	$Hex_6HexNAc_5$	7
1444.5090	$Hex_4HexNAc_3dHex_1$	1	2053.7325	$Hex_4HexNAc_6dHex_1$	6
1460.5029	$Hex_5HexNAc_3$	0.6	2067.6873	$Hex_{10}HexNAc_2$	0
1485.5353	$Hex_3HexNAc_4dHex_1$	1	2101.7387	$Hex_5HexNAc_4dHex_3$	8
1501.5285	$Hex_4HexNAc_4$	0	2117.7344	$Hex_6HexNAc_4dHex_2$	7
1542.5542	$Hex_3HexNAc_5$	0.6	2158.7634	$Hex_5HexNAc_5dHex_2$	6
1581.5279	$Hex_7HexNAc_2$	0.2	2174.7576	$Hex_6HexNAc_5dHex_1$	5
1590.5617	$Hex_4HexNAc_3dHex_2$	2	2199.7842	$Hex_4HexNAc_6dHex_2$	8
1606.5598	$Hex_5HexNAc_3dHex_1$	0	2215.7813	$Hex_5HexNAc_6dHex_1$	7
1622.5511	$Hex_6HexNAc_3$	2	2304.8148	$Hex_5HexNAc_5dHex_3$	8
1647.5874	$Hex_4HexNAc_4dHex_1$	0.6	2320.8122	$Hex_6HexNAc_5dHex_2$	7
1663.5806	$Hex_5HexNAc_4$	0.5	2361.8348	$Hex_5HexNAc_6dHex_2$	9
1688.6129	$Hex_3HexNAc_5dHex_1$	0	2377.8285	$Hex_6HexNAc_6dHex_1$	9
1704.6021	$Hex_4HexNAc_5$	3	2466.8614	$Hex_6HexNAc_5dHex_3$	10
1743.5799	$Hex_8HexNAc_2$	0.7	2507.8749 (+H)	$Hex_8HexNAc_3NeuAc_1dHex_2$	6
1752.6130	$Hex_5HexNAc_3dHex_2$	3	2523.8793 (+H)	$Hex_9HexNAc_3NeuAc_1dHex_1$	2
1768.6097	$Hex_6HexNAc_3dHex_1$	0	2539.8764 (+H)	$Hex_{10}HexNAc_3NeuAc_1$	1
1793.6421	$Hex_4HexNAc_4dHex_2$	1	2612.9016 (+H)	$Hex_9HexNAc_2NeuAc_1dHex_3$	7
1809.6370	$Hex_5HexNAc_4dHex_1$	1	2669.9299 (+H)	$Hex_9HexNAc_3NeuAc_1dHex_2$	5
1850.6631	$Hex_4HexNAc_5dHex_1$	1.5	2685.9269 (+H)	$Hex_{10}HexNAc_3NeuAc_1dHex_1$	4
1866.6469	$Hex_5HexNAc_5$	7.5	2831.9749 (+H)	$Hex_{10}HexNAc_3NeuAc_1dHex_2$	7
1891.6793	$Hex_3HexNAc_6dHex_1$	7			

Table S2. N-glycans reported in previous studies (Refs. 32 & 33) and detected by SubAP/MALDI-MS platform from FFPE mouse brain section. H: Hexose; N: N-Acetylglucosamine; F: Fucose; S: Sialic acid. √: Detected; ×: Undetected

N-glycan	Powers TW, et. al (Ref 32)	Toghi. ES, et. al (Ref 33)	This study	N-glycan	Powers TW, et. al (Ref 32)	Toghi. ES, et. al (Ref 33)	This study
H3N3	×	×	√	H6N5	×	×	√
H5N2	√	√	√	H4N6F1	×	√	√
H3N3F1	×	√	√	H10N2	×	×	√
H4N3	×	√	√	H5N4F3	√	√	√
H3N4	×	√	√	H6N4F2	√	√	√
H5N2F1	×	√	√	H5N5F2	√	√	√
H6N2	√	√	√	H6N5F1	×	√	√
H4N3F1	×	√	√	H4N6F2	×	√	√
H5N3	×	√	√	H5N6F1	√	√	√
H3N4F1	×	√	√	H6N6	×	√	×
H4N4	√	√	√	H5N5F3	×	√	√
H3N5	√	√	√	H6N5F2	×	√	√
H7N2	×	√	√	H5N6F2	×	√	√
H4N3F2	×	×	√	H6N6F1	×	×	√
H5N3F1	×	√	√	H6N5F3	×	√	√
H6N3	×	×	√	$H_5N_6F_1S_1$	√	×	×
H4N4F1	√	√	√	$H_5N_4F_1S_2$	√	×	×
H5N4	×	√	√	$H_5N_6F_2S_1$	√	×	×
H3N5F1	√	√	√	H5N6F3	×	√	×
H4N5	×	×	√	H6N6F2	×	√	×
H8N2	√	√	√	$H_7N_6S_1$	√	×	×
H5N3F2	×	×	√	$H_6N_5F_3S_1$	√	×	×
H6N3F1	×	√	√	$H_6N_5F_2S_2$	√	×	×
H4N4F2	√	√	√	$H_7N_6F_3S_1$	√	×	×
H5N4F1	√	√	√	$H_7N_6F_4S_1$	√	×	×
H4N5F1	×	√	√	H6N5F4	×	√	×
H5N5	×	×	√	H6N6F3	×	√	×
H3N6F1	×	×	√	H8N3F2S1	×	×	√
H9N2	×	√	√	H9N3F1S1	×	×	√
H6N3F2	×	×	√	H10N3S1	×	×	√
H5N4F2	√	√	√	H9N2F3S1	×	×	√

H6N4F1	×	√	√	H9N3F2S1	×	×	√
H4N5F2	√	√	√	H10N3F1S1	×	×	√
H5N5F1	×	√	√	H10N3F2S1	×	×	√

Table S3. List of N-glycans detected from FFPE mouse ovarian cancer tissue section.

m/z (+Na)	Identified N-glycans	Mass deviation (ppm)	m/z (+Na)	Identified N-glycans	Mass deviation(ppm)
933.3164	$Hex_3HexNAc_2$	0.6	1663.5847	$Hex_5HexNAc_4$	2
1079.3758	$Hex_3HexNAc_2dHex_1$	1	1743.5833	$Hex_8HexNAc_2$	1
1095.3714	$Hex_4HexNAc_2$	1.5	1768.6121	$Hex_6HexNAc_3dHex_1$	0
1257.4233	$Hex_5HexNAc_2$	0.5	1784.6084	$Hex_7HexNAc_3$	0.4
1282.4527	$Hex_3HexNAc_3dHex_1$	1	1809.6423	$Hex_5HexNAc_4dHex_1$	2
1298.4498	$Hex_4HexNAc_3$	0.5	1825.6335	$Hex_6HexNAc_4$	0
1419.4774	$Hex_6HexNAc_2$	1	1905.6362	$Hex_9HexNAc_2$	1
1444.5070	$Hex_4HexNAc_3dHex_1$	0	1971.6936	$Hex_6HexNAc_4dHex_1$	1
1460.5031	$Hex_5HexNAc_3$	0.7	2028.7106	$Hex_6HexNAc_5$	1.5
1485.5338	$Hex_3HexNAc_4dHex_1$	0	2067.6856	$Hex_{10}HexNAc_2$	0.5
1501.5296	$Hex_4HexNAc_4$	1	2133.7442	$Hex_7HexNAc_4dHex_1$	0
1581.5307	$Hex_7HexNAc_2$	1	2174.7692	$Hex_6HexNAc_5dHex_1$	1
1606.5614	$Hex_5HexNAc_3dHex_1$	1	2336.8212	$Hex_7HexNAc_5dHex_1$	1.3
1622.5553	$Hex_6HexNAc_3$	0	2539.8946	$Hex_7HexNAc_6dHex_1$	3
1647.5867	$Hex_4HexNAc_4dHex_1$	0			

Chapter 7

Mass Spectrometric Imaging Reveals Temporal and Spatial Dynamics of Bioactive Lipids in Arteries Undergoing Restenosis

Adapted from: Yatao Shi, Jillian Johnson, Bowen Wang, Bingming Chen, Gregory L. Fisher, Go Urabe, Xudong Shi, K. Craig Kent, Lian-Wang Guo and Lingjun Li. Mass Spectrometric Imaging Reveals Temporal and Spatial Dynamics of Bioactive Lipids in Arteries Undergoing Restenosis. *J. Proteome Res* **2019**, 18, (4), 1669-1678.

Abstract

Restenosis, or re-narrowing of the arterial lumen, is a common recurrent disease following balloon angioplasty and stenting treatments for cardiovascular disease. A major technical barrier for deciphering restenotic mechanisms is the dynamic, spatial profiling of bioactive lipids in the arterial wall, especially in small animals. Here, applying matrix-assisted laser desorption/ionization mass spectrometric imaging (MALDI-MSI), we conducted the first lipidomic study of temporal-spatial profiling in a small animal model of angioplasty-induced restenosis. Cross-sections were collected 3, 7, and 14 days after balloon angioplasty of rat carotid arteries. MALDI-MSI analyses showed that diacylglycerols (DAGs), signaling lipids associated with restenosis, and lysophosphatidylcholines (LysoPCs), whose function was uncharacterized in restenosis, dramatically increased at post-angioplasty day 7 and day 14 in the neointimal layer of balloon-injured arteries compared to uninjured control. In contrast, sphingomyelins (SMs) did not increase, but rather decreased at day 3, day 7, and day 14 in injured arteries versus the uninjured controls. These results revealed previously unexplored distinct temporal-spatial lipid dynamics in the restenotic arterial wall. Additionally, we employed time-of-flight secondary ion mass spectrometry (TOF-SIMS) tandem MS imaging for both molecular identification and imaging at high spatial resolution. These imaging modalities provide powerful tools for unraveling novel mechanisms of restenosis involving lipids or small signaling molecules.

Introduction

Cardiovascular disease is the leading cause of death worldwide.[1] The most common treatment to resume circulation is balloon angioplasty, performed in over a million patients each year. Although acutely effective, this intervention can injure the artery leading to the development of neointimal lesion in the inner vessel wall that re-narrows the lumen - a recurrent disease known as restenosis. Even with stent implantation, rates of restenosis remain at ~12-30%.[2, 3] It is therefore imperative to better understand mechanisms of restenosis for achieving more effective treatments.

The artery is composed of three distinct layers: adventitia, media, and the inner lining of endothelium which would be removed due to angioplasty followed by the development of neointima lesion. Three major cell types, fibroblasts, smooth muscle cells (SMCs), and endothelial cells reside in these three layers, respectively. The networks formed by these cells and their signaling molecules are crucial to the pathogenesis and progression of restenosis. While a wealth of knowledge is available on many pathways involved in restenosis,[4-6] spatial-temporal dynamics of signaling molecules and their networks remain poorly understood due to technological limitations. Particularly challenging is the spatial profiling of small molecules including bioactive lipids in the vessel wall of small animal models, the thickness of which is typically in the sub-millimeter range.

Bioactive lipids are important signaling molecules involved in a variety of biological processes.[7,8] A number of studies have implicated abnormal synthesis and distribution of these lipids associated with a wide range of diseases, particularly in the cardiovascular system.[9,10] However, the majority of these studies have relied on lipid detection in whole tissue homogenates or body fluid. Critical information concerning lipid spatial dynamics is missing. Aside from lability, chemical diversity and complexity, and generally low abundance of bioactive lipids in the

arterial wall, the small arteries in experimental animals (e.g., ~1 mm diameter in rats) impose an additional technical challenge. The emerging technology of mass spectrometric imaging (MSI) brings about a unique opportunity to overcome these long-standing hurdles in spatially profiling lipids in the arterial wall. MSI spatial profiling of lipids has been reported, mostly in lipid-rich organs such as brain, kidney, lung and some types of tumors,[11-17] and also in arteries of human patients, but not in arteries from small animals.

Here, we describe the first MSI-based in situ lipidomic imaging in a rat model mimicking angioplasty-elicited clinical restenosis. This model features well-characterized stages of disease progression and corresponding dynamic signaling events with a variety of bioactive lipids involved. Our data reveals distinct temporal-spatial profiles of different bioactive lipid species during the development of restenosis, including those not previously linked to restenosis.

Material and methods

Chemical reagents

Methanol and formic acid (FA) were purchased from Fisher Scientific (Pittsburgh, PA). Difco™ gelatin was purchased from Becton, Dickinson and Company (Sparks, MD). 2, 5-dihydroxybenzoic acid (DHB) was from Acros Organics (Morris Plains, NJ). Microscope glass slides were from VWR international, LLC (Radnor, PA). Reagents used for Verhoeff-Van Gieson staining were purchased from Sigma Aldrich (cat#: 1159740002). Distilled water was Milli-Q water from a Millipore filtration system (Bedford, MA). All reagents were used without additional purification.

Animal experiments

Animal experiments were conducted following institutional guidelines (University of Wisconsin-Madison IACUC). Balloon angioplasty for carotid arteries was performed in male

Sprague–Dawley rats (350-400g) as previously described.[18] Briefly, rats were anesthetized with isoflurane followed by a longitudinal incision in the neck to expose carotid arteries. A 2-F balloon catheter (Edwards Lifesciences, Irvine, CA) was inserted through an arteriotomy on the left external carotid artery to the common carotid artery. To produce arterial injury, the balloon was inflated at a pressure of 2 atm and withdrawn from proximal common carotid artery to the carotid bifurcation. This action was repeated three times. The external carotid artery was then permanently ligated, and blood flow was resumed. A total of 9 rats were used for these experiments. Three injured and three control arteries were collected for each condition: Day 3, Day 7, and Day 14 post balloon angioplasty.

Tissue collection and sectioning

For each rat, the balloon angioplasty-treated left common carotid artery was collected while the contralateral common carotid artery without angioplasty was harvested as control. Collected carotid arteries were embedded in gelatin (100 mg/mL in warm MilliQ water) and snap-frozen on dry ice immediately after dissection. The frozen tissues were stored at -80 ˚C until cryosectioning into 12 µm thick slices on a cryostat (Thermo Scientific Microm HM 525，Waltham, MA). A total of 18 tissue slices (3 from each condition) were mounted onto two microscope glass slides (75x25x1 mm). The glass slides were then dried in a desiccator for 30 minutes at room temperature to avoid condensation of water vapor.

Matrix application

A robotic sprayer system (TM sprayer, HTX Technologies, Carrboro, NC) was used to homogeneously apply matrix on two tissue slides: 24 passes of DHB matrix (40 mg/mL in 0.1% FA / methanol/water (50:50, v:v)) was applied with a flow rate of 0.05mL/min, nozzle temperature

of 80˚C, and a moving velocity of 800 mm/min. The slides were dried in a desiccator at room temperature following matrix application and stored at -80˚C until analysis.

MALDI MS imaging

All MALDI-MSI experiments were performed on a MALDI-LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, Bremen, Germany) equipped with a 60Hz 337 nm N2 laser. Full MS experiment was performed in positive-ion mode with m/z range of 450-1000, mass resolution of 30,000 (at m/z 400), using a laser energy of 20 µJ and 2 microscans/step. A spatial resolution of 75 µm was used for all acquisitions. The precursor ions of target lipids in the full MS scan were isolated and subsequently fragmented by higher-energy collisional dissociation (HCD) with the normalized collisional energy of 25. Instrumental methods were set up in Xcalibur software (Thermo Scientific, Bremen, Germany) and the imaging position file was defined with TunePlus software (Thermo Scientific, Bremen, Germany).

Histology

After MALDI-MSI acquisition, artery sections mounted onto glass slides were fixed in 4% paraformaldehyde solution at 4˚C overnight. Slides were then air-dried in a chemical hood and rinsed in a series of ethanol solutions (100%, 95%, and 75%). Gelatin was removed by incubating the slides at 37 ˚C for 45 min. Verhoeff-Van Gieson (VVG) staining was performed as we previously described.[19] Briefly, processed sections were stained in a working solution containing alcoholic hematoxylin, ferric chloride, and Verhoeff's iodine solution for 15 min, followed by rinsing in running water for 5 min. Slides were differentiated in fresh ferric chloride solution before mounted with a coverslip. Images were acquired on a Nikon Eclipse TE2000-inverted microscope (Minato, Tokyo, Japan) at 4X magnification.

Data processing and statistical analysis

MS data was processed using Xcalibur (Thermo Scientific, Bremen, Germany) and ImageQuest (Thermo Scientific, Bremen, Germany). MS images were analyzed in MSiReader (North Carolina State University, Raleigh, NC) following exporting data into imzML in ImageQuest. Images were normalized to total ion current (TIC) and the mass tolerance window for image construction is 5 ppm. MS data as an imzML format was then loaded into SCiLS software 2018 (Bruker Daltonics, Bremen, Germany) for statistical analysis. Using SCiLS, discriminative analysis was used to calculate Receiver Operator Characteristic (ROC) curves and Area Under the Curve (AUC) values. Each AUC value represents the ability of a single m/z value to act as binary classifier for the injured versus control/sham arteries. Additionally, hypothesis testing using t-tests was used to calculate the p-value for the differences in the mean spectra between the control and injured arteries. SCiLS software was also used for spatial analysis by co-registration of the VVG stained histology to the TIC of the artery, annotating regions of interest based on the histology, and then performing hypothesis testing on the 3 layers of the injured arteries: the adventitia, media, and neointima based on the histological staining overlay. METLIN metabolite database (Scripps Center for Metabolomics, La Jolla, CA) was used for accurate mass matching and lipid assignment with 5 ppm tolerance.

TOF-SIMS tandem MS imaging for MS/MS identification of lipids

The analyses were performed using a PHI nanoTOF II TOF-SIMS Parallel Imaging MS/MS instrument (Physical Electronics, Minnesota, USA). A detailed description of this TOF-TOF instrument has been reported previously.[20-22] In the present study, all the spectra and images were recorded with electrodynamically bunched pulses of a 30 keV Bi3+ primary ion beam of which the DC current was $\approx$ 16 nA. The ion beam was operated in the high resolution-squared

(HR^2) mode to achieve ⩽1 μm lateral resolution at high mass resolving power. The field-of-view of each analytical area was 500 μm × 500 μm divided by 256 x 256 image pixels; to image the entire tissue cross section a stage raster motion was used with software stitching of 4 x 4 imaging tiles to view a total image area of 2.0 mm x 2.0 mm encompassing the tissue section. TOF-SIMS data were acquired over a m/z range of 0-2,000 in positive ion mode, and the ion fluence for each acquisition was at least an order of magnitude below the static limit of analysis. For the TOF-SIMS analysis, the m/z 0-2,000 range was used because that was the duty cycle (approx. 121 μsec) required for stable charge neutralization and which also encompasses the mass range of the targeted lipids. During the analyses, low energy electrons (15 eV) and low energy Ar^+ ions (10 eV) were applied for charge compensation. The lateral resolution (Δl) was measured using 80/20 bounds to be ≈ 1 μm; however, the pixel size (500 μm/256 pixels) was 1.95 μm. Data acquisition was achieved using PHI SmartSoft-TOF software and data processing was performed using PHI TOF-DR (Physical Electronics, MN) software.

Results and Discussion

MALDI imaging MS was applied to investigate the spatiotemporal changes of lipids during the progress of restenosis (**Figure 1A**). We used the balloon angioplasty technique to model restenosis. Typically, a balloon catheter is inserted into the rat carotid artery and then inflated and pulled. This action inflicts injury by overstretching the artery wall while damaging the endothelium. Disruption of this endothelial barrier between blood and the vessel wall exposes SMCs in the medial layer to various stimulants in circulation such as cytokines and growth factors that trigger a myriad of signaling events. Consequently, SMCs become migratory and proliferative forming a new layer of tissue called neointima, which occupies and narrows the lumen space (so called restenosis). To illustrate the distinct layers of the arterial wall structure and neointima

formation, diagrams in **Figure 1B** show the different layers of the artery present in the 3 different injured conditions and in the uninjured condition. In contrast to the uninjured control and injured day-3 arteries, the injured artery at day-7 and day-14 post-balloon angioplasty has an additional thick neointima layer.

This widely-used balloon angioplasty model recapitulates distinct stages of post-angioplasty development of restenosis and associated dynamic signaling events. Samples including injured and uninjured (sham control) carotid arteries were collected at three post-angioplasty time points that are well-defined in the literature. [23] The day-3 time point features acute activation of proliferation, as well as apoptosis of SMCs in the medial layer, with neointima not yet formed. The day-7 time point represents a highly active signaling stage marked by diminished cell death and a dramatic increase of proliferative and migratory SMCs that form a prominent neointima layer. At the day-14 time point, cellular activities decline, while the flow-obstructing neointimal lesion is fully developed. Here, we will discuss the trends of three main lipid groups that previously little was known about in restenosis: diacyglycerols (DAGs), lysophosphylcholines (LPCs), and sphingomyelins (SMs) (**Table 1**).

Our study is the first to directly identify the dynamic changes of DAG, LPC, and SM species during the pathogenesis of restenosis. The previous knowledge on their contribution to cardiovascular diseases was largely limited to either in vitro cell culture studies or measurements in atherosclerotic plaques.[24, 25] For example: (1) altered LPC production led to dysfunction of cultured SMCs[26]; (2) deletion of a phospholipase (iPLA2β, responsible for synthesis of arachidonic acid and LPC) ameliorated restenosis in a transgenic mouse model[27]; (3) while DAG is a known activator of certain Protein Kinase C (PKC) family members, aberrant Protein Kinase C activation was closely linked to SMC pathology and restenosis[28]; (4) SM is correlated with SMC

apoptosis, and its level was elevated in whole-tissue homogenates of atherosclerotic plaques.[29] Taken together, these studies indirectly linked the three lipid groups to restenosis.

MALDI-MSI indicates up-regulation of diacylglycerols in the injured arterial wall

We first analyzed the temporal dynamics of DAGs for a possible relevance to restenosis. To sort out the lipids that showed prominent temporal dynamics with a close relevance to restenosis, we analyzed diacylglycerols (DAGs), which in the literature have been linked to pathogenic pathways in restenosis.[30] The data shows that at post-angioplasty day 3, there was a statistically significant increase of DAGs in the injured artery compared to the uninjured control. The signal was localized in the middle of the artery section, correlating with the medial layer; the neointima was not yet formed at this time point. At day 7, coincident with fast growth of neointima in the injured artery, all measured DAG lipid species showed a statistically significant increase with p-values < 0.0001. Additionally, most DAG lipids had AUC values > 0.75 at day 7, indicating that the changes were great enough for these DAG species to discriminate between injured and uninjured tissues. DAG signal in the injured artery was substantially attenuated at day 14, while remaining upregulated compared to the day-14 uninjured control. **Figure 2B** shows the relative intensity of DAG (36:1), m/z 627.5381, a representative DAG species illustrating the dynamic trends of this lipid class. Briefly, each figure shows injured carotid arteries compared side-by-side with their respective contralateral uninjured arteries from same animal (n=3, data shown in **Supplemental Figure S1**). At day 3, day 7, and day 14, m/z 627.5381 is significantly upregulated compared to uninjured control, with the increase peaking at day 7. Statistics in **Supplemental Figure S1** confirm that this trend is conserved for most DAG lipid species identified. To reveal the temporal changes of different DAG lipid classes, the day-3, day-7, and day-14 injured tissues were statistically compared. DAG (36:2), m/z 625.5219 is shown in **Figure 3A** as an example of

this lipid class. From Day 3 to Day 7, DAG (36:2) signal increased, and then declined at day 14 but remained above the day-3 level. Statistics is provided for all DAG lipid species (**Table S1**), demonstrating a temporal regulation in injured arteries.

DAGs are one of the most widely studied secondary signaling messengers. They are indispensable for maintaining the normal physiology of various organs and tissues, mainly through activating multiple classes of protein kinase C (PKC).[30, 31] Their aberrant abundance and/or distribution are associated with pathogenic conditions.[32] The activation of various isoforms of PKC, in particular, PKC delta and PKC beta, has been reported to contribute to the development of restenosis.[33-35] However, how DAGs change during this process has not been clearly revealed.[36,37] Herein, for the first time, our MALDI-MSI study characterized restenosis-associated temporal-spatial DAG profiles that closely correlate to previously reported dysregulated PKC activities. As such, these data also serve as a positive control validating our MALDI-MSI methodology for profiling the changes of other lipids correlated with various stages of restenosis.

MALDI-MSI indicates up-regulation of lysophosphatidycholines in the injured arterial wall

We next analyzed other prominently regulated lipids, including lysoPCs. Multiple lysoPCs were found to be greatly accumulated following arterial injury, predominantly in the neointima/medial area where smooth muscle cells (SMCs) undergo a phenotypic switching that is primarily responsible for the generation of the neointimal lesion. **Figure 2C** shows the relative intensity of LysoPC (18:1), m/z 522.3578, as an example of this lipid class. At day 3, day 7, and day 14, m/z 522.3578 was significantly upregulated compared to the control, with the greatest increase in intensity occurring at day 7. Statistics in **Supplement Figure S2** confirm that this trend is conserved for other LysoPC lipids. Therefore, the general pattern of changes in LysoPCs is very similar to that of DAGs. Compared to uninjured control, most of the LysoPC species significantly

increased in injured arteries at all 3 time points. The ROC curves indicate that the highest AUC values are present at day 7, likely due to the dramatic injury-induced LysoPC lipid upregulation in the neointimal lesions at this time point. To understand the temporal dynamic regulation of the LysoPC lipid classes, day-3, day-7, and day-14 injured tissues were statistically compared. As an example, the LysoPC (16:1), m/z 494.325 signal (**Figure 3B**) increased from day 3 to day 7, and then slightly decreased at day 14 compared to day 7, likely due to a resolution of the injury-related processes. To further confirm the identification of lipids, representative on-tissue fragmentation of a LysoPC lipid ion is provided in the supplemental section (**Figure S5**), demonstrating the presence of m/z 184.0727 as the diagnostic ion of the phospholipid head group of LPC.

LysoPCs are a group of highly active phospholipids that are involved in a wide range of physiological and pathological events.[38] Recently, the functions of LysoPCs in cardiovascular disease have been extensively studied,[39,40] and a positive correlation between increased circulating LysoPCs and risks of coronary diseases has been revealed.[41,42] However, while in vitro studies showed that SMC pro-restenotic phenotypes were stimulated by LysoPCs, the responses of different vascular cell types were highly variable,[43,44] and the restenosis-related dynamics of LysoPCs in the arterial wall has not been reported. Our study provides the first temporal dynamic profiling of LysoPCs closely correlated to the disease context of restenosis, thus contributing novel information for further mechanistic studies.

MALDI MSI indicates down-regulation of sphingomyelins in the injured arterial wall

Aside from upregulated lipids, MALDI-MSI also identified several down-regulated lipids, such as sphingomyelins (SMs) during the development of restenosis. **Figure 2D** shows the relative intensity boxplots of SM (38:1), m/z 781.6229, as an example, to demonstrate the trends of intensity changes of SM species. Compared to uninjured control, SM (38:1) is significantly down-

regulated in injured arteries at day 3, day 7, and day 14. Statistics in **Supplemental Figure S3** confirm that this trend is conserved for most of the SM species identified. As we have seen with the DAGs and LysoPCs, SMs also exhibit the most dramatic changes at day 7 post injury, as indicated by the highest AUC values at day 7. To illustrate the temporal changes of SM lipids, we use SM (40:1), m/z 787.6760, as an example in **Figure 3C**. SM (40:1) is significantly upregulated at day 14, compared with day 3 and day 7. Additional dynamic statistics are available for all SM species in **Table S1**, which confirm that day 14 is significantly upregulated compared to day 3 and day 7. This dynamic profile of SM is distinct from those of DAGs and LysoPCs, further confirming the specificity and effectiveness of our MALDI-MSI methodology.

As the most abundant sphingolipids are stored in plasma membranes, sphingomyelins can be converted into ceramides, which ultimately lead to the synthesis of sphingosine and sphingosine-1-phosphate (S1P). S1P is a well-established mitogenic stimulus implicated tumorigenesis and metastasis. Based on studies with genetic deletion of S1P receptors, [45,46] it has been postulated that S1P could accumulate in injured arteries and contribute to the signaling events that lead to restenosis. Consistently, the observed post-injury decrease of SMs, the primary substrates for the synthesis of S1P, might be accounted for an increase of S1P. While our study provides an interesting correlation with literature evidence, further studies are needed to define the functional roles of SMs in restenosis.

Histology reveals distinct stages of restenotic development in injured rat carotid arteries

Following MALDI MSI analysis, artery cross-sections were subjected to VVG staining (**Figure 2A**). VVG stains collagen as red and elastin as black, which can be used to distinguish the adventitial layer (shown in red collagenous regions) and medial layer (shown in black elastin region). While neointima was not seen on day-3 injured artery sections, it developed into a layer

of significant thickness at day 7 and progressed to the maximum thickness at day 14 in injured tissues (denoted with the presence of a blue arrow in **Figure 2A**). This neointimal growth at day 7 and day 14 exhibited in the histology is consistent with previous reports,[19, 47] thus validating the accuracy of the angioplasty model used. Furthermore, since the same artery section was used first for MALDI-MSI and then VVG staining, results from these two can be readily co-registered and correlated (**Figure 4**) with the stained section, providing spatial information and relative lipid abundance within the distinct layers of the artery.

MALDI-MSI is combined with histology to reveal neointima specific dynamic lipid changes

As indicated by histology, the neointima was not seen on injured day 3 tissues but developed into a visibly thick inner layer of the artery in injured tissues collected at day 7 and day 14. We thus compared lipid signal intensities specifically in the neointima layer to provide new information in more detail. Statistical analysis of annotated neointima lesions in injured arteries of day 7 and day 14 indicates that the abundance of DAGs and LysoPCs was higher at day 7 compared with day 14 (**Figure 5**). As shown in **Figure 5A-B**, example DAG (34:2), m/z 597.4889, and example LysoPC (18:2), m/z 502.3420 both exhibit the same trend of decrease from day 7 to day 14. In contrast, example SM (40:1), m/z 809.6544, shown in **Figure 5C**, is higher at day 14 compared with day 7. Further statistical analysis of the intensity of each layer of the carotid artery (the neointima, the media, and the adventita) are provided in **Supplemental Table S3** for injured day 7 arteries and **Supplemental Table S4** for injured day 14 arteries. Statistics provided in **Supplemental Table S3** confirm that SMs were overall more abundant in the day-14 neointima compared with the day-7 neointima. Additional spatial statistics comparing the adventitia, media, and neointima regions are provided in the supplemental sections for day-7 and day-14 injured arteries (**Table S3** and **Table S4**, respectively). The data indicates that the neointimal levels of

DAGs (day 7), LysoPCs (day 7 and day 14), and SMs (day 14) are statistically different from both the adventitia and the media. It should be noted that there is technical limitation in spatially resolving all three layers: the adventitia, media and neointima, given the small size of a rat artery (~1 mm) and the minimum spatial resolution of 75μm on the MALDI-Orbitrap LTQ XL instrument without oversampling from neighboring pixels. Moreover, many of the lipids investigated here are of relatively low abundance in the tissue, potentially compromising detectability from increasing spatial resolution with smaller raster steps. Further MSI analysis with improved spatial resolution using TOF-SIMS was necessary to fully resolve all three layers with minimal overlapping between neighboring pixels that are shared by the outlined regions of interest. Nevertheless, we are able to use MALDI-MSI to visualize statistically significant temporal level changes of DAGs, LysoPCs, and SMs in the neointima, from post-injury day 7 to day 14, time points when substantial neointima has developed.

TOF-SIMS tandem MS imaging for identification and high-resolution imaging

TOF-SIMS tandem MS imaging was employed to observe the neointima lesions developed in day 7 injured tissue sections and to obtain MS/MS fragmentation for some low abundance lipid species. In **Figure 6A** we have produced an ion image of a 2.0 mm x 2.0 mm large-area mosaic tile map encompassing an entire tissue section; **Figures 6(B-C)** represent successive enlargement of the indicated regions of 1.0 mm x 1.0 mm and 500 μm x 500 μm areas, respectively.

The set of images was collected using an ion fluence of < 1.50 x1011 Bi_3^+/cm^2 which is two orders of magnitude below the static limit, i.e., the threshold for molecular damage. In each ion image the lateral resolution was between 1.01 μm and 1.85 μm. At this high spatial resolution, we are able to clearly observe the collagen (adventitia), elastin (media) and neointima layers in the cross-section of an injured artery showing the distribution of $C3H8N^+$ ion. High resolution MS

imaging enabled us to fully resolve the different regions of the artery. Additionally, the identification of specific DAG and SM moieties was also obtained from the TOF-SIMS tandem MS imaging data (**Figures S6 and S7**) to confirm tentative lipid identifications.

Conclusions

In this study, we performed the first MALDI-based lipidomic imaging on cross-sections of artery from a small animal model of restenosis, which revealed distinct temporal-spatial dynamics of lipids that are functionally related to restenosis. This imaging methodology exhibited unique strength in mapping dynamic lipid distribution in different layers of the arterial wall. Moreover, since unfixed cryo-sections were used, in situ display of various lipids could be captured. We were subsequently able to visualize at high resolving power the neointima lesions of the injured artery sections with the complementary MSI technique of TOF-SIMS, and to confirm the identification of DAGs and of SMs, making use of the tandem MS molecular identification. The combined information of the temporal and spatial dynamics of the lipidome is important for discovering potential novel mechanisms underlying restenosis. With the resolution and detection sensitivity continuously improving, the MALDI-MSI and TOF-SIMS tandem MS imaging technologies are becoming powerful tools in vascular studies that will ultimately contribute to the development of next-generation therapeutic methods for vascular diseases and beyond.

Acknowledgments

The research was supported in part by NIH R01 DK071801 and R56MH110215 (to LL), R01 HL 068673, R01 HL 129785, and R01 HL143469 (to KCK) and R01 HL 133665 (to L-WG). The MALDI Orbitrap instrument was purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531). LL acknowledges a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy. BW acknowledges AHA Pre-doctoral fellowship (Award #16PRE30160010).

References

1. Heron, M.; Anderson, R. N., Changes in the Leading Cause of Death: Recent Patterns in Heart Disease and Cancer Mortality. NCHS data brief 2016, (254), 1-8.

2. Cassese, S.; Byrne, R. A.; Tada, T.; Pinieck, S.; Joner, M.; Ibrahim, T.; King, L. A.; Fusaro, M.; Laugwitz, K. L.; Kastrati, A., Incidence and predictors of restenosis after coronary stenting in 10 004 patients with surveillance angiography. Heart 2014, 100 (2), 153-9.

3. Byrne, R. A.; Joner, M.; Kastrati, A., Stent thrombosis and restenosis: what have we learned and where are we going? The Andreas Gruntzig Lecture ESC 2014. Eur Heart J 2015, 36 (47), 3320-31.

4. Wei, G. L.; Krasinski, K.; Kearney, M.; Isner, J. M.; Walsh, K.; Andres, V., Temporally and spatially coordinated expression of cell cycle regulatory factors after angioplasty. Circ Res 1997, 80 (3), 418-26.

5. Li, Q.; Li, G.; Lan, X.; Zheng, M.; Chen, K. H.; Cao, C. M.; Xiao, R. P., Receptor interacting protein 3 suppresses vascular smooth muscle cell growth by inhibition of the phosphoinositide 3-kinase-Akt axis. J Biol Chem 2010, 285 (13), 9535-44.

6. Pyles, J. M.; March, K. L.; Franklin, M.; Mehdi, K.; Wilensky, R. L.; Adam, L. P., Activation of MAP kinase in vivo follows balloon overstretch injury of porcine coronary and carotid arteries. Circ Res 1997, 81 (6), 904-10.

7. Hannun, Y. A.; Obeid, L. M., Principles of bioactive lipid signalling: lessons from sphingolipids. Nat Rev Mol Cell Biol 2008, 9 (2), 139-50.

8. de Pablo, M. A.; Puertollano, M. A.; de Cienfuegos, G. Á., Biological and clinical significance of lipids as modulators of immune system functions. Clin Diagn Lab Immunol 2002, 9 (5), 945-950.

9. Wymann, M. P.; Schneiter, R., Lipid signalling in disease. Nat Rev Mol Cell Biol 2008, 9 (2), 162-76.

10. Hutchins, P. M.; Moore, E. E.; Murphy, R. C., Electrospray MS/MS reveals extensive and nonspecific oxidation of cholesterol esters in human peripheral vascular lesions. J Lipid Res 2011, 52 (11), 2070-83.

11. Roux, A.; Muller, L.; Jackson, S. N.; Post, J.; Baldwin, K.; Hoffer, B.; Balaban, C. D.; Barbacci, D.; Schultz, J. A.; Gouty, S.; Cox, B. M.; Woods, A. S., Mass spectrometry imaging of rat brain lipid profile changes over time following traumatic brain injury. J Neurosci Methods 2016, 272, 19-32.

12. Berry, K. A.; Li, B.; Reynolds, S. D.; Barkley, R. M.; Gijon, M. A.; Hankin, J. A.; Henson, P. M.; Murphy, R. C., MALDI imaging MS of phospholipids in the mouse lung. J Lipid Res 2011, 52 (8), 1551-60.

13. Muller, L.; Kailas, A.; Jackson, S. N.; Roux, A.; Barbacci, D. C.; Schultz, J. A.; Balaban, C. D.; Woods, A. S., Lipid imaging within the normal rat kidney using silver nanoparticles by matrix-assisted laser desorption/ionization mass spectrometry. Kidney Int 2015, 88 (1), 186-192.

14. Jiang, L.; Chughtai, K.; Purvine, S. O.; Bhujwalla, Z. M.; Raman, V.; Pasa-Tolic, L.; Heeren, R. M.; Glunde, K., MALDI-Mass Spectrometric Imaging Revealing Hypoxia-Driven Lipids and Proteins in a Breast Tumor Model. Anal Chem 2015, 87 (12), 5947-56.

15. Deininger, S. O.; Ebert, M. P.; Futterer, A.; Gerhard, M.; Rocken, C., MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. J Proteome Res 2008, 7 (12), 5230-6.

16. Zemski Berry, K. A.; Murphy, R. C.; Kosmider, B.; Mason, R. J., Lipidomic Characterization and Localization of Phospholipids in the Human Lung. J Lipid Res 2017, 58 (5), 926-933.

17. Thompson, C. G.; Bokhart, M. T.; Sykes, C.; Adamson, L.; Fedoriw, Y.; Luciw, P. A.; Muddiman, D. C.; Kashuba, A. D.; Rosen, E. P., Mass spectrometry imaging reveals heterogeneous efavirenz distribution within putative HIV reservoirs. Antimicrob Agents Chemother 2015, 59 (5), 2944-8.

18. Guo, L. W.; Wang, B.; Goel, S. A.; Little, C.; Takayama, T.; Shi, X. D.; Roenneburg, D.; DiRenzo, D.; Kent, K. C., Halofuginone stimulates adaptive remodeling and preserves re-endothelialization in balloon-injured rat carotid arteries. Circ Cardiovasc Interv 2014, 7 (4), 594-601.

19. Wang, B.; Zhang, M.; Takayama, T.; Shi, X.; Roenneburg, D. A.; Kent, K. C.; Guo, L. W., BET Bromodomain Blockade Mitigates Intimal Hyperplasia in Rat Carotid Arteries. EBioMedicine 2015, 2 (11), 1650-61.

20. Fisher, G. L.; Bruinen, A. L.; Ogrinc Potocnik, N.; Hammond, J. S.; Bryan, S. R.; Larson, P. E.; Heeren, R. M., A New Method and Mass Spectrometer Design for TOF-SIMS Parallel Imaging MS/MS. Anal Chem 2016, 88 (12), 6433-40.

21. Fisher, G. L.; Hammond, J. S.; Bryan, S. R.; Larson, P. E.; Heeren, R. M. A., The Composition of Poly(Ethylene Terephthalate) (PET) Surface Precipitates Determined at High Resolving Power by Tandem Mass Spectrometry Imaging. Microsc Microanal 2017, 23 (4), 843-848.

22. Chini, C. E.; Fisher, G. L.; Johnson, B.; Tamkun, M. M.; Kraft, M. L., Observation of endoplasmic reticulum tubules via TOF-SIMS tandem mass spectrometry imaging of transfected cells. Biointerphases 2018, 13 (3), 03B409.

23. Shi, X.; Guo, L. W.; Seedial, S.; Takayama, T.; Wang, B.; Zhang, M.; Franco, S. R.; Si, Y.; Chaudhary, M. A.; Liu, B.; Kent, K. C., Local CXCR4 Upregulation in the Injured Arterial Wall Contributes to Intimal Hyperplasia. Stem cells 2016, 34 (11), 2744-2757.

24. Martin-Lorenzo, M.; Balluff, B.; Maroto, A. S.; Carreira, R. J.; van Zeijl, R. J.; Gonzalez-Calero, L.; de la Cuesta, F.; Barderas, M. G.; Lopez-Almodovar, L. F.; Padial, L. R.; McDonnell, L. A.; Vivanco, F.; Alvarez-Llamas, G., Molecular anatomy of ascending aorta in atherosclerosis by MS Imaging: Specific lipid and protein patterns reflect pathology. J Proteomics 2015, 126, 245-51.

25. Martin-Lorenzo, M.; Balluff, B.; Maroto, A. S.; Carreira, R. J.; van Zeijl, R. J.; Gonzalez-Calero, L.; de la Cuesta, F.; Barderas, M. G.; Lopez-Almodovar, L. F.; Padial, L. R.; McDonnell, L. A.; Vivanco, F.; Alvarez-Llamas, G., Lipid and protein maps defining arterial layers in atherosclerotic aorta. Data Brief 2015, 4, 328-31.

26. Aiyar, N.; Disa, J.; Ao, Z.; Ju, H.; Nerurkar, S.; Willette, R. N.; Macphee, C. H.; Johns, D. G.; Douglas, S. A., Lysophosphatidylcholine induces inflammatory activation of human coronary artery smooth muscle cells. Mol Cell Biochem 2007, 295 (1-2), 113-20.

27. Liu, S.; Xie, Z.; Zhao, Q.; Pang, H.; Turk, J.; Calderon, L.; Su, W.; Zhao, G.; Xu, H.; Gong, M. C.; Guo, Z., Smooth muscle-specific expression of calcium-independent phospholipase A2β (iPLA2β) participates in the initiation and early progression of vascular inflammation and neointima formation. J Biol Chem 2012, 287 (29), 24739-53.

28. Ding, R. Q.; Tsao, J.; Chai, H.; Mochly-Rosen, D.; Zhou, W., Therapeutic potential for protein kinase C inhibitor in vascular restenosis. J Cardiovasc Pharmacol Ther 2011, 16 (2), 160-7.

29. Edsfeldt, A.; Dunér, P.; Ståhlman, M.; Mollet, I. G.; Asciutto, G.; Grufman, H.; Nitulescu, M.; Persson, A. F.; Fisher, R. M.; Melander, O.; Orho-Melander, M.; Borén, J.; Nilsson, J.; Gonçalves, I., Sphingolipids Contribute to Human Atherosclerotic Plaque Inflammation. Arterioscler Thromb Vasc Biol 2016, 36 (6), 1132-40.

30. Griendling, K. K.; Tsuda, T.; Alexander, R. W., Endothelin stimulates diacylglycerol accumulation and activates protein kinase C in cultured vascular smooth muscle cells. J Biol Chem 1989, 264 (14), 8237-40.

31. Yang, C.; Kazanietz, M. G., Divergence and complexities in DAG signaling: looking beyond PKC. Trends in Pharmacol Sci 2003, 24 (11), 602-8.

32. Eichmann, T. O.; Lass, A., DAG tales: the multiple faces of diacylglycerol--stereochemistry, metabolism, and signaling. Cell Mol Life Sci 2015, 72 (20), 3931-52.

33. Ananthanarayanan, B.; Stahelin, R. V.; Digman, M. A.; Cho, W., Activation mechanisms of conventional protein kinase C isoforms are determined by the ligand affinity and conformational flexibility of their C1 domains. J Biol Chem 2003, 278 (47), 46886-94.

34. Ha, H.; Yu, M. R.; Choi, Y. J.; Lee, H. B., Activation of protein kinase c-delta and c-epsilon by oxidative stress in early diabetic rat kidney. Am J Kidney Dis 2001, 38 (4 Suppl 1), S204-7.

35. Parmentier, J. H.; Zhang, C.; Estes, A.; Schaefer, S.; Malik, K. U., Essential role of PKC-zeta in normal and angiotensin II-accelerated neointimal growth after vascular injury. Am J Physiol Heart Circ Physiol 2006, 291 (4), H1602-13.

36. Yamanouchi, D.; Kato, K.; Ryer, E. J.; Zhang, F.; Liu, B., Protein kinase C delta mediates arterial injury responses through regulation of vascular smooth muscle cell apoptosis. Cardiovasc Res 2010, 85 (3), 434-43.

37. Andrassy, M.; Belov, D.; Harja, E.; Zou, Y. S.; Leitges, M.; Katus, H. A.; Nawroth, P. P.; Yan, S. D.; Schmidt, A. M.; Yan, S. F., Central role of PKCbeta in neointimal expansion triggered by acute arterial injury. Circ Res 2005, 96 (4), 476-83.

38. Bansal, P.; Gaur, S. N.; Arora, N., Lysophosphatidylcholine plays critical role in allergic airway disease manifestation. Sci Rep 2016, 6, 27430.

39. Aoyama, T.; Chen, M.; Fujiwara, H.; Masaki, T.; Sawamura, T., LOX-1 mediates lysophosphatidylcholine-induced oxidized LDL uptake in smooth muscle cells. FEBS Lett 2000, 467 (2-3), 217-20.

40. Rong, J. X.; Berman, J. W.; Taubman, M. B.; Fisher, E. A., Lysophosphatidylcholine stimulates monocyte chemoattractant protein-1 gene expression in rat aortic smooth muscle cells. Arterioscler Thromb Vasc Biol 2002, 22 (10), 1617-23.

41. Lavi, S.; McConnell, J. P.; Rihal, C. S.; Prasad, A.; Mathew, V.; Lerman, L. O.; Lerman, A., Local production of lipoprotein-associated phospholipase A2 and lysophosphatidylcholine in the coronary circulation: association with early coronary atherosclerosis and endothelial dysfunction in humans. Circulation 2007, 115 (21), 2715-21.

42. Iwase, M.; Sonoki, K.; Sasaki, N.; Ohdo, S.; Higuchi, S.; Hattori, H.; Iida, M., Lysophosphatidylcholine contents in plasma LDL in patients with type 2 diabetes mellitus: relation with lipoprotein-associated phospholipase A2 and effects of simvastatin treatment. Atherosclerosis 2008, 196 (2), 931-6.

43. Serrano-Pertierra, E.; Benavente, L.; Blanco-Gelaz, M. Á.; Fernández-Martín, J. L.; Lahoz, C. H.; Calleja, S.; López-Larrea, C.; Cernuda-Morollón, E., Lysophosphatidylcholine Induces Vascular Smooth Muscle Cell Membrane Vesiculation: Potential Role in Atherosclerosis through Caveolin-1 Regulation. J Proteomics Bioinform 2014, 7 (10), 332.

44. Colles, S. M.; Chisolm, G. M., Lysophosphatidylcholine-induced cellular injury in cultured fibroblasts involves oxidative events. J Lipid Res 2000, 41 (8), 1188-98.

45. Shimizu, T.; De Wispelaere, A.; Winkler, M.; D'Souza, T.; Caylor, J.; Chen, L.; Dastvan, F.; Deou, J.; Cho, A.; Larena-Avellaneda, A.; Reidy, M.; Daum, G., Sphingosine-1-phosphate

receptor 3 promotes neointimal hyperplasia in mouse iliac-femoral arteries. Arterioscler Thromb Vasc Biol 2012, 32 (4), 955-61.

46. Shimizu, T.; Nakazawa, T.; Cho, A.; Dastvan, F.; Shilling, D.; Daum, G.; Reidy, M. A., Sphingosine 1-phosphate receptor 2 negatively regulates neointimal formation in mouse arteries. Circ Res 2007, 101 (10), 995-1000.

47. Arai, H.; Gordon, D.; Nabel, E. G.; Nabel, G. J., Gene transfer of Fas ligand induces tumor regression in vivo. Proc Natl Acad Sci U S A 1997, 94 (25), 13862-7.

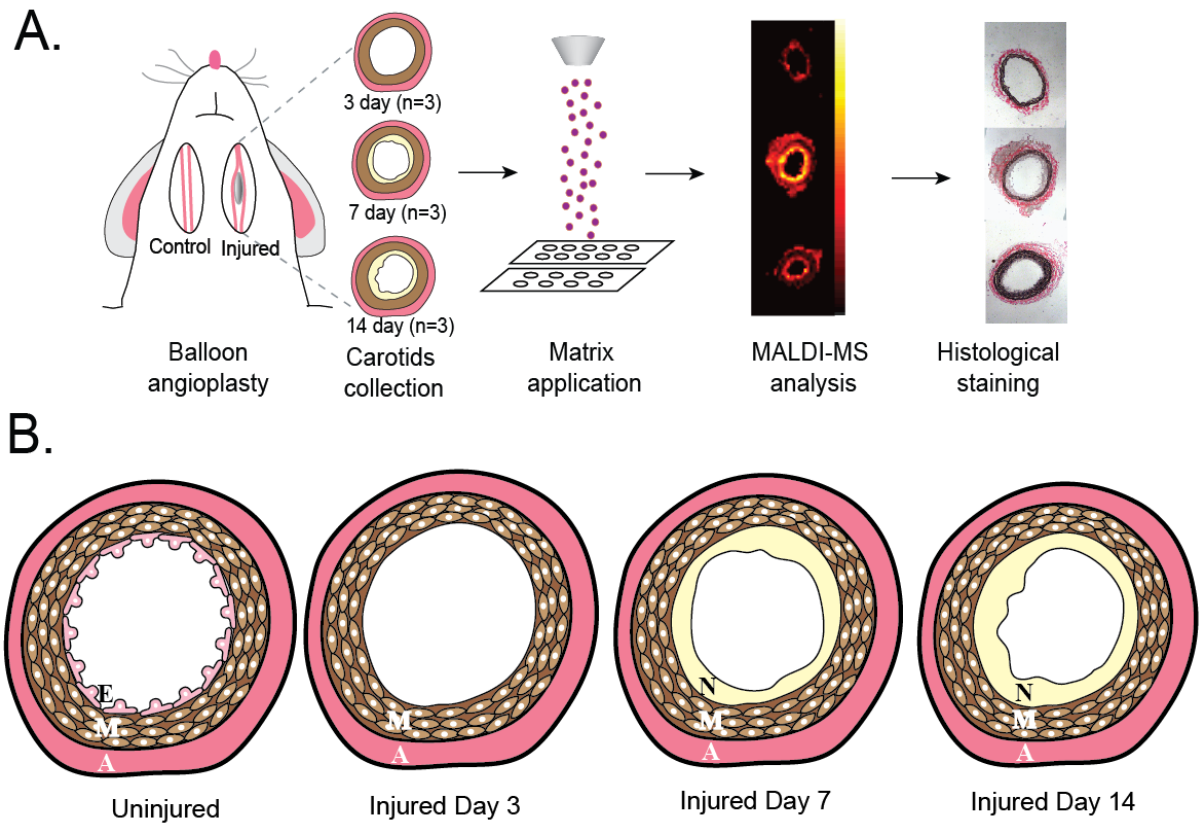


Figure 1. (A) Workflow for testing spatiotemporal changes of lipids in the development of neointima using a rat model of restenosis. (B) Diagram of the different layers of the artery in uninjured, injured day 3, injured day 7, and injured day 14. The layers are labeled as A: adventitia; M: media; E: epithelial layer; N: neointima layer.

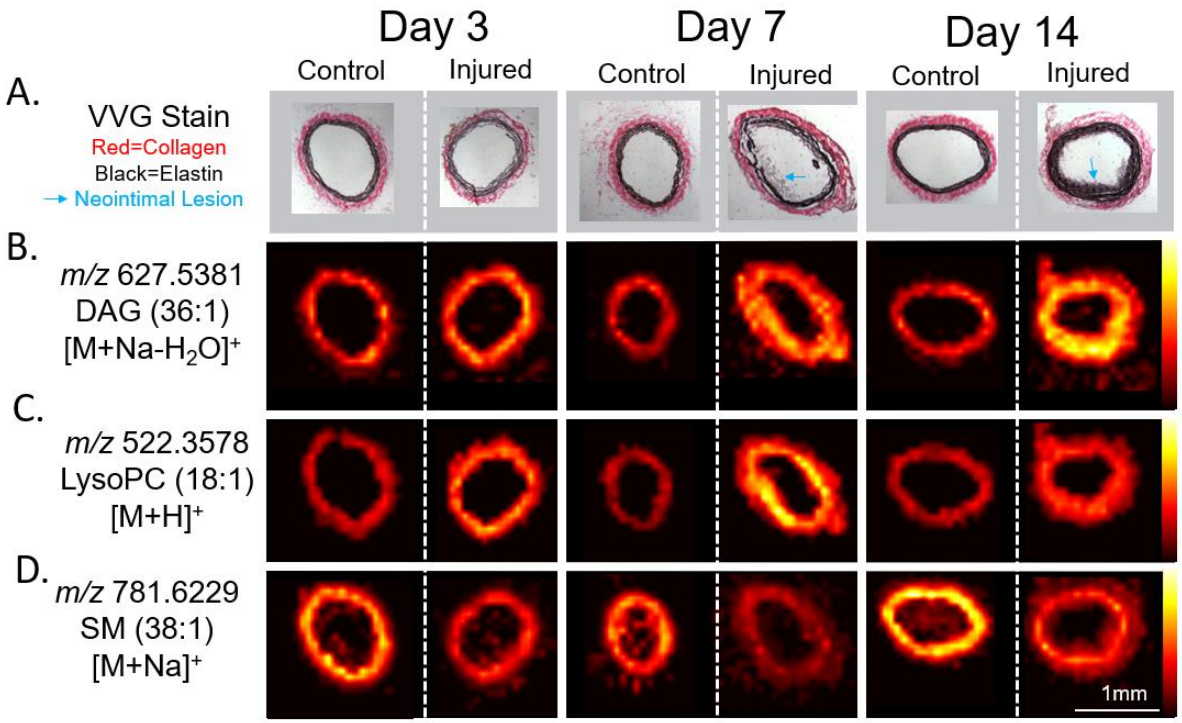


Figure 2. (A) VVG stains elastin rich regions black and collagen rich regions red in histological analysis to distinguish the media (middle) and adventitia (outer) regions. Neointima (inner) regions are shown with a blue arrow and are seen on day 7 and day 14 in injured arteries following balloon angioplasty. (B) DAG MALDI images of a representative DAG species, *m/z* 627.5381, demonstrate an increase in DAG (36:1) intensity on day 3, day 7 and day 14, compared with corresponding control arteries of one representative rat. (C) LysoPC MALDI images of a representative LysoPC, *m/z* 522.3578, demonstrate an increase in LysoPC (18:1) on day 3, day 7, and day 14, compared with corresponding control arteries. (D) MALDI images of a representative SM species, *m/z* 781.6229, demonstrate a decrease in SM (38:1) intensity on day 3, day 7 and day 14, compared with corresponding control arteries.

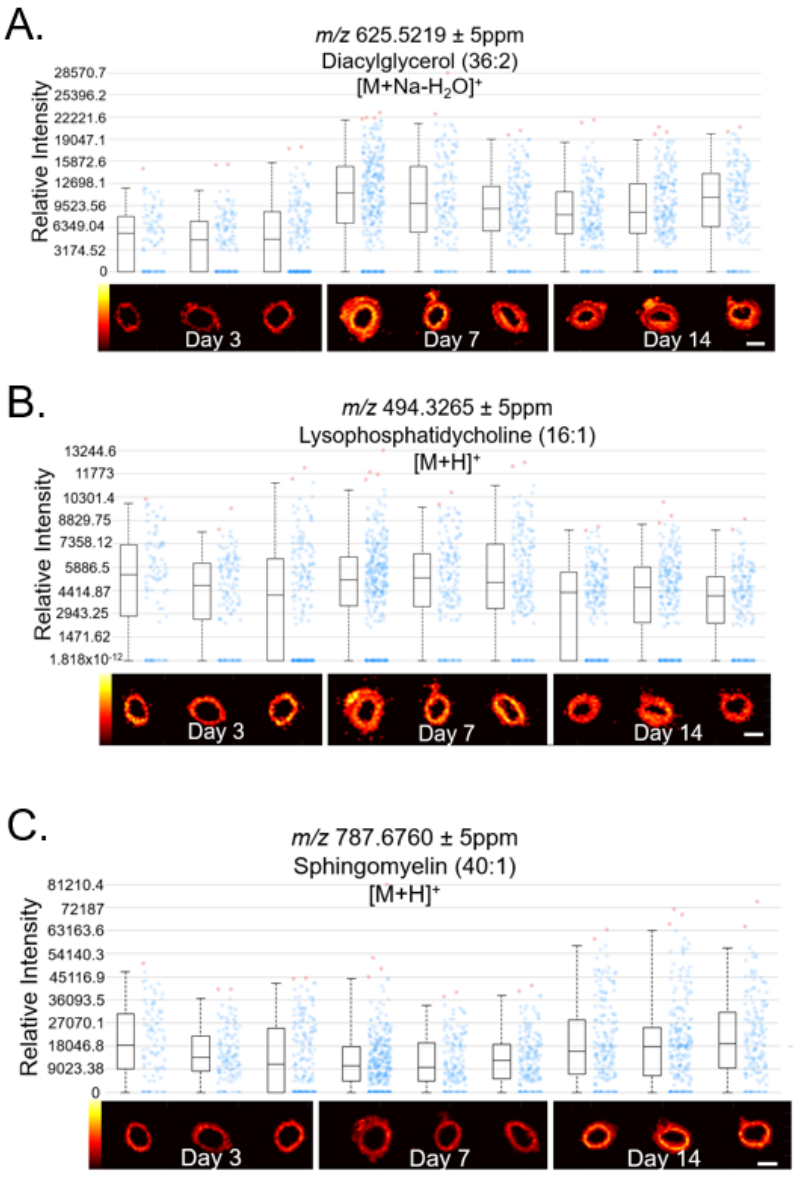


Figure 3. Temporal dynamics of three representative lipid species in injured arteries: Relative intensity box and whisker plots and MALDI images of (A) *m/z* 625.5219, DAG (36:2), in injured tissue (B) *m/z* 494.3265, LysoPC (16:1), in injured tissue (C) *m/z* 787.6760, SM(40:1), in injured tissue on day 3, day 7 and day14 for 3 biological replicates to illustrate post-injury temporal dynamics. Scale bar=1mm

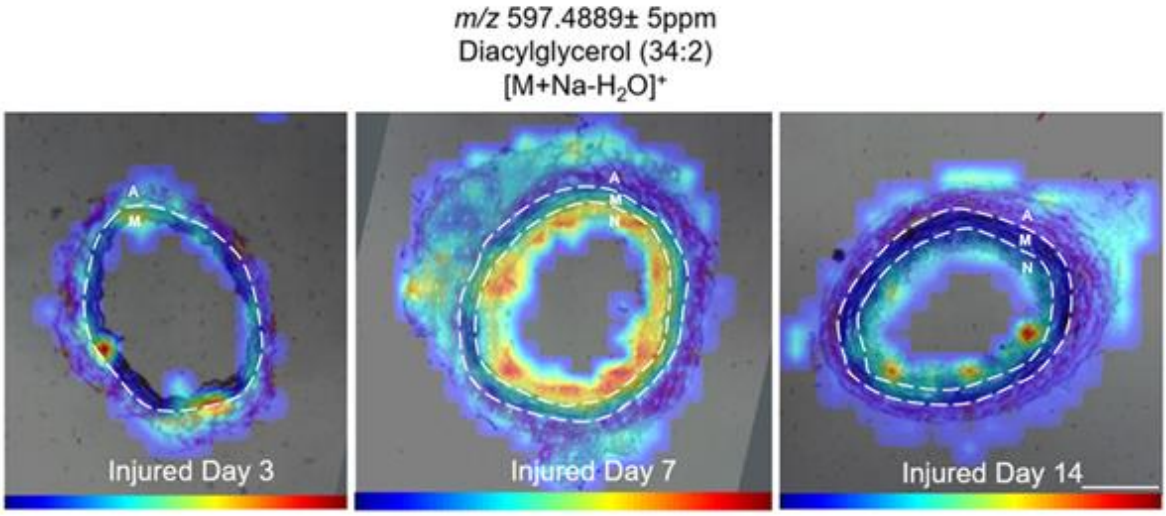


Figure 4. Histology overlay with mass spectrometry imaging reveals neointimal accumulation of diacylglycerol lipids: MALDI images and histology were co-registered for spatial analysis, as shown with *m/z* 597.489, DAG (34:2), overlaid with the VVG histology. This lipid is enriched in the neointima regions compared with the media and adventitia regions in day-7 and day-14 arteries. Scale bar=0.5mm

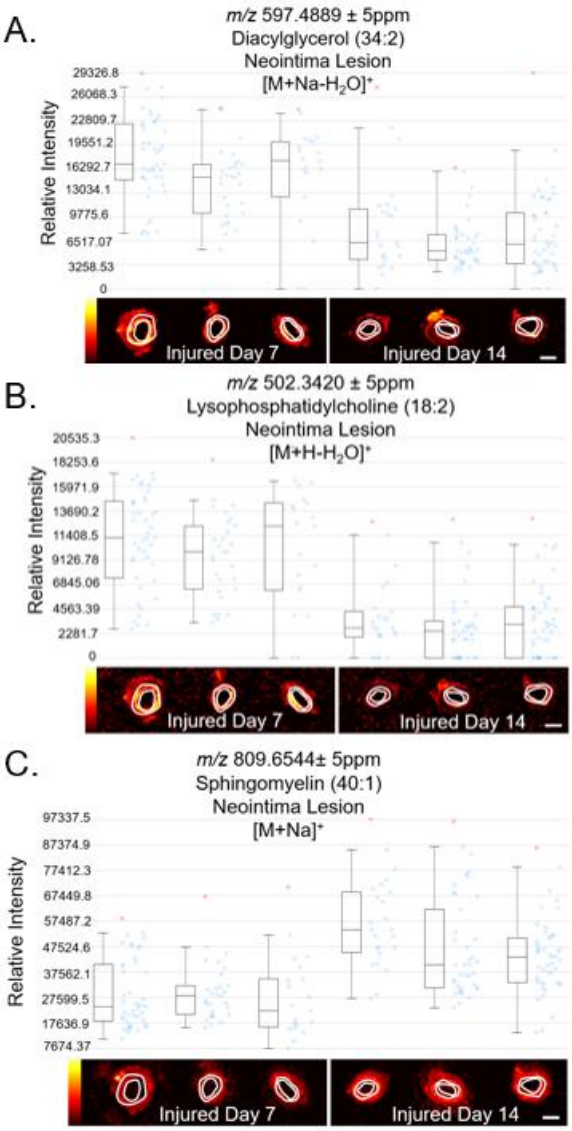


Figure 5. Comparison of neointimal abundance of lipids between post-injury day-7 and day-14 artery sections: Relative intensity box and whisker plots and MALDI images of (A) *m/z* 597.4889, DAG(34:2) (B) *m/z* 502.3420, LysoPC(18:2) and (C) *m/z* 809.6544, SM(40:1), shown in the day-7 and day-14 neointimal lesions defined through histology overlays. More statistical information about spatial distribution of each lipid species in the three layers of the artery can be found in the supplemental information (**Table S2** for injured day-7 tissue and **Table S3** for injured day-14 tissue). Scale bar=1mm

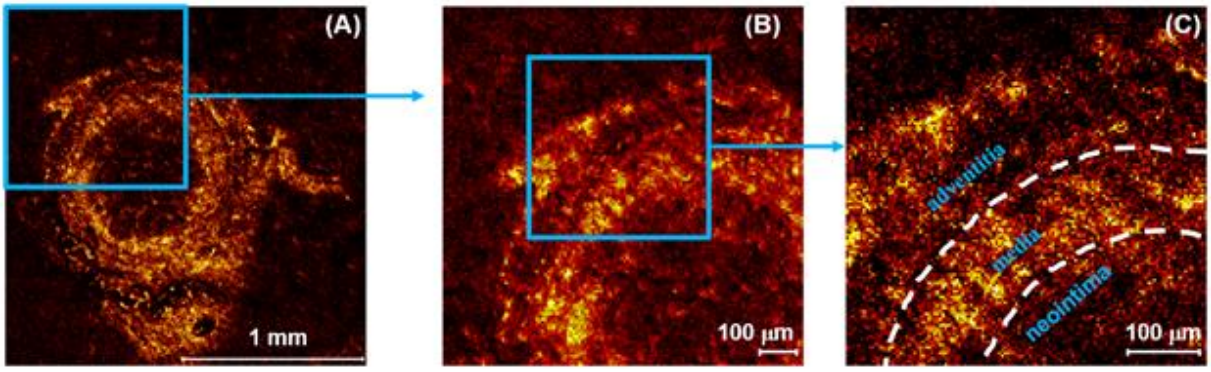


Figure 6. High Resolution MS Imaging: High resolution MS images showing the distribution of a non-specific ion of biological origin, C3H8N+ (choline head group) in the day-7 injured artery cross section. The image areas are (A) 2.0 mm x 2.0 mm, (B) 1.0 mm x 1.0 mm, and (C) 500 µm x 500 µm, wherein the layered structure including the neointima lesion is clearly observed.

Table 1. Identified lipids involved in the development of restenosis.

Table 1A. List of diacylglycerols (DAGs) involved in the development of restenosis

m/z	Ions	Tentative identification	Δppm
575.5052/597.4889	$[M+H-H_2O]^+$ / $[M+Na-H_2O]^+$	DAG (34:2)	2
577.5225/599.5063	$[M+H-H_2O]^+$ / $[M+Na-H_2O]^+$	DAG (34:1)	5
601.5208/623.5066	$[M+H-H_2O]^+$ /$[M+Na-H_2O]^+$	DAG (36:3)	2
603.5382/625.5219	$[M+H-H_2O]^+$ / $[M+Na-H_2O]^+$	DAG (36:2)	4
605.5538/627.5381	$[M+H-H_2O]^+$ / $[M+Na-H_2O]^+$	DAG (36:1)	4

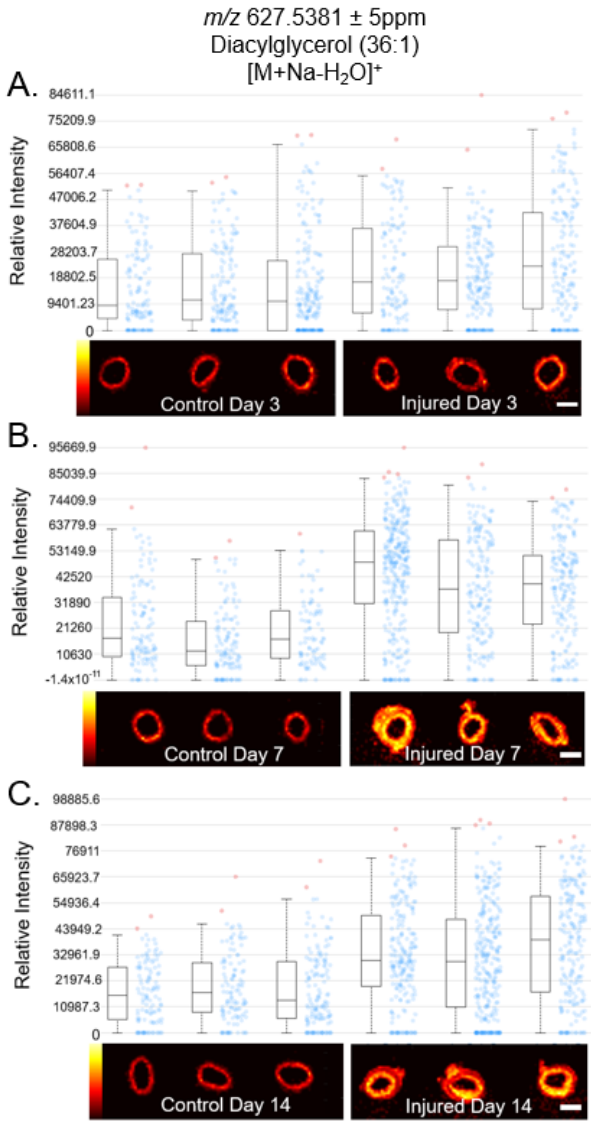
Table 1B. List of lysophosphatidycholines (LPCs) involved in the development of restenosis

m/z	Ions	Tentative identification	Δppm
476.3158/494.3265	$[M+H-H_2O]^+$ / $[M+H]^+$	LysoPC (16:1)	3
478.3292/496.3397	$[M+H-H_2O]^+$ / $[M+H]^+$	LysoPC (16:0)	1
502.3314/520.3420	$[M+H-H_2O]^+$ /$[M+H]^+$	LysoPC (18:2)	3
504.3471/522.3578	$[M+H-H_2O]^+$/$[M+H]^+$	LysoPC (18:1)	3

Table 1C. List of sphingomyelins (SMs) involved in the development of restenosis

m/z	Ions	Tentative identification	Δppm
731.6062/753.5921	$[M+H]^+$/ $[M+Na]^+$	SM (36:1)	0
759.641/781.6229	$[M+H]^+$/ $[M+Na]^+$	SM (38:1)	4
787.6706/809.6544	$[M+H]^+$/ $[M+Na]^+$	SM (40:1)	2
801.6844/823.671	$[M+H]^+$/ $[M+Na]^+$	SM (41:1)	0
815.7036/837.6862	$[M+H]^+$/ $[M+Na]^+$	SM (42:1)	4

Supplemental Information



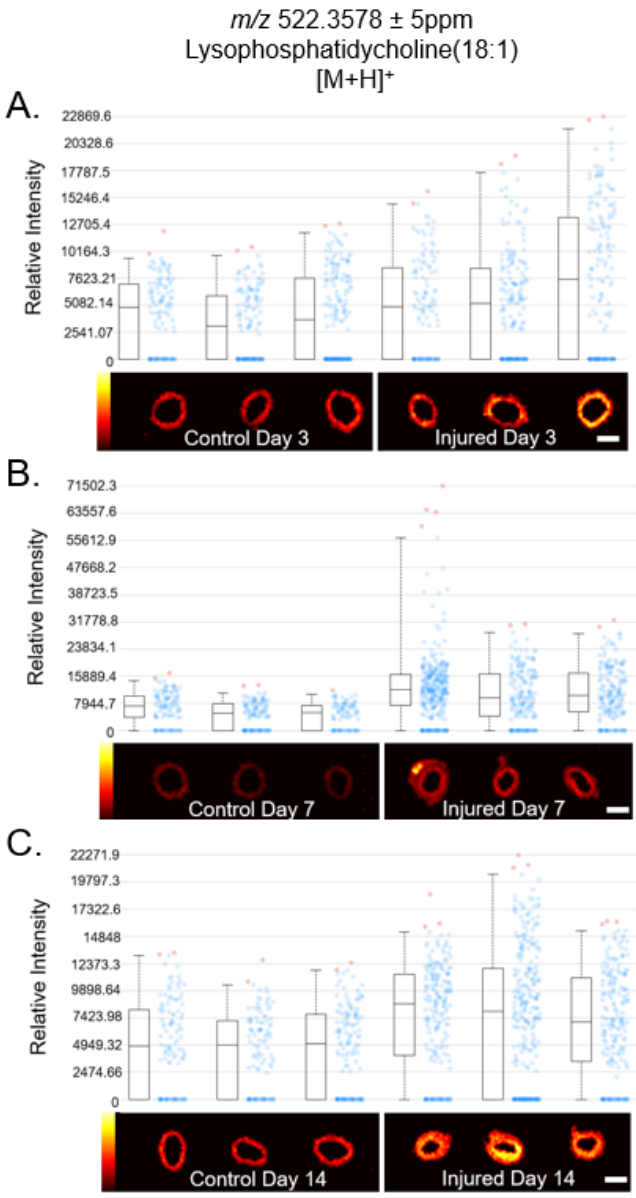
D.

m/z	Tentative Identification	Day 3 AUC	Day 7 AUC	Day 14 AUC
575.505	DAG (34:2)	0.619	0.8	0.734
597.489	DAG (34:2)	0.614	0.813	0.723
577.523	DAG (34:1)	0.602	0.791	0.725
599.506	DAG (34:1)	0.534	0.756	0.726
601.451	DAG (36:3)	0.581	0.797	0.755
623.507	DAG (36:3)	0.572	0.75	0.698
603.538	DAG(36:2)	0.677	0.806	0.769
625.522	DAG (36:2)	0.607	0.795	0.771
605.554	DAG(36:1)	0.633	0.703	0.662
627.538	DAG (36:1)	0.606	0.737	0.702

E.

m/z	Tentative Identification	Day 3 t-test	Day 7 t-test	Day 14 t-test
575.505	DAG (34:2)	>=0.05	<0.001	<0.001
597.489	DAG (34:2)	<0.001	<0.001	<0.001
577.523	DAG (34:1)	<0.01	<0.001	<0.001
599.506	DAG (34:1)	>=0.05	<0.001	<0.001
601.451	DAG (36:3)	>=0.05	<0.001	<0.001
623.507	DAG (36:3)	<0.001	<0.001	<0.001
603.538	DAG(36:2)	<0.001	<0.001	<0.001
625.522	DAG (36:2)	<0.001	<0.001	<0.001
605.554	DAG(36:1)	<0.001	<0.001	<0.001
627.538	DAG (36:1)	<0.001	<0.001	<0.001

Supplemental Figure S1. Trends in diacylglycerols in the development of restenosis.



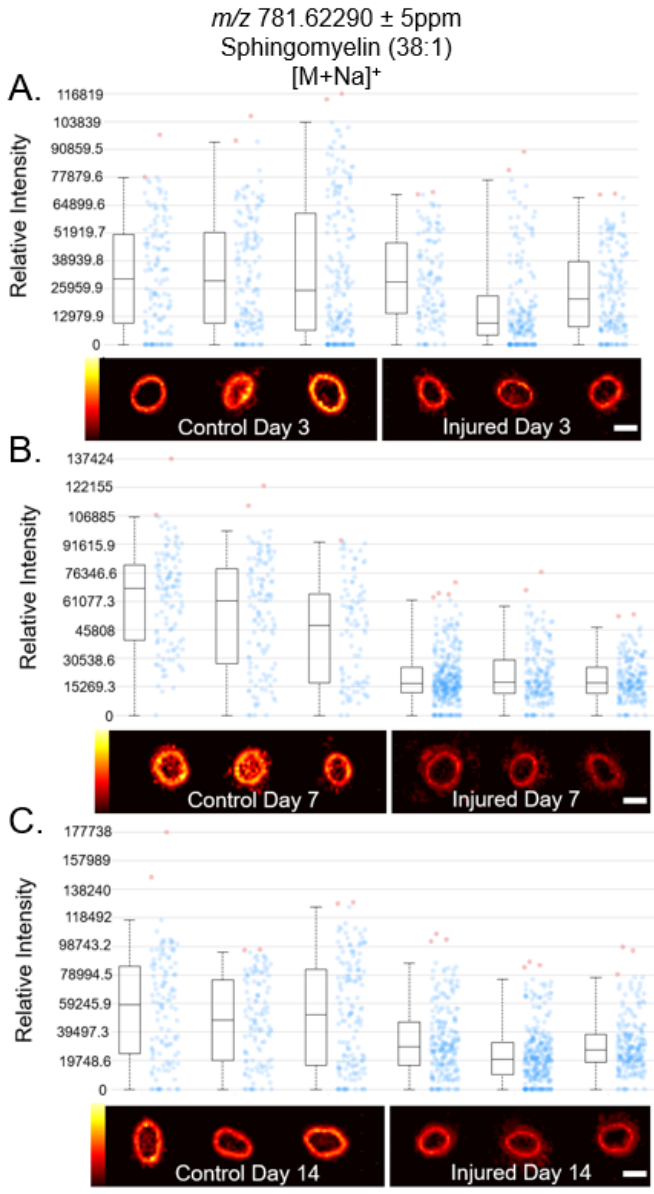
D.

m/z	Tentative Identification	Day 3 AUC	Day 7 AUC	Day 14 AUC
476.3158	LysoPC (16:1)	0.531	0.648	0.526
494.3265	LysoPC (16:1)	0.61	0.657	0.582
478.3158	LysoPC(16:0)	0.68	0.771	0.681
496.3397	LysoPC(16:0)	0.487	0.297	0.413
502.3314	LysoPC (18:2)	0.657	0.773	0.684
520.342	LysoPC(18:2)	0.554	0.499	0.549
504.3471	LysoPC(18:1)	0.646	0.787	0.758
522.3578	LysoPC(18:1)	0.602	0.706	0.652

E.

m/z	Tentative Identification	Day 3 t-test	Day 7 t-test	Day 14 t-test
476.3158	LysoPC (16:1)	<0.001	<0.001	<0.001
494.3265	LysoPC(16:1)	<0.001	<0.001	<0.001
478.3158	LysoPC(16:0)	<0.001	<0.001	<0.001
496.3397	LysoPC(16:0)	>=0.05	<0.001	<0.001
502.3314	LysoPC (18:2)	<0.001	<0.001	<0.001
520.342	LysoPC(18:2)	<0.001	>=0.05	<0.001
504.3471	LysoPC(18:1)	<0.001	<0.001	<0.001
522.3578	LysoPC(18:1)	<0.001	<0.001	<0.001

Supplemental Figure S2. Trends in lysophosphatidycholines in the development of restenosis.



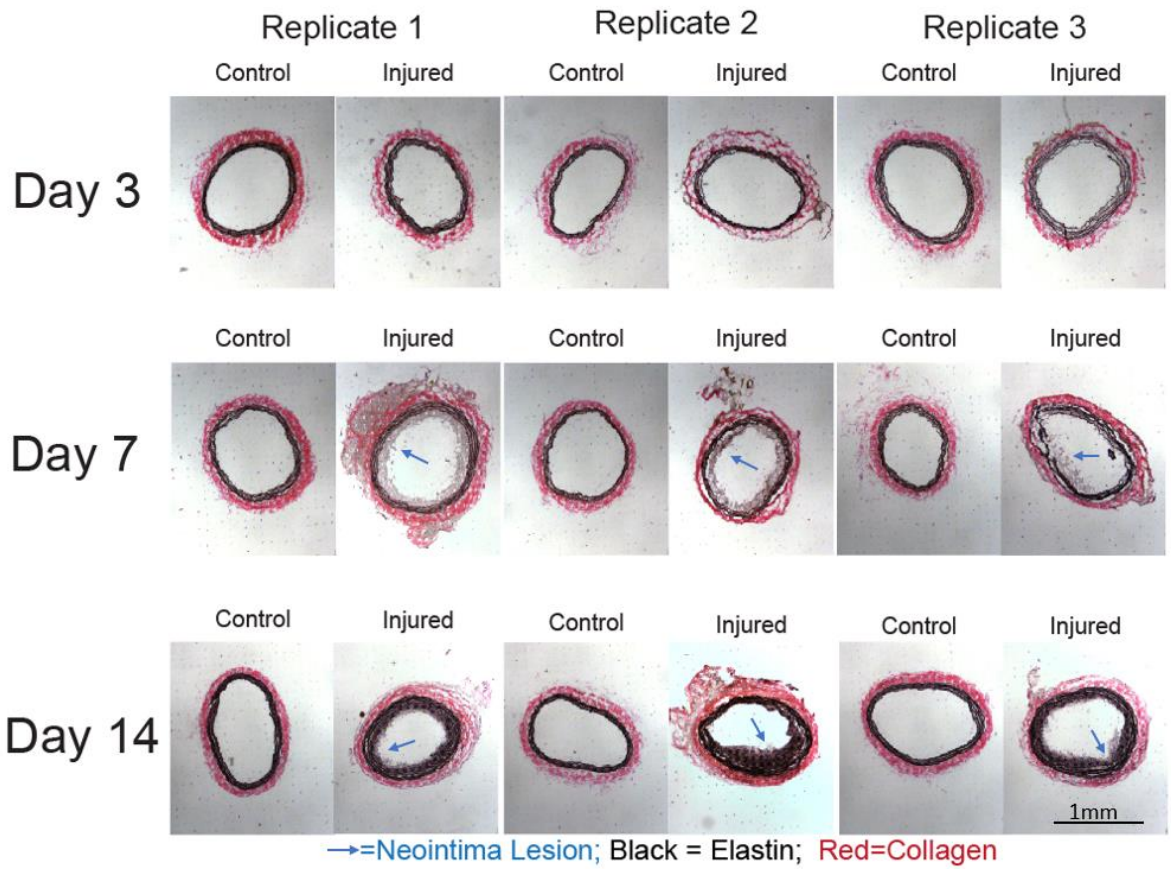
D.

m/z	Tentative Identification	Day 3 AUC	Day 7 AUC	Day 14 AUC
731.6062	SM (36:1)	0.537	0.636	0.537
753.5921	SM (36:1)	0.571	0.75	0.634
759.641	SM (38:1)	0.559	0.596	0.563
781.6229	SM (38:1)	0.589	0.778	0.691
787.6709	SM (40:1)	0.526	0.653	0.562
809.6544	SM (40:1)	0.535	0.745	0.668
801.6844	SM (41:1)	0.524	0.632	0.467
823.671	SM (41:1)	0.507	0.574	0.554
815.7036	SM(42:1)	0.522	0.63	0.509
837.6862	SM(42:1)	0.55	0.778	0.661

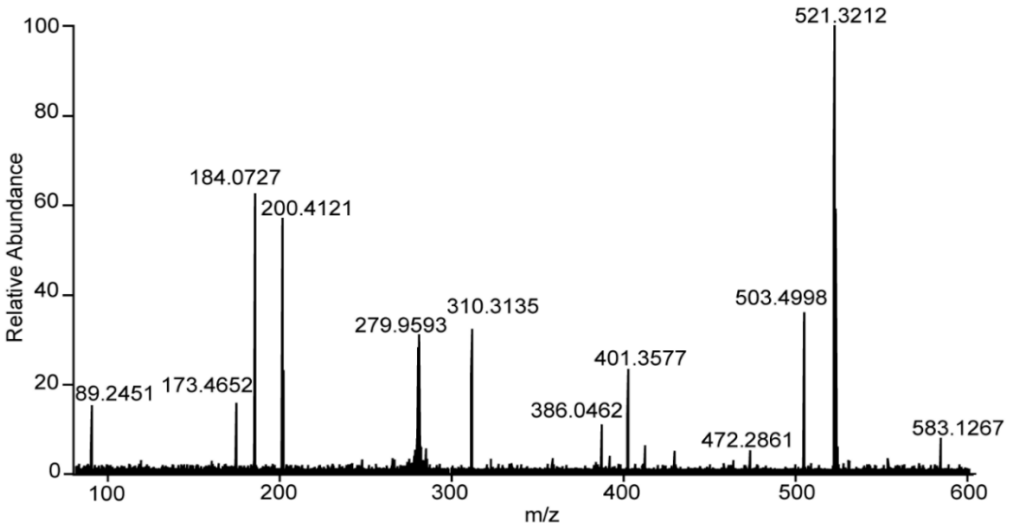
E.

m/z	Tentative Identification	Day 3 t-test	Day 7 t-test	Day 14 t-test
731.6062	SM (36:1)	<0.001	<0.001	<0.01
753.5921	SM (36:1)	<0.001	<0.001	<0.001
759.641	SM (38:1)	<0.001	<0.001	<0.001
781.6229	SM (38:1)	<0.001	<0.001	<0.001
787.6709	SM (40:1)	<0.05	<0.001	<0.001
809.6544	SM (40:1)	<0.01	<0.001	<0.001
801.6844	SM (41:1)	>=0.05	>=0.05	>=0.05
823.671	SM (41:1)	>=0.05	>=0.05	<0.001
815.7036	SM(42:1)	<0.01	<0.001	>=0.05
837.6862	SM(42:1)	<0.001	<0.001	<0.001

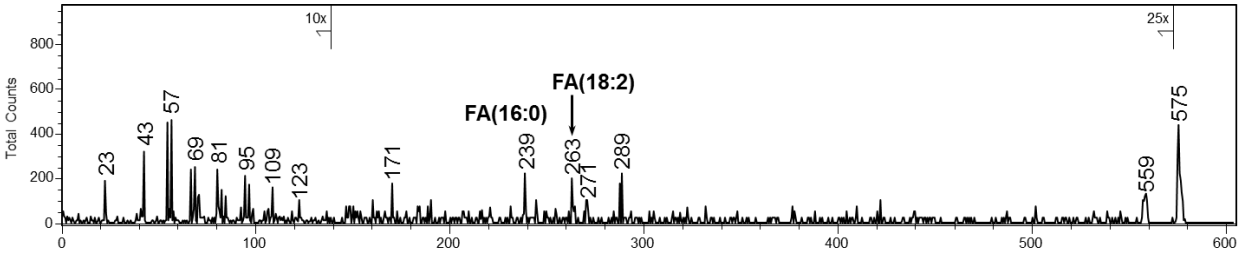
Supplemental Figure S3. Trends in sphingomyelins (SMs) in the development of restenosis.



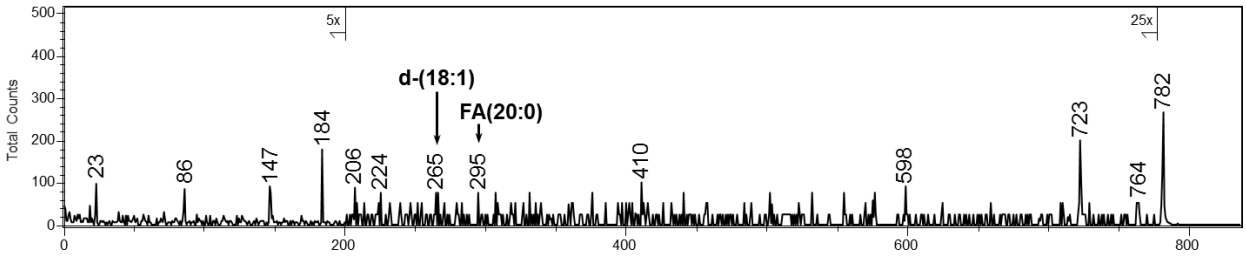
Supplemental Figure S4. Histology annotation of control arteries and arteries with restenosis.



Supplemental Figure S5. Tandem MS/MS spectrum of *m/z* 522.3578 confirming the identification of LysoPC (18:1).



Supplemental Figure S6. Tandem MS/MS spectrum of *m/z* 575.50 confirming the identification of DAG (34:2).



Supplemental Figure S7. Tandem MS/MS spectrum of *m/z* 781.62 confirming the identification of SM (38:1).

Supplemental Table S1. Dynamic trends in injured tissues: t-test statistical comparison of injured arteries: Day 3 vs. Day 14, Day 7 vs. Day 14, and Day 3 vs. Day 7

m/z value	Tentative Lipid ID	Day 3 *vs.* Day 14 t-test	Day 7 *vs.* Day 14 t-test	Day 3 *vs.* Day 7 t-test
575.5052	DAG (34:2)	<0.001	<0.01	<0.001
577.5225	DAG (34:1)	<0.001	<0.001	<0.001
597.4889	DAG (34:2)	<0.001	<0.001	<0.001
599.5063	DAG (34:1)	<0.001	>=0.05	<0.001
601.5208	DAG (36:3)	<0.001	>=0.05	<0.001
603.5382	DAG (36:2)	<0.001	<0.01	<0.001
605.5538	DAG (36:1)	<0.001	<0.05	<0.001
623.5066	DAG (36:3)	<0.001	<0.001	<0.001
625.5219	DAG (36:2)	<0.001	<0.01	<0.001
627.5381	DAG (36:1)	<0.001	>=0.05	<0.001
476.3158	LysoPC (16:1)	>=0.05	<0.001	<0.001
478.3292	LysoPC (16:0)	<0.05	<0.001	<0.001
494.3265	LysoPC (16:1)	>=0.05	<0.001	<0.001
496.3397	LysoPC (16:0)	>=0.05	<0.001	<0.001
502.3314	LysoPC (18:2)	>=0.05	<0.001	<0.001
504.3471	LysoPC (18:1)	<0.001	<0.001	<0.001
520.342	LysoPC (18:2)	>=0.05	<0.01	<0.001
522.3578	LysoPC (18:1)	<0.001	<0.001	<0.001
731.6062	SM (36:1)	<0.001	<0.001	<0.01
753.5921	SM (36:1)	<0.001	<0.001	<0.001
759.641	SM (38:1)	<0.001	<0.001	<0.001
781.6229	SM (38:1)	<0.001	<0.001	<0.001
787.6706	SM (40:1)	<0.001	<0.001	<0.001
801.6844	SM (41:1)	<0.001	<0.05	>=0.05
809.6544	SM (40:1)	<0.01	<0.001	<0.001
815.7036	SM (42:1)	<0.001	<0.001	<0.05
823.671	SM (41:1)	<0.001	<0.01	>=0.05
837.6862	SM (42:1)	<0.001	<0.001	<0.001

Supplemental Table S2. Statistical significance of injured day 7 and day 14 neointima lesions

m/z value	Tentative Identification	t-test Neointima Day 7 *vs.* Neointima Day 14	AUC Neointima Day 7 *vs.* Neointima Day 14
575.5052	DAG (34:2)	<0.001	0.877
597.4889	DAG (34:2)	<0.001	0.882
577.5225	DAG (34:1)	<0.001	0.880
599.5063	DAG (34:1)	<0.01	0.662
601.4508	DAG (36:3)	<0.001	0.898
623.5066	DAG (36:3)	<0.001	0.830
603.5382	DAG (36:2)	>=0.05	0.544
625.5219	DAG (36:2)	<0.001	0.834
605.5538	DAG (36:1)	<0.001	0.767
627.5381	DAG (36:1)	>=0.05	0.500
476.3158	LysoPC (16:1)	<0.001	0.809
494.3265	LysoPC (16:1)	<0.001	0.878
478.3158	LysoPC (16:0)	<0.001	0.707
496.3397	LysoPC (16:0)	<0.001	0.733
502.3314	LysoPC (18:2)	<0.001	0.883
520.342	LysoPC (18:2)	<0.001	0.882
504.3471	LysoPC (18:1)	<0.001	0.738
522.3578	LysoPC (18:1)	<0.001	0.762
731.6062	SM (36:1)	<0.001	0.809
753.5921	SM (36:1)	<0.001	0.842
759.641	SM (38:1)	<0.001	0.791
781.6229	SM (38:1)	<0.001	0.813
787.6706	SM (40:1)	<0.001	0.823
801.6844	SM (41:1)	<0.001	0.801
809.6544	SM (40:1)	<0.001	0.812
815.7036	SM (42:1)	<0.001	0.868
823.671	SM (41:1)	<0.001	0.877
837.6862	SM (42:1)	<0.001	0.913

Supplemental Table S3. Statistical significance of spatial region of injured day 7: comparison of neointima, media, and adventitia regions

m/z value	Tentative Lipid ID	Neointima *vs.* Adventitia	Neointima *vs.* Media	Adventitia *vs.* Media
		t-test	t-test	t-test
575.5052	DAG (34:2)	<0.001	<0.001	<0.05
577.5225	DAG (34:1)	<0.001	<0.001	<0.05
597.4889	DAG (34:2)	<0.001	<0.001	<0.001
599.5063	DAG (34:1)	<0.001	<0.001	<0.05
601.5208	DAG (36:3)	<0.001	<0.001	<0.05
603.5382	DAG (36:2)	<0.001	<0.001	<0.01
605.5538	DAG (36:1)	<0.001	<0.01	<0.01
623.5066	DAG (36:3)	<0.001	<0.001	<0.01
625.5219	DAG (36:2)	<0.001	<0.001	<0.001
627.5381	DAG (36:1)	<0.001	<0.001	<0.001
476.3158	LysoPC (16:1)	<0.001	<0.001	<0.05
478.3292	LysoPC (16:0)	<0.001	<0.001	<0.01
494.3265	LysoPC (16:1)	<0.001	>=0.05	<0.001
496.3397	LysoPC (16:0)	<0.001	<0.001	<0.001
502.3314	LysoPC (18:2)	<0.001	<0.001	<0.001
504.3471	LysoPC (18:1)	<0.001	<0.001	>=0.05
520.342	LysoPC (18:2)	<0.001	<0.001	<0.001
522.3578	LysoPC (18:1)	<0.001	<0.001	<0.05
731.6062	SM (36:1)	>=0.05	<0.001	<0.001
753.5921	SM (36:1)	>=0.05	<0.001	<0.001
759.641	SM (38:1)	>=0.05	<0.001	<0.001
781.6229	SM (38:1)	>=0.05	<0.001	<0.001
787.6706	SM (40:1)	>=0.05	<0.001	<0.001
801.6844	SM (41:1)	<0.05	<0.001	<0.05
809.6544	SM (40:1)	>=0.05	<0.001	<0.001
815.7036	SM (42:1)	>=0.05	<0.001	<0.001
823.671	SM (41:1)	<0.01	<0.001	<0.05
837.6862	SM (42:1)	<0.001	<0.001	<0.001

Supplemental Table S4. Statistical significance of spatial regions in injured day 14 arteries: comparison of neointima, media, and adventitia regions

m/z value	Tentative Lipid ID	Neointima *vs.* Adventitia	Neointima *vs.* Media	Adventitia *vs.* Media
		t-test	t-test	t-test
575.5052	DAG (34:2)	<0.001	>=0.05	<0.001
577.5225	DAG (34:1)	>=0.05	>=0.05	>=0.05
597.4889	DAG (34:2)	<0.001	<0.001	<0.001
599.5063	DAG (34:1)	>=0.05	<0.01	>=0.05
601.5208	DAG (36:3)	<0.05	>=0.05	<0.05
603.5382	DAG (36:2)	<0.001	>=0.05	<0.01
605.5538	DAG (36:1)	>=0.05	<0.001	<0.001
623.5066	DAG (36:3)	<0.001	<0.01	>=0.05
625.5219	DAG (36:2)	<0.001	<0.001	<0.01
627.5381	DAG (36:1)	>=0.05	<0.001	<0.001
476.3158	LysoPC (16:1)	<0.01	<0.001	>=0.05
478.3292	LysoPC (16:0)	<0.001	<0.001	<0.01
494.3265	LysoPC (16:1)	>=0.05	<0.001	<0.001
496.3397	LysoPC (16:0)	>=0.05	<0.001	<0.001
502.3314	LysoPC (18:2)	<0.001	<0.001	<0.01
504.3471	LysoPC (18:1)	<0.01	>=0.05	>=0.05
520.342	LysoPC (18:2)	<0.001	<0.001	<0.001
522.3578	LysoPC (18:1)	<0.001	<0.001	<0.001
731.6062	SM (36:1)	<0.001	<0.001	<0.001
753.5921	SM (36:1)	<0.001	<0.001	<0.001
759.641	SM (38:1)	<0.01	<0.001	<0.001
781.6229	SM (38:1)	<0.01	<0.001	<0.001
787.6706	SM (40:1)	<0.001	<0.001	<0.001
801.6844	SM (41:1)	>=0.05	<0.001	<0.001
809.6544	SM (40:1)	<0.001	<0.001	<0.001
815.7036	SM (42:1)	<0.01	<0.001	<0.001
823.671	SM (41:1)	>=0.05	<0.05	<0.01
837.6862	SM (42:1)	<0.01	>=0.05	<0.001

Chapter 8

Conclusions and Future Directions

Conclusions

In this dissertation, multiple mass spectrometry (MS)-based enabling tools, such as novel detection techniques or new instrumental platform, have been developed for the high throughput analysis of various protein post-translational modifications (PTMs). Moreover, this work also demonstrates the potential applications of current analytical methodologies in pharmaceutical industry and clinical settings.

Protein citrullination is an important protein PTM involved in many devastating diseases. However, due to the lack of method for large-scale and high-throughput analysis, systematic exploration of this disease-related PTM, including its cellular distribution, physiological functions and roles on the onset and progression of diseases, has been rarely studied. In Chapter 2, I designed a novel biotin thiol tag for the specific modification and enrichment of citrullinated peptides. Meanwhile, the biotin tag labeling significantly increased the mass shift of protein citrullination, and thereby avoided the interference of protein deamidation and ^{13}C isotopic peaks for the identification of citrullinated peptide and citrullination site assignment. By combining the biotin tag labeling strategy and high-resolution accurate mass (HRAM) mass spectrometry, a novel biotin tag-assisted, MS-based method was successfully established as the first method enabling large-scale identification of citrullinated proteins in biological samples. The novel method was assessed and optimized using citrullinated peptide standard. Further application of this method to mouse tissue analysis enabled the identification of numerous citrullinated proteins and citrullination sites that have not been reported previously, and the discovery of mouse-tissue specific distribution patterns of citrullinated proteins, expanding our knowledge about this important PTM.

Following the development of biotin tag-assisted, MS-based method (Chapter 2), I developed a novel method, which integrated the biotin tag-assisted, MS-based method with stable-

isotope dimethyl labeling technique, for simultaneous qualitative and quantitative analyses of citrullinated proteins in biological samples. The integration of those two approaches was firstly optimized and demonstrated by using citrullinated peptide standard. The quantitative performance of the integrated method was further evaluated on the citrullinated peptide standard and protein digest from mouse brain. Reliable quantitation accuracy could be archieved although fewer identifications of citrullinated proteins were obtained. To the best of our knowledge, this integrated method is the first method enabling the comprehensive analysis of citrullinated proteins from different biological samples. In addition, in order to increase multiplexing capability, I also tested the coupling of biotin tag-assisted, MS-based method with our 12-plex isobaric DiLeu labeling tags, which has been applicable to citrullinated peptide analysis, but have not been determined on biological samples.

I noticed that another disease-related protein PTM, called protein homocitrullination, could also result in the formation of a ureido group at the side chain of lysine residue not arginine residue. MS-based in-depth profiling of protein substrates with homocitrullination modification is not available due to the lack of enrichment method. Given that the novel biotin thiol tag could target ureido group, it is entirely possible that the biotin tag-assisted, MS-based method that I developed in Chapter 2 could also be used for the qualitative analysis of homocitrullinated proteins in samples. I carried out the mouse tissue-specific protein homocitrullination analysis in Chapter 4, and identified many homocitrullinated proteins for the first time. This result indicated that the biotin tag-assisted, MS-based method was also amenable to the qualitative analysis of homocitrullinated proteins in biological samples, providing a viable solution to global analysis of homocitrullination for the first time.

Quantitative proteomics and phosphoproteomics have been performed in Chapter 5 to

investigate the global changes of proteins and phosphoproteins and explore important signal pathways involved in the abnormal accumulation of IFN-β in LPS/TPG-treated macrophages. 12-plex DiLeu isobaric labeling tags developed by our lab were used for the quantitative analysis. Proteins digest from mouse macrophages treated with four different conditions (Control, LPS, TPG and LPS/TPG) were labeled by different DiLeu channels. Phosphopeptides were enriched by using Titanium (IV) immobilized metal affinity chromatography (Ti(IV)-IMAC). Several signal pathways were upregulated in the LPS/TPG-treated macrophages, such as DNA replication, RNA transport, ribosome biogenesis and splicesome, while TNF signaling pathway and necroptosis were down-regulated.

In the chapter 6, a novel platform was developed by coupling a novel sub atmospheric pressure (SubAP)/ MALDI source with high-resolution accurate mass (HRAM) mass spectrometry for the *in-situ* imaging of N-glycans from formalin-fixed paraffin-embedded (FFPE) tissue sections. Parameters of the SubAP MALDI source were optimized using a mixture of N-glycans released from glycoprotein standard. A standard workflow was established for the preparation of FFPE tissue sections for MALDI imaging. High-resolution N-glycan images were collected from FFPE mouse brain section. I further applied the novel platform to map the spatial distribution of N-glycans from mouse FFPE ovarian cancer tissue section. Different spatial distribution patterns of N-glycans, especially the abnormal accumulation of high-mannose N-glycans in cancer area, were observed, demonstrating the applicability of the new platform to biological sample analysis.

MALDI imaging MS was also applied to reveal the spatial distribution of biomolecules in rat arteries undergoing restenosis (Chapter 7). The distinct temporal-spatial dynamics of bioactive lipids in the restenotic arterial wall, including diacylglycerols (DAGs), lysophosphatidylcholines (LysoPCs) and sphingomyelins (SMs), were explored for the first time. Additionally, we employed

time-of-flight secondary ion mass spectrometry (TOF-SIMS) tandem MS imaging for both molecular identification and imaging at high spatial resolution. These imaging modalities provide powerful tools for unraveling novel mechanisms of restenosis involving lipids or small signaling molecules.

Future Directions

1. Development of a novel FATRAC method for qualitative analysis of citrullinated and homocitrullinated proteins

In the Chapter 2 and Chapter 4, we reported synthesis of a novel biotin thiol tag in one step by using biotin-NHS ester and cysteamine. We further demonstrated that the novel biotin thiol tag, together with 2, 3-butanedione, could specifically modify the ureido group of the citrullinated peptide or homocitrullinated peptide. Based on this characteristic reaction, we developed a novel biotin tag-assisted, MS-based method, which is the first method enabling the large-scale identification of citrullinated and homocitrullinated proteins in biological samples.

Although effective, the biotin tag-assisted MS-based approach developed in the Chapter 2 and Chapter 4 suffers from several challenges including: 1) Reduced hydrophilicity and solubility of biotin tag-modified citrullinated peptides interferes with the following enrichment process and MS analysis; 2) Streptavidin beads are required for enrichment; 3) To release biotin tag-labeled citrullinated peptides or homocitrullinated peptides from beads, harsh conditions must be used to denature streptavidin resulting high noise background upon MS analysis; 4) Biotin head group decreases the ionization efficiency of enriched citrullinated peptide upon MS analysis; 5) MS fragmentation of the biotin head group suppresses peptide backbone fragmentation, resulting in a large number of low-quality tandem MS spectra.

To address these challenges, we propose to develop a novel second-generation method, called filter-aided, thiol resin-assisted chemoenzymatic (FATRAC) method, for the improved qualitative analysis of citrullinated proteins in biological samples.

The principle of the novel FATRAC method for MS-based qualitative analysis of citrullinated proteins in biological samples is summarized in **Figure 1**. We will firstly customize novel thiol resin beads using commercial NHS-activated resin beads and a short peptide containing Arg-Cys (RC) at the C terminus (**Figure 1a**). As shown in **Figure 1b**, the novel thiol beads contain three functional groups: 1) A free thiol group at the cysteine residue, could work together with 2,3-butanedione to capture citrullinated or homocitrullinated peptides in biological samples; 2) The amide bond between arginine and cysteine is trypsin cleavable, and thereby could be used for the release of enriched citrullinated and homocitrullinated peptides; 3) Resin beads act as solid support, contributing to the easy removal of non-citrullinated peptides and the immobilization of citrullinated and homocitrullinated peptides. Hence, the novel thiol beads enable simultaneous chemical derivatization and enrichment of citrullinated peptides from biological samples, avoiding the usage of biotin tag and streptavidin beads for enrichment. More importantly, citrullinated and homocitrullinated peptides enriched by the FATRAC method will have a small chemical modification instead of an entire biotin tag, avoiding all the MS detection challenges mentioned above caused by the biotin head group (**Figure 1c)**.

In addition, as shown in **Figure 1d**, the application of a spin filter will further simplify the sample pretreatment process. All procedures, including the synthesis of thiol resin beads, chemical derivatization, removal of excess derivatization reagent, the enrichment and release of citrullinated or homocitrullinated peptides could be performed in a small spin filter. The enriched peptides could be simply desalted with C_{18} reverse phase columns before MS analysis.

We have tested the FATRAC method using our synthesized citrullinated peptide standard (SAVRACitSSVPGVR). As shown in **Figure 2**, the thiol resin could be successfully prepared by mixing 50µl NHS-activated resin beads with 0.33mg GGGGRC for 1h at room temperature. The thiol resin was further reacted with 1µg citrullinated peptide standard and 100mM 2,3-butanedione in 12.5% TFA solution. On-resin chemical derivatization and sample cleaning were performed as shown in **Figure 3a.** After trypsin digestion overnight, the enriched citrullinated peptide standard could be successfully released from the resin beads (**Figure 3b**). This preliminary data demonstrated the feasibility of the FATRAC method for large-scale identification of citrullinated proteins in biological samples.

2. Integration of multiplexed dimethylated leucine (DiLeu)-based quantitation strategy with FATRAC method for high throughput protein citrullination analysis.

Given the fact that there is currently no MS-based method for quantitative protein citrullination analysis, we reported the development of the first quantitative method in Chapter 3 by combining the biotin tag-assisted method and stable-isotope reductive dimethyl labeling strategy. However, stable-isotope reductive dimethyl labeling strategy suffers from low-throughput and is not amenable to examine larger number of clinical specimens simultaneously, such as human CSF and plasma samples collected from healthy control and AD patients at different stages. Hence, it is highly desirable to develop high throughput MS-based quantitative method for the parallel analysis of a large number of biological samples. We also tested the integration of 12-plex DiLeu isobaric tags and the biotin tag-assisted method for multiplex quantitative analysis of citrullinated proteins from different biological samples, which only evaluated by using the citrullinated peptide standard.

Here, one of our future direction is to integrate the 12-plex DiLeu labeling strategy with the novel FATRAC method mentioned above for the multiplex and high throughput quantitative analysis of citrullinated proteins in samples, which is shown in **Figure 4**.

Tryptic digested peptides from different samples will firstly be differentially labeled with different DiLeu channels, and then be combined for further FATRAC analysis. The 12-plex DiLeu labeled citrullinated peptides will finally be released from beads by trypsin digestion. Although the quantitation accuracy of DiLeu labeling strategy has been demonstrated, the quantitation accuracy of the integrated method will still be systematically evaluated using the citrullinated peptide standard and complex biological samples.

3. Comparative analysis of citrullinated and homocitrullinated proteins in human CSF and plasma using the FATRAC method and DiLeu labeling strategy.

Abnormal accumulation of citrullinated proteins and the colocalization of citrullinated proteins with neurotoxic *Aβ*42 have been observed in the brain of AD patients, suggesting the association of protein citrullination with AD.[1-3] However, previous studies mainly focused on the changes of citrullinated proteins in the brain, while the alternation of citrullinated proteins in both CSF and plasma from control and AD patients at different stages, have not been studied.

Given the important role of protein citrullination in AD, one future direction of our work will be the comparative analysis of citrullinated and homocitrullinated proteins in CSF and plasma from different and control and AD patients at different stages using the FATRAC method and DiLeu labeling strategy, which could be crucial for unraveling the role of protein citrullination and homicitrullination in AD (**Figure 5**).

A pilot study has been performed to test the citrullinated proteins in the CSF samples of a healthy control, a mild cognitive impairment (MCI) and an AD patient. From each 400μg protein

digest, we successfully identified 169 citrullination sites from 94 citrullinated proteins, 173 citrullination sites from 92 citrullinated proteins and 174 citrullination sites from 88 citrullinated proteins in the CSF sample of healthy control, MCI and AD patients, respectively (**Figure 6**). Some citrullination sites were detected only in the CSF samples of MCI and AD, but not in the healthy control. One particularly interesting citrullinated protein was the Apolipoprotein E (ApoE). As shown in **Figure 7a**, three citrullination sites at R108, R152 and R185 were only identified from ApoE in the CSF samples of MCI and AD patient. A representative HCD spectrum of the citrullinated peptide (Cit_152) from the ApoE in CSF of AD patient, is shown in **Figure 7b**. Two diagnostic ions at *m/z* 227 and 304 produced from biotin tag group were clearly observed. This preliminary data provided an initial glimpse at the citrullinated proteins in human CSF and revealed some promising evidence for the discovery of citrullinated protein biomarkers of AD. More systematic and in-depth analysis will likely reveal more citrullinated proteins as potential diagnostic molecular signatures of AD pathogenesis and progression.

References

(1). Acharya, N. K.; Nagele, E. P.; Han, M.; Coretti, N. J.; DeMarshall, C.; Kosciuk, M. C.; Boulos, P. A.; Nagele, R. G., Neuronal PAD4 expression and protein citrullination: possible role in production of autoantibodies associated with neurodegenerative disease. *J Autoimmun* 2012, 38, (4), 369-80.

(2). Ishigami, A.; Ohsawa, T.; Hiratsuka, M.; Taguchi, H.; Kobayashi, S.; Saito, Y.; Murayama, S.; Asaga, H.; Toda, T.; Kimura, N.; Maruyama, N., Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease. *J Neurosci Res* 2005, 80, (1), 120-8.

(3). Nicholas, A. P., Dual immunofluorescence study of citrullinated proteins in Alzheimer diseased frontal cortex. *Neurosci Lett* 2013, 545, 107-11.

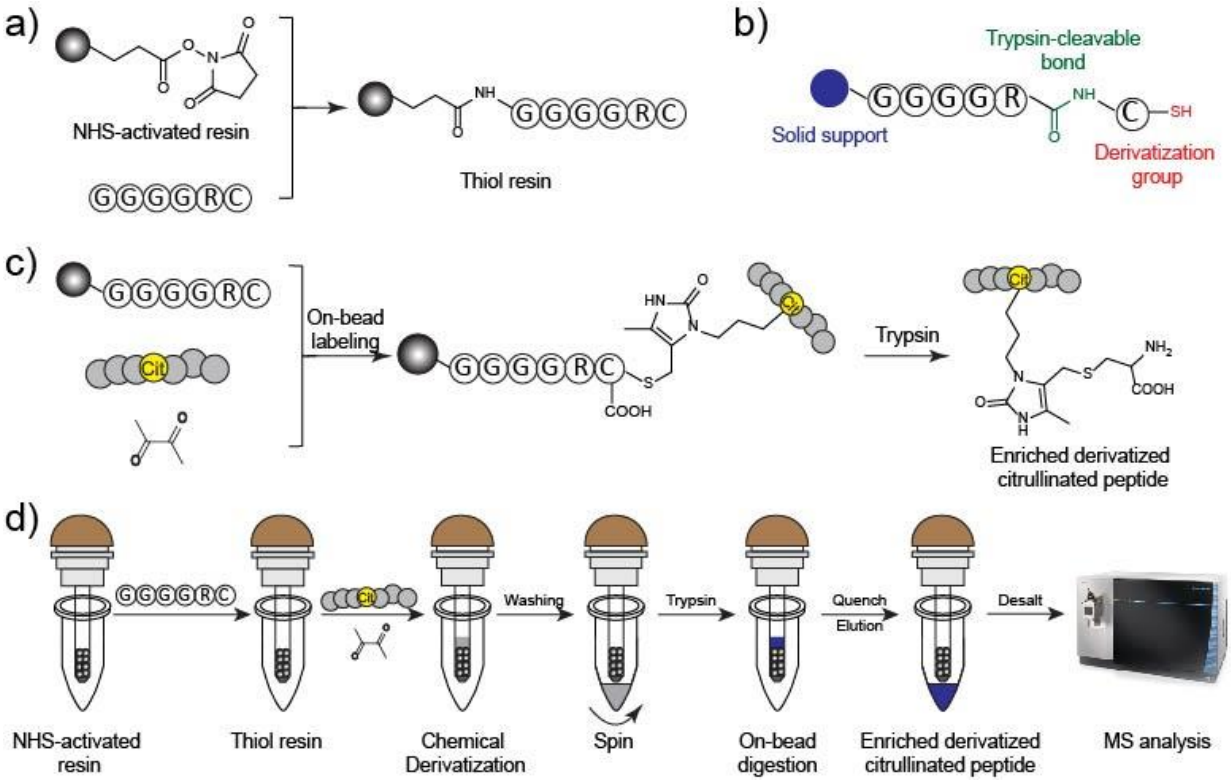


Figure 1. Principle of the novel FATRAC method for MS-based protein citrullination and homocitrullination analysis. a) Schematic illustration of the preparation of thiol resin beads; b) Three functional units of thiol resin beads; c) Principle of thiol resin-assisted derivatization and enrichment of citrullinated peptides; d) Workflow of the novel FATRAC method for large-scale protein citrullination and homocitrullination analysis.

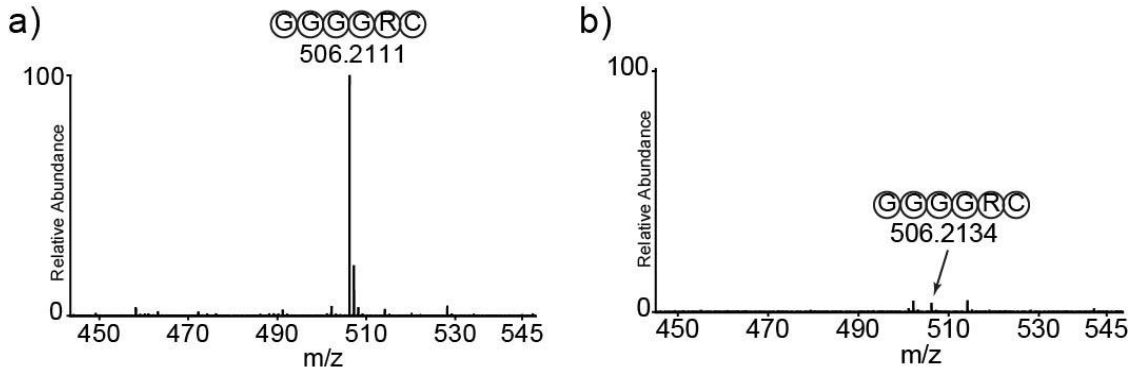


Figure 2. Synthesized peptide GGGGRC before (a) and after (b) reacting with NHS-activated resin. MS spectra were collected by using MALDI-Orbitrap mass spectrometer.

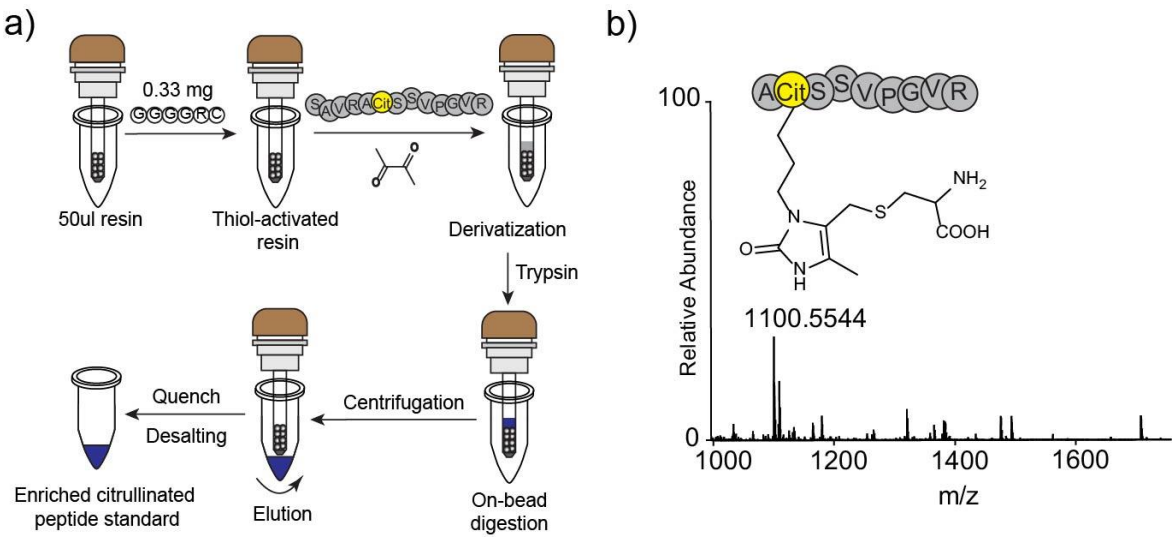


Figure 3. Evaluation of the FATRAC method using citrullinated peptide standard (a) and Citrullinated peptide standard enriched by FATRAC method. MS spectrum were collected by using MALDI-Orbitrap mass spectrometer.

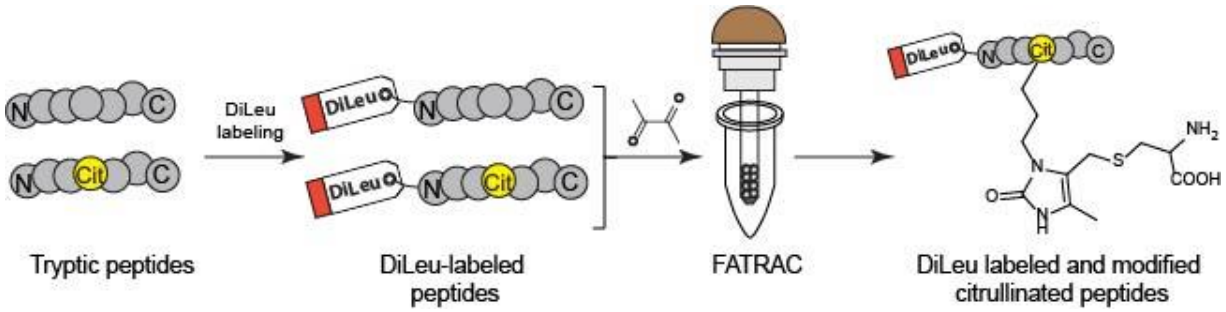


Figure 4. Combination of DiLeu tag labeling with the FATRAC method for multiplex and high throughput quantitative analysis of citrullinated proteins.

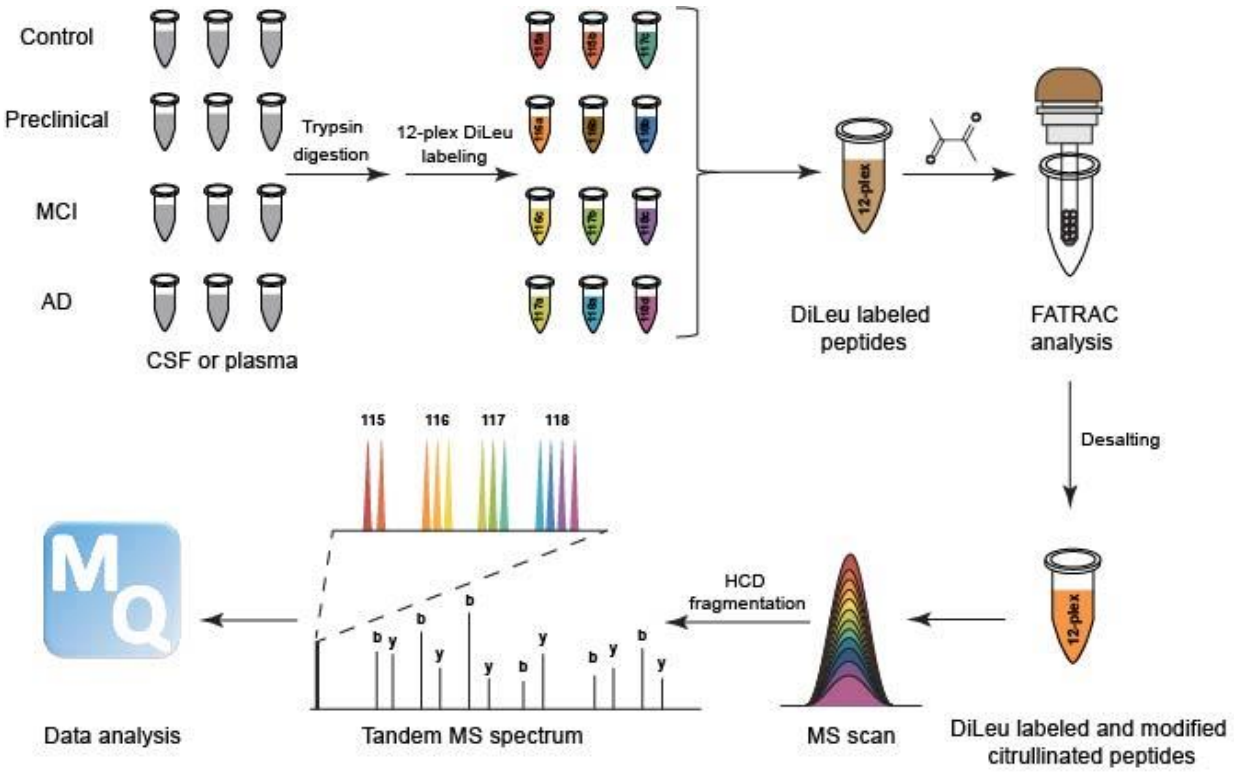


Figure 5. Comparative analysis of citrullinated and homocitrullinated proteins in human CSF and plasma using the FATRAC method and DiLeu labeling strategy.

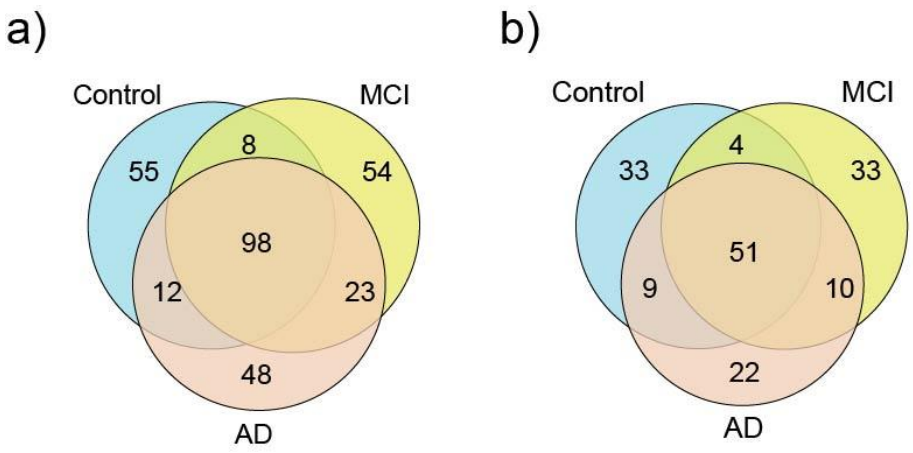


Figure 6. Overlap of citrullination sites (a) and citrullinated proteins (b) identified from CSF sample of healthy control, MCI and AD patient, respectively.

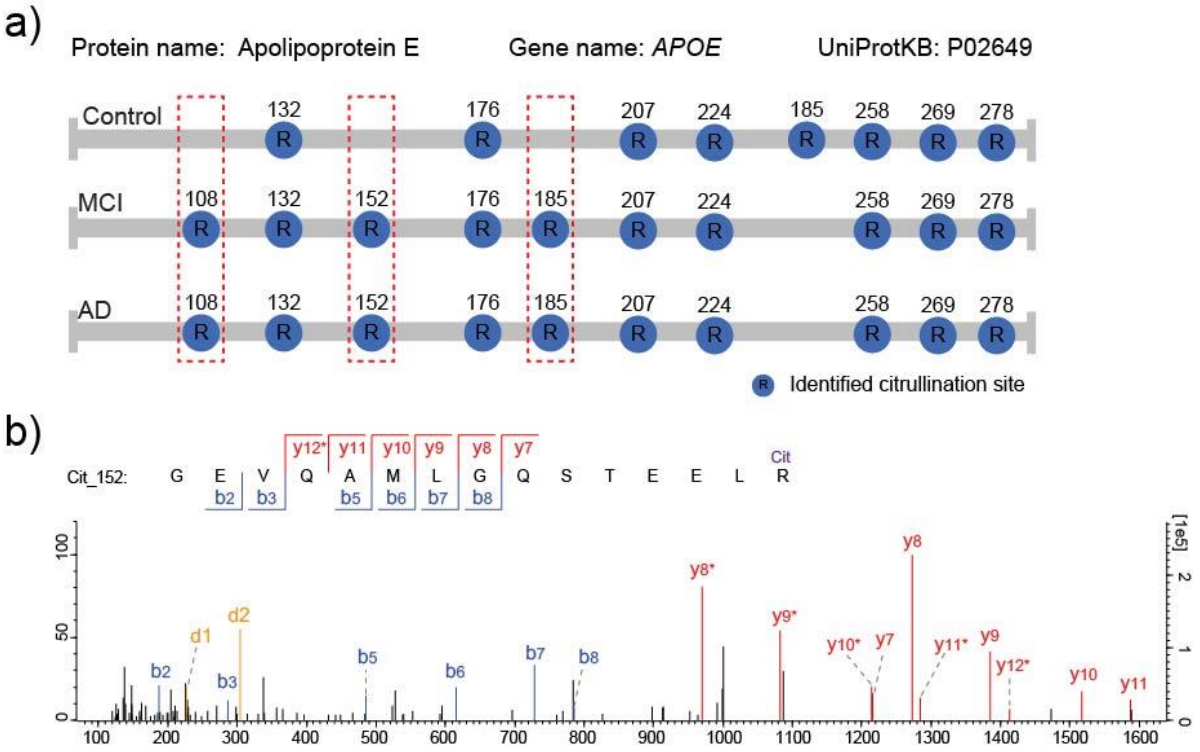


Figure 7. Citrullination sites identified from ApoE in the CSF of control, MCI and AD patient, respectively (a). HCD spectrum of citrullinated peptide with citrullination site at R152 from the ApoE in CSF of AD patient (b). d1: diagnostic ion at *m/z* 227; d2 diagnostic ion at *m/z* 304;

Appendix I

Publications and Presentations

Peer-Reviewed Publications

1. **Shi, Y.**; Li, Z.; Wang, B.; Ye, Z.; Ye, H; Shi, X.; Chen, Z.; Ma, F; Li, L.; Simultaneous mass spectrometry-based large-scale identification and relative quantitation of citrullinated proteins from complex biological samples. *Manuscript in preparation.*

2. **Shi, Y.**; Li, Z.; Felder, M.; Yu, Q.; Shi, X.; Peng, Y.; Cao Q.; Wang, B.; Puglielli, L.; Patankar, M.; Li, L.; High-resolution mass spectrometry imaging of N-glycans from formalin-fixed paraffin-embedded tissue sections using a novel sub-atmospheric pressure ionization source. *Anal Chem*. **2019**, 91, 20, 12942-12947.

3. **Shi, Y.**; Cui, Y.; Li, Z.; Huang J; Wang, B.; Shi, X.; Li, L.; Investigation of the abnormal accumulation of IFN-β in LPS/TPG-treated macrophages by using mass spectrometry-based quantitative proteomics. *Manuscript in preparation.*

4. **Shi, Y.**; Johnson, J.; Wang, B.; Chen, B.; Fisher, G. L.; Urabe, G.; Shi, X.; Kent, K. C.; Guo, L. W.; Li, L.; Mass Spectrometric Imaging Reveals Temporal and Spatial Dynamics of Bioactive Lipids in Arteries Undergoing Restenosis. *J. Proteome Res* **2019**, 18, (4), 1669-1678.

5. Hao, L.; **Shi, Y.**; Thomas, S.; Vezina, C. M.; Bajpai, S.; Ashok, A.; Bieberich, C. J.; Ricke, W. A.; Li, L.; Comprehensive urinary metabolomic characterization of a genetically induced mouse model of prostatic inflammation. *Int J Mass Spectrom* **2018**, 434, 185-192.

6. Hao, L.; Greer, T.; Page, D.; **Shi, Y.**; Vezina, C. M.; Macoska, J. A.; Marker, P. C.; Bjorling, D. E.; Bushman, W.; Ricke, W. A.; Li, L.; In-Depth Characterization and Validation of Human Urine Metabolomes Reveal Novel Metabolic Signatures of Lower Urinary Tract Symptoms. *Sci Rep* **2016**, 6, 30869.

7. Vestal, M.; Li, L.; Dobrinskikh, E.; **Shi, Y.**; Wang, B.; Shi, X.; Li, S.; Vestal, C.; Parker, K.; Rapid MALDI-TOF Molecular Imaging: Instrument Enhancements & Their Practical Consequences. *J. Mass Spectrom* **2019,** 1-8.

8. Stallcop, L.E.; Álvarez-García, Y.R.; Reyes-Ramos, A.M.; Ramos-Cruz, K.P.; Morgan, M.M.; **Shi, Y.**; Li, L.; Beebe, D.J.; Domenech, M.; Warrick, J.W.; Razor-printed sticker microdevices for cell-based applications. *Lab Chip.* **2018** Jan 30;18(3):451-462.

9. Zhu, Y.; Takayama, T.; Wang, B.; Kent, A.; Zhang, M.; Binder, B.Y.; Urabe, G.; **Shi, Y.**; DiRenzo, D.; Goel, S.A.; Zhou, Y.; Little, C.; Roenneburg, D.A.; Shi, X.; Li, L.; Murphy, W.L.; Kent, K.C.; Ke, J.; Guo, LW.; Restenosis Inhibition and Re-differentiation of TGFβ/Smad3-activated Smooth Muscle Cells by Resveratrol. *Sci Rep.* **2017** Feb 6; 7: 41916.

10. Lu, G.; Xu, X.; Li, G.; Sun, H.; Wang, N.; Zhu, Y.; Wan, N.; **Shi, Y.**; Wang, G.; Li, L.; Hao, H.; Ye, H.; Sub-residue Resolution Footprinting of Ligand-Protein Interactions by Carbene Chemistry and Ion Mobility-Mass Spectrometry. *Anal Chem. Accepted.*

Patent

- Li, L.; **Shi, Y.**; Li, Z; "Development of a Novel Mass Spectrometry-Based Method for Simultaneous Qualitative and Quantitative Protein Citrullination Analysis of Complex Biological Samples". (Submitted)

Conference Presentations

Oral Presentation

- **Shi, Y.**; Li, Z.; Shi, X; Wang, B.; Li, L., "Mass spectrometry-based large-scale and precise identification of citrullinated proteins from complex biological samples." Oral presentation at The 67th American Society for Mass Spectrometry (ASMS) Annual Conference, 2019, Atlanta. GA.

Poster Presentation

- **Shi, Y.**; Li, Z.; Chen, Z.; Li, G.; Cao, Q.; Li, L., "High-resolution imaging of N-glycans from FFPE tissue sections using sub-atmospheric pressure MALDI source." Poster presentation at The 66th American Society for Mass Spectrometry Annual Conference, 2018, San Diego. CA.
- **Shi, Y.**; Chen, Z.; Yu, Q.; Wang, B.; Glover, M.; Shi, X.; Guo, L.; Kent, K.C.; Li, L., "Large-scale Characterization and Quantitation of Citrullinated Proteins Involved in Restenosis by HCD Product Ion Triggered EThcD Mass Spectrometry." Poster presentation at The 65th American Society for Mass Spectrometry Annual Conference, 2017, Indianapolis. IN.
- **Shi, Y.**; Wang, B.; Chen, B.; Shi, X.; Guo, L.; Kent, K.C.; Li, L., "MALDI-MS visualization reveals the involvement of diacylglycerols and lysophosphatidylcholines in the progression of restenosis." Poster presentation at The 64th American Society for Mass Spectrometry Annual Conference, 2016, San Antonio. TX.
- Li, Z; **Shi, Y.**; Ma, F; Gant, K.L.; Patankar, M.S.; Li, L.; "MS-Based Strategies Reveal Extracellular Matrix Alterations and N-Glycan Spatial Distribution Changes with the Progression of Ovarian Cancer." Poster presentation at The 67th American Society for Mass Spectrometry (ASMS) Annual Conference, 2019, Atlanta. GA.
- Ma, M.; **Shi, Y.**; Cui, Y.; Huang, J.; Liu, Y.; Smith, A.J.; Li, L.; "Systematic Proteomic Analysis of the interaction between UPR and LPS regulated Phosphorylation Establishes Novel Connections to Innate Immunity." Poster presentation at The 67th American Society for Mass Spectrometry (ASMS) Annual Conference, 2019, Atlanta. GA.
- Wang, B.; **Shi, Y.**; Li, Z.; Shi. X.; Tao, N.; Li L.; "Comparative Study of Pancreatic Insulin and N-Glycans between Lean and Obese Zucker Rats by MALDI Imaging Mass Spectrometry." Poster presentation at The 67th American Society for Mass Spectrometry (ASMS) Annual

Conference, 2019, Atlanta. GA.

- Lu, G.; Wang, N.; Tian, Y.; Wan, N.; **Shi, Y.**; Li, G.; Li, L.; Hao, H.; Ye, H.; "Sub-Residue Resolution Footprinting of LigandProtein Interactions Enabled by Ion Mobility Mass Spectrometry." Poster presentation at The 67th American Society for Mass Spectrometry (ASMS) Annual Conference, 2019, Atlanta. GA.
- Xu, M.; Johnson, J.; **Shi, Y.;** Shi, X.; Keikhosravi, A.; Eliceiri, K.; Skala, M. C.; Kao, W. J.; Li, L.; "Combining Enzymes to Increase Glycan Profiling in Human Pancreatic Cancer Tissues Using Mass Spectrometry Imaging." Poster presentation at The 66th American Society for Mass Spectrometry Annual Conference, 2018, San Diego. CA.
- Chen, Z.; Zhong, X; Cai, T; **Shi, Y.**; Zhang, X.; Li, L.; "Hydrophilic Interaction Liquid Chromatography-Mass Spectrometric Imaging Platform for N-glycan Relative Quantitation using Stable-Isotope Labeled Hydrazide Reagents." Poster presentation at The 64th American Society for Mass Spectrometry Annual Conference, 2016, Indianapolis. IN.