Development of Novel Chemical Probes to Facilitate the Qualitative and Quantitative Analysis of Biomolecules Enabled by Mass Spectrometry

By

Miyang (Mike) Li

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Analytical Chemistry)

at the

University of Wisconsin-Madison

2022

Date of final oral examination: 06/16/2022

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

Lingjun Li, Professor, Chemistry and School of Pharmacy

Ying Ge, Professor, Chemistry and Cell and Regenerative Biology

Manish S Patankar, Professor, Obstetrics and Gynecology

Martha M Vestling, Director of Mass Spec Lab, Chemistr

Acknowledgements

I would first like to express my sincere gratitude to my advisor Dr. Lingjun Li, without whom these works would not be possible. She has given me continuous support and freedom to allow me to achieve my academic and professional goals. Her patience and hard work have always been motivating me through my Ph.D. career and will continue to do so in my pursuit of becoming an independent scientist. It has been a great honor to study and work in Li lab and I will forever be thankful for this opportunity and experience.

I would also like to thank my committee members, Dr. Ying Ge, Dr. Manish S Patankar, and Dr. Martha M Vestling for their time and feedback regarding the prelim exam, annual committee meeting, and this dissertation. Their support has been instrumental to my development as a scientist.

Furthermore, I would like to acknowledge Dr. Yu Feng, Dr. Junfeng Huang and Dr. Min Ma for either mentoring or working closely with me in Li lab. Yu introduced me to the Li lab as well as mentored me and taught me a lot about mass spectrometry and organic synthesis during my junior years. This experience stimulated my scientific curiosity and helped me become an independent researcher. It is because of them that I could achieve so much progress on my projects and their teaching and help had a great impact toward setting me on my scientific career path.

Through my entire graduate career, I have always been helped by my colleagues. I would like to thank a few of them who have made great contributions to my projects. Dr. Xiaorong Lin and Dr. Yuanyuan Lin were of great help in doing the experiments and working with me. Dr. Xiaofang Zhong has provided her expertise on the enrichment of glycopeptides and data analysis. Dr. Fengfei Ma introduced me to proteomics and taught me how to perform sample preparation.

I would also like to extend my thanks to Dr. Stephen Block and Dr. Liana Lamont whom I have been honored to work with as a teaching assistant. In addition, I would like to give special thanks to Arrietta W Clauss and Taylor Mathewson, the fantastic graduate program coordinators who have been extremely helpful providing their support for all graduate students.

My final thanks go to my family. My mother and father have provided me unconditional love and support throughout my life. Without their support and understanding, I would not have been able to make these achievements.

Table of Contents

Development of Novel Chemical Probes to Facilitate the Qualitative and Quantitative Analysis of Biomolecules Enabled by Mass Spectrometry

Miyang (Mike) Li

Under the supervision of Professor Lingjun Li
At the University of Wisconsin-Madison

Abstract

Liquid chromatography coupled with mass spectrometry (LC-MS) has evolved as a powerful tool in the identification and quantification of various biomolecules. It has the capability to monitor thousands of metabolites and peptides within hours at tremendous sensitivity and therefore plays a critical role in the study of underlying biological mechanisms and the discovery of novel therapeutic targets. However, sometimes it requires the assistance of chemical derivatization to enhance analytical performance or obtain quantitative information, as the analytes of interest may not be suitable for direct analysis at their native states or would carry quantitative information after introducing isotopic labels. Therefore, there have been continuous interests and efforts to develop specialized chemical probes to meet the demands. This dissertation is devoted to the development and application of novel chemical probes to facilitate the analysis of biomolecules such as glycans, peptides and proteins.

The first part of this dissertation focuses on the quantitative glycomics analysis. Two specialized chemical probes, the mass defect isobaric multiplex reagents for carbonyl containing compound (mdSUGAR) tags and dimethyl leucine containing pyrazolone analog (DiLeuPMP) were made to analyze the released N-glycans and O-glycans, respectively. Although both tags

derivatize the reducing ends of the released glycans, they achieved the labeling process in a different manner. While mdSUGAR tags utilized reductive amination reaction to label released N-glycans from PNGaseF, DiLeuPMP derivatized O-glycans through beta-elimination in a one-pot reaction. In addition, mdSUGAR tags achieved quantification on MS1 level with the aid of ultra-high resolution mass spectrometer but DiLeuPMP relied on reporter ions in MS2 spectrum.

The second part of this dissertation discussed about qualitative and quantitative analysis of sialylglycopeptides (SGPs). It is known that protein sialylation plays pivotal roles in cellular communication and biological recognition. Taking the advantage of the special chemical structure of sialic acids, a novel pipeline combining hydrazide chemistry, click chemistry and dynamic covalent exchange to selectively enrich SGPs was developed. As for the quantitative analysis, a set of isobaric *N*,*N*-dimethyl Leucine derivatized ethylenediamine (DiLeuEN) tags were designed to introduce not only the isotopic labels but also extra positive charges through charge reversal derivatization onto SGPs. The elevated charge states of precursors were much preferable to electron-transfer/higher-energy collision dissociation (EThcD) fragmentation and abundant *c/z* ions were produced to enable unambiguous site-specific mapping of glycosylation.

The last part of this dissertation shifts focus from peptidome to proteome. A novel tetrafunctional chemical probe named cleavable DiLeu-biotin-Azide (cDBA) was designed and synthesized. The cDBA contains a dimethylated leucine moiety as the reporter group during MS/MS, and an azide group for the copper-catalyzed alkyne azide cycloaddition (CuAAC) click chemistry conjugation, as well as a biotin group for the biotin-streptavidin enrichment. Within the structure, there is also a cleavable linker for the facile release of enriched peptides. The delicate design made cDBA tag a powerful chemical probe for the biorthogonal incorporation and enrichment of the analytes of interests, namely peptides and proteins, in complex samples such as

cell lysates, and performing multiplex quantitative analysis in downstream LC-MS/MS. It has the great potential to be developed as a platform for not only profiling amino acids residues' reactivity but also high-throughput quantitative pan-PTM analysis.

List of Abbreviations and Acronyms

ACN	Acetonitrile
AD	Alzheimer's disease
AGC	Automatic gain control
ANP	atrial natriuretic peptide
BCA	Bicinchoninic acid assay
BST	Biotin switch technique
CID	Collision-induced dissociation
cDBA	Cleavable DiLeu-Biotin-Azide
CuAAC	Copper catalyzed alkyne azide cycloaddition
Da	Dalton, unit of mass measurement (1 Da = 1 g/mol)
DCC	Dicyclohexylcarbodiimide
DiLeu	N,N-dimethyl Leucine
DiLeuC	DiLeu as complementary ion based isobaric tags
DiLeuEN	N,N-dimethyl Leucine derivatized ethylenediamine
DiLeuPMP	dimethyl leucine containing pyrazolone analog
DTT	Dithiothreitol
EDTA	Ethylenediaminetetracetic acid
EDCI	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide

ESI	Electrospray ionization
ETD	Electron transfer dissociation
EThcD	Electron-transfer/higher-energy collision dissociation
FA	Formic acid
FDR	False discovery rate
GO	Gene ontology
HA	Hydrazide alkyne
HCD	Higher-energy collisional dissociation
HILIC	Hydrophilic interaction chromatography
HPLC	High-performance liquid chromatography
HRAM	High-resolution, accurate mass
HOBt	Hydroxybenzotriazole
IT	Ion trap
iTRAQ	Isobaric tag for relative and absolute quantification
isoTOP-ABPP	isotopic tandem orthogonal proteolysis–activity-based protein profiling
isoBOP-ABPP	isobaric tandem orthogonal proteolysis activity-based protein profiling
IPM	2-indo-N-(prop-2-yn-yl)acetamide
LC	Liquid chromatography
Lys-C	Endoproteinase lysine C

m/z	mass-to-charge ratio
mdSUGAR	mass defect isobaric multiplex reagents for carbonyl containing compound
MALDI	Matrix-assisted laser desorption ionization
MRM	multiple reaction monitoring
MS	Mass spectrometry
MS^1	Precursor mass analysis
MS^2	Tandem mass spectrometry
NCE	Normalized collision energy
NO	Nitric oxide
NMM	N-methylmorpholine
oxiPTM	oxidative post-translational modification
PD	Parkinson's diseases
ppm	Parts per million
PSM	Peptide-spectrum match
PMP	1- phenyl-3-methyl-5-pyrazolone
protoDBA	prototype of DBA tags
PRM	Parallel reaction monitoring
PTM	Post-translational modification
PyBOP	benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate

RPLC	Reversed phase liquid chromatography
RT	Retention time
ROS/RNS	Reactive oxygen and nitrogen species
SNO	S- nitrosylation
SOH	S- sulphenation
SO2H	S- sulphination
SRM	Selected reaction monitoring
S/N	Signal-to-noise ratio
SCX	Strong cation exchange chromatography
SGP	sialylglycopeptides
SPS-MS^3	Synchronous precursor selection MS^3
TEV	tobacco etch virus
TFA	Trifluoroacetic acid
TIC	Total ion chromatogram
TMT	Tandem mass tag
XIC	Extracted ion chromatogram
z	charge
TCEP	tris(2-carboxyethyl) phosphine
MWCO	Molecular weight cut-off

Chapter 1

Introduction and Research Summary

Introduction and Research Summary

Liquid chromatography-mass spectrometry (LC-MS) has evolved as a powerful tool in the identification and quantification of various biomolecules. It has the capability to monitor thousands of metabolites and peptides within hours at tremendous sensitivity and therefore plays a critical role in the study of underlying biological mechanisms and the discovery of novel therapeutic targets. However, sometimes it requires the assistance of chemical derivatization to enhance analytical performance or obtain quantitative information, as the analyte of interest may not be suitable for direct analysis at their native states or would carry quantitative information after introducing isotopic labels. Therefore, there have been continuous interests and efforts to develop specialized chemical probes to meet the demands. This dissertation is devoted to the development and application of novel chemical probes to facilitate the analysis of biomolecules such as glycans, peptides and proteins.

The first two chapters in this thesis center around the topic of glycomics (Chapter 2-3). Glycosylation is one of the most important post-translational modifications (PTMs) with essential physiological functions, including protein folding, cell signaling and immune response. Thus, various qualitative and quantitative glycomics analysis strategies have been developed. Recently, isobaric multiplex reagents for carbonyl containing compound (SUGAR) tags were developed for quantitative glycomics with multiplexing capacity and increased reporter ion yield. To further improve quantification efficiency and enable quantifying low abundance species, mass defect based triplex SUGAR (mdSUGAR) tag has been designed and introduced in Chapter 2. In addition, we also introduce additional reaction sites for mdSUGAR at terminal sialic acid by periodate oxidation of polyhydroxy chain to extend the mass difference and lower the requirement for resolving power. As a result, mdSUGAR tags show complete labeling efficiency, improved

fragmentation pattern and accurate quantification. Moreover, the quantitative performance of the mdSUGAR tags in complex system has been systematically evaluated and demonstrated reliable results. While extensive efforts have been devoted to the analysis of N-glycans, high throughput quantitative analysis of O-glycans is often overlooked and underexplored. This is partially due to the lack of universal enzyme for the release of O-glycans from protein backbone. Furthermore, the traditional chemical releasing method suffers from severe side reactions and involves tedious sample preparation procedures. Here, a multiplexed isobaric labeling method enabled by *N*, *N*-dimethyl leucine containing pyrazolone analog (DiLeuPMP) is introduced in Chapter 3. This method combines the release and labeling of O-glycans in a one-pot reaction and achieves accurate MS2-based relative quantification with the ability to process four samples at a time. The method has been applied to core-1 O-glycan standard and three glycoproteins first and the results demonstrated its validity. Then complex biological specimen, the human serum sample was analyzed. Overall, this method provides an effective and reliable approach for the profiling and high-throughput quantitative analysis of O-glycans in complex samples.

Chapter 4 discusses our development of a streamlined workflow to selectively enrich sialylglycopeptides. Despite the important roles of protein sialylation in biological processes such as cellular interaction and cancer progression, simple and effective methods for the analysis of intact sialylglycopeptides (SGPs) are still limited. Analyses of low-abundance SGPs typically require efficient enrichment prior to comprehensive liquid chromatography-mass spectrometry (LC-MS)-based analysis. Here, a novel workflow combining mild periodate oxidation, hydrazide chemistry, copper-catalyzed azide/alkyne cycloaddition (CuAAC) click chemistry, and dynamic covalent exchange, has been developed for selective enrichment of SGPs. The intact SGPs could be separated easily from protein tryptic digests and the signature ions were produced during LC-

MS/MS for unambiguous identification. The structure of the signature ions and corresponding dynamic covalent exchange was confirmed by using isotopic reagent. Under the optimized condition, over 70% enrichment efficiency of SGPs was achieved for bovine fetuin digests, and the method was successfully applied to complex biological samples, such as a mouse lung tissue extract. The high enrichment efficiency, good reproducibility, and easily adopted procedure without the need to generate specialized materials make this method a promising tool for broad applications to SGPs analysis.

Chapter 5 shifts focus from qualitative analysis of sialylglycopeptides to quantitative analysis. To characterize intact glycopeptide, mass spectrometry (MS)-based glycoproteomics have employed versatile fragmentation methods, among which Electron-Transfer/Higher-Energy Collision Dissociation (EThcD) has gained great popularity. However, the inherent limitation of EThcD in fragmenting low-charge ions has prevented its wide applications. Furthermore, it is in need to develop a high-throughput strategy for comparative glycoproteomics in a large cohort of samples. Herein, we developed isobaric *N*,*N*-dimethyl Leucine derivatized ethylenediamine (DiLeuEN) tags to increase the charge states of glycopeptides, thereby improving the fragmentation efficiency and allowing for in-depth intact glycopeptide analysis, especially sialoglycopeptides. Moreover, the unique reporter ions of DiLeuEN-labeled glycopeptides generated in tandem MS spectra enable relative quantification of up to four samples in a single analysis, which represents a new high-throughput method for quantitative glycoproteomics.

A novel chemical probe named cleavable DiLeu-Biotin-Azide (cDBA) tag was introduced in Chapter 6 and a versatile chemical proteomics platform based on the cDBA tag is established. The idea of this project stems from extending the analytical capability of traditional isotopic tandem orthogonal proteolysis–activity-based protein profiling (isoTOP-ABPP) strategy. By

incorporating isobaric labeling concept and our DiLeu tag design, several candidates of cDBA tag were designed and synthesized. After evaluating the performance of these candidates including labeling/enrichment efficiency and fragment pattern in tandem mass spectrometry (MS/MS), the best candidate was selected. The usage of cDBA tag was further explored due to the exquisite design of its chemical structure and unique features of biorthogonal reactions. By altering alkyne probe to chemical selectively label different PTM sites, various PTM could be derivatized, captured, enriched, and analyzed. The versatility of the application of cDBA tag makes it an ideal pan-PTM high throughput quantitative analysis platform. Furthermore, the combination of different PTM constitutes the next level of complexity for proteomics research and this PTM crosstalk research field could also benefit from our pan-PTM proteomics platform.

Chapter 7 summarizes results and key findings from each previous chapter and presents future directions of my work. Through extensive method development and applications of various novel chemical probes that integrates various sample preparation strategies and LC-MS methods, we explored the underlying mechanisms in several physiological or pathological systems. It is my hope that these method development work can be continued, and they can be of help in future biological research.

References:

- **Li, M.**; Huang, J.; Ma, M.; Shi, X.; Li, L. Selective Enrichment of Sialylglycopeptides Enabled by Click Chemistry and Schiff-base Covalent Exchange. *Anal. Chem.* **2022**, 94, 18, 6681–6688.
- **Li, M.**[#]; Gu, T.[#]; Lin, X.; Li, L. DiLeuPMP: a multiplexed isobaric labeling method for quantitative analysis of O-glycans. *Anal. Chem.* **2021**, 93, 28, 9845–9852.
- Feng, Y.[#]; **Li, M.**[#]; Lin, Y.; Chen, B.; Li, L. Multiplex Quantitative Glycomics Enabled by Periodate Oxidation and Triplex Mass Defect Isobaric Multiplex Reagents for Carbonyl Containing Compound Tags (mdSUGAR). *Anal. Chem.* **2019**. 91, 18, 11932-11937.
- **Li, M.**[#]; Zhong, X.[#]; Feng, Y.; Li, L. Novel Isobaric Tagging Reagent Enabled Multiplex Quantitative Glycoproteomics via Electron-Transfer/Higher-Energy Collisional Dissociation (EThcD) Mass Spectrometry. *Ready for submission.*
- Feng, Y.[#]; **Li, M.**[#]; Li, L. 12-plex SUGAR tag enabled pancreatic cancers cells glycomics. *Ready for submission.*
- **Li, M.**[#]; Ma, M.[#]; Li, L. Cleavable DiLeu-Biotin-Azide (cDBA) Tag Enabled isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP) Chemical Proteomics Platform for High-throughput Quantitative pan-PTM Analysis. *Manuscript in preparation.*
- **Li, M.**[#]; Ma, M.[#]; Li, L. DiLeuC enabled complementary ion based quantitative proteomics. *Manuscript in preparation.*
- Ma, M. [#]; **Li, M.**[#]; Wang, Z.; Li, L. 6-plex mdSUGAR tag enabled single-cell glycomics. *Manuscript in preparation.*

Chapter 2

Multiplex Quantitative Glycomics Enabled by Periodate Oxidation and Triplex Mass Defect Isobaric Multiplex Reagents for Carbonyl Containing Compound Tags (mdSUGAR)

Adapted from: Yu Feng[#], **Miyang Li[#]**, Yuanyuan Lin, Bingming Chen, Lingjun Li. Multiplex Quantitative Glycomics Enabled by Periodate Oxidation and Triplex Mass Defect Isobaric Multiplex Reagents for Carbonyl Containing Compound Tags *Anal. Chem.* **2019**, *91*, 11932−11937. (Author contribution: study was designed by Y. Feng, L. Li; experiment was performed by Y. Feng, M. Li; data was analyzed by Y. Feng, M. Li, Y. Lin; manuscript was written by Y. Feng and edited by M. Li, Y. Lin, B. Chen, L. Li.)

Abstract

Glycosylation is one of the most important post-translational modifications (PTMs) with essential physiological functions, including protein folding, cell signaling and immune response. Thus, various qualitative and quantitative glycomics analysis strategies have been developed. Recently, isobaric multiplex reagents for carbonyl containing compound (SUGAR) tags were developed for quantitative glycomics with multiplexing capacity and increased reporter ion yield. To further improve quantification efficiency and enable quantifying low abundance species, mass defect based triplex SUGAR (mdSUGAR) tag has been designed. In addition, we also introduce additional reaction sites for mdSUGAR at terminal sialic acid by periodate oxidation of polyhydroxy chain to extend the mass difference and lower the requirement for resolving power. As a result, mdSUGAR tags show complete labeling efficiency, improved fragmentation pattern and accurate quantification. Moreover, the quantitative performance of the mdSUGAR tags in complex system has been systematically evaluated and demonstrated reliable results.

Introduction

Glycosylation is a prevalent post-translational modification (PTM) with essential structural and biological functions including protein folding, cell signaling and immunity response.[1] Abnormal glycosylation could lead to multiple diseases, such as immunological disorders, cardiovascular diseases and various cancers.[2] Due to important functions of glycans, precise quantification approach is essential to understand complex biological processes. As glycan molecules cannot be easily detected by fluorescence or mass spectrometer due to lack of chromophore or poor ionization efficiency, derivatization is crucial for glycan analysis. Over the past few decades, various qualitative and quantitative glycomics analysis strategies have been

developed. In addition to fluorescence labeling, such as 2-aminobenzoic acid (2-AA) and 2-aminobenzamide (2-AB),[3] using stable isotope labeling method to improve the detection by mass spectrometry (MS) has become popular as it has the potential for highly multiplexed analysis with glycan structure elucidation capability.

The quantitative analysis with isotope labeling strategy relies on either reporter ion generated upon tandem MS fragmentation or peak area in full MS spectra. Isobaric tags, for example, aminoxyTMT,[4] iART,[5] QUANTITY[6] and SUGAR,[7] demonstrate the extensive sample multiplexing capacity for quantitative glycomics. To obtain quantitative result, the precursor needs to be isolated and fragmented to generate reporter ions. However, poor reporter ion yield was observed occasionally.[8] Moreover, the need for additional fragmentation further limits the quantification rate, especially for low abundance glycans using commonly employed data dependent acquisition strategy. Isotopic[9-13] and mass defect[14-15] labeling strategies rely on peak area for relative quantification. Data interpretation can be problematic for isotopic labeling due to increased spectral complexity. By contrast, the subtle mass difference between mass defect tags enables relative quantification by high resolving power MS without increasing the spectral complexity.

The isobaric multiplex reagents for carbonyl containing compound (SUGAR) tag has been developed for quantitative glycomics recently.[7] The simple building blocks of naturally occurring amino acids make it a cost-effective chemical probe with extensive multiplexing capacity. The tertiary amine structure in the SUGAR tag enhances ionization efficiency while stepwise reductive amination with hydrazide group enables complete labeling to further increase the detection sensitivity. With higher-energy collisional dissociation (HCD), reporter ions can be generated from SUGAR labeled glycans. Thus, the relative quantification can be obtained by comparing the

intensities of reporter ions. Although SUGAR tags significantly improve the performance of isobaric tags for glycan analysis, they still suffer from drawbacks associated with the isobaric tagging approach: quantification rate is limited due to the required MS^2 scans for quantification and low abundance glycan species tend to be difficult to quantify with lack of MS^2 scans. Here we report the development of mass defect SUGAR (*md*SUGAR) tags to eliminate the need for additional fragmentation for quantitative analysis, which is especially beneficial for quantitative analysis of low abundance glycans. Moreover, additional conjugation sites at sialic acid by periodate oxidation[16-18] on polyhydroxy chain further diminish instrument resolving power requirement and improve fragmentation patterns of labeled N-glycans. The excellent quantification performance with broad dynamic range offered by the *md*SUGAR tag has been demonstrated by labeling glycoprotein and complex human serum samples.

Materials and Methods

Materials

Acetic acid (AA), acetonitrile (ACN), dimethyl sulfoxide (DMSO), formic acid (FA), methanol (MeOH), and water were purchased from Fisher Scientific (Pittsburgh, PA). Sodium cyanoborohydride, sodium *meta*-periodate, triethylammonium bicarbonate buffer (TEAB, 1.0 M) and tris(2-carboxy-ethyl) phosphine hydrochloride (TCEP) were purchased from Sigma-Aldrich (St. Louis, MO). PNGase F was purchased from Promega (Madison, WI). Oasis HLB $1cm^3$ cartridges were purchased from Waters Corporation (Milford, MA). Bovine thyroglobulin (BTG) and human serum protein (HSP) were purchased from Thermo Fisher Scientific (Rockford, IL). Microcon-30kDa centrifugal filters (30K MWCO) were purchased from Merck Millipore Ltd. (Darmstadt, Germany). PolyGLYCOPLEX A^{TM} beads (3 µm) were purchased from PolyLC Inc.

(Columbia, MD). Fused silica capillary tubing (inner diameter 75 µm, outer diameter 375 µm) was purchased from Polymicro Technologies (Phoenix, AZ). All reagents were used without additional purification.

Enzymatic release of N-glycans by Filter-Aided N-Glycan Separation (FANGS)

Enzymatic release of N-glycans from protein samples was adapted from FANGS protocol.[19] Briefly, for N-glycan release without oxidation, protein samples were dissolved at a concentration of 5 µg/µL in 0.5 M TEAB buffer with 5 µL 0.5 M TCEP and heat-denatured by switching sample tubes between 95 °C and room temperature water baths every 15 seconds for 2 minutes. For N-glycan release with oxidation, protein samples were dissolved at a concentration of 5 µg/µL in 0.1 M sodium acetate buffer at pH 5.5 and heat-denatured. Then sodium *meta*-periodate was added to the final concentration of 1 mM. The solution was incubated at dark for 30 minutes at room temperature. For both conditions, 30 K MWCO filters were used for 3 times buffer exchange with 200 µL of 0.5 M TEAB. The prepared protein samples on the MWCO filter were then incubated with 5 µL PNGase F in 95 µL 0.5 M TEAB buffer at 37 °C overnight. The filters were washed with 100 µL 0.5 M TEAB buffer for three times. The fractions were combined and dried *in vacuo*. To convert glycosylamine to glycan with free reducing end, 200 µL of 1% AA was added and incubated for 4 h and dried *in vacuo*.

N-glycan labeling by mass-defect SUGAR tags

Released N-glycans were mixed with 1 mg *md*SUGAR tag in 100 µL MeOH containing 1% FA. The solvent was removed *in vacuo* after 10 min incubation. The reduction was performed in 100 µL 1 M $NaBH_3CN$ in DMSO: AA (7:3 v/v) at 70 ºC for 2 h.

Oasis HLB 1 cm^3 cartridge was used to remove excess *md*SUGAR tag and chemical reagents. The cartridge was conditioned with 1 mL 95% ACN, water, and 95% ACN respectively.

The reaction mixture was loaded to the cartridge with pre-filled 1 mL of 95% ACN. Then, the cartridge was washed with 1 mL of 95% ACN for 3 times, and the *md*SUGAR labeled N-glycans were eluted with 1 mL 50% ACN and 1 mL water. The eluting fractions were combined and dried *in vacuo*, reconstituted in 50 µL of 75% ACN, and analyzed by MALDI-MS or LC-MS/MS immediately.

Matrix-assisted laser desorption/ionization (MALDI)-MS analysis

Samples were prepared by premixing 1 µL of *md*SUGAR-labeled N-glycans with 1 µL 2,5-dihydroxybenzoic acid matrix (150 mg/mL in 2% N, N-dimethylaniline, 49% MeOH and 49% water), and 1 µL of each matrix/sample mixture was spotted onto the MALDI target plate. A MALDI-LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, Bremen, Germany) was used for MALDI-MS analysis. Ionization was performed using a laser energy of 15 µJ. Spectra were acquired in the Orbitrap mass analyzer within a mass range of *m/z* 1,000-4,000 at a mass resolution of 30 K (at *m/z* 400).

LC-MS/MS analysis

A self-fabricated nano-hydrophilic interaction chromatography (HILIC) column (15 cm, 75 µm i.d., 3 µm PolyGlycoPlex A HILIC beads) was used for glycan separation. A Dionex Ultimate 3000 nanoLC system was coupled to an Orbitrap Fusion Lumos Tribrid quadrupole-ion trap-Orbitrap Mass Spectrometer with a NanoSpray Flex ion source (Thermo Scientific, Bremen, Germany) for all LC-MS/MS analyses. Mobile phase A was water with 0.1% FA, and mobile phase B was ACN with 0.1% FA. The flow rate was set at 0.3 µL/min, and the injection volume was 2 µL. The following gradient was used (time, % mobile phase B) unless otherwise specified: (0 min, 75%), (18 min, 75%), (78 min, 15%), (78.1 min, 10%), (90 min, 10%), (90.1 min, 75%), (100 min, 75%).

The following mass spectrometer parameters were used for all data acquisition. Samples were ionized in positive ion mode with a spray voltage of 3 kV. S-lens radio frequency (RF) level was set to be 30, and capillary temperature was set to be 300 °C. Full MS scans were acquired at *m/z* 500-2000 with resolving power of 500 K (at *m*/*z* 200). Maximum injection time of 50 ms, automatic gain control (AGC) target value of 1e5, and 1 microscan were used for full MS scans. Top 5 data-dependent MS^2 analysis was performed at a resolving power of 30 K (at *m*/*z* 200) with collision-induced dissociation (CID) operating with a normalized collision energy of 30. The dynamic exclusion of acquired precursors was set to 15 sec with a ± 20 ppm tolerance.

N-glycan data analysis

The raw data was compared against an in-house database including most possible combinations of N-glycan units (Hexose (H), HexNAc (N), Fucose (F), and NeuAc (S)). *md*SUGAR-labeled N-glycans were identified by accurate mass matching in full MS with a mass tolerance of 10 ppm and fragmentation match in MS/MS spectra. Peak areas for *md*SUGAR-labeled glycans were used for relative quantification. Microsoft Excel and Origin were used for calculations and statistical analyses.

Results and Discussion

The structures of triplex *md*SUGAR tags are shown in **Figure 1** with three configurations (L, M and H) labeled with colored dots representing heavy isotope positions. In total, 12 heavy isotopes including ^{2}H, ^{13}C and ^{15}N, have been incorporated into triplex *md*SUGAR tags. The mass difference remains 23.8 mDa between adjacent channels. The synthetic route of *md*SUGAR is illustrated in **Figure S1**. The triplex *md*SUGAR tags were synthesized in three steps with commercially available heavy isotope coded starting materials respectively.

Due to subtle mass difference between mass defect chemical tags, ultrahigh resolving power is essential for mass defect based quantification strategy. The multiplexing capacity is often limited to duplex for quantitative glycomics with mass defect tags.[14-15] To increase the multiplexing capacity for mass defect strategy, extended mass difference can effectively reduce the resolving power requirement. However, with commonly used reductive amination derivatization, only reducing end can react with *md*SUGAR tag for conjugation. As larger glycans require higher resolving power for accurate quantification while they often contain one or multiple sialic acids at terminal position, periodate oxidation of polyhydroxy chain in sialic acid could be utilized to generate additional aldehyde for *md*SUGAR conjugation. With multiple reaction sites to extend mass difference, the triplex *md*SUGAR tags have been developed for quantitative glycomics.

To generate additional reaction sites with *md*SUGAR tag, aldehyde group could be readily introduced by mild periodate oxidation, which is known to selectively oxidize the polyhydroxy chain of sialic acid.[16-18] Then, reducing end of glycan and newly generated aldehyde group were labeled with hydrazide group of *md*SUGAR tag. To prevent reduction of aldehyde group prior to hydrazone formation with absence of high concentration of $NaBH_3CN$, stepwise reductive amination was developed. The glycans were mixed with *md*SUGAR tag in methanol containing 1% formic acid for 10 minutes to accelerate hydrazone formation. Then the reduction was performed in dimethyl sulfoxide and acetic acid with 1 M $NaBH_3CN$ to achieve complete irreversible conjugation. The workflow for *md*SUGAR labeling is shown in **Scheme S1**.

Periodate oxidation and stepwise reductive amination with *md*SUGAR tags were further applied to label N-glycans released from bovine thyroglobulin (BTG) glycoprotein standard. **Figure 2** depicts the workflow for quantitative glycomics with *md*SUGAR tags. Mild periodate

oxidation is performed prior to enzymatic release of N-glycans by PNGase F. With optimized stepwise reductive amination, N-glycans are labeled completely with *md*SUGAR tags at both reducing end and terminal sialic acid. The quantitative analysis can be achieved with high resolving power MS. **Figure 3A** highlighted multiple types of N-glycans labeled with *md*SUGAR including high mannose, fucosylated and sialylated N-glycans. The performance of periodate oxidation was also illustrated in **Figure 3B-E**. Almost identical spectra were observed for high-mannose and fucosylated N-glycans before and after periodate oxidation (**Figure 3B-C**) due to lack of polyhydroxy chain while 62 Da mass loss was observed (**Figure 3D-E**) for N-glycans containing single sialic acid (H_5N_4FS and H_6N_4FS) and 124 Da mass loss for $H_5N_4FS_2$.

In order to maximize the separation of *md*SUGAR labeled glycans with minimum ion coalescence, the resolving power and automatic gain control (AGC) were carefully optimized. Triplex *md*SUGAR labeled glycan samples with 1:1:1 ratio were evaluated with resolving power at 120K, 240K and 500K (at *m/z* 200). At 120K, the triplex peaks were unresolved, while at 240K, most peaks were partially resolved. At 500K, baseline separation was observed for most glycans but not for larger glycans such as $H_5N_4FS_2$ with only reducing end labeling (**Figure S2**). With periodate oxidation, additional reaction sites were introduced to label *md*SUGAR for extending mass difference, all triplex peak groups were baseline resolved at 500K (**Figures S3-4**). Space-charge-induced ion coalescence occurs with excessively large ion populations accumulating in the orbitrap mass analyzer.[20] The AGC target was evaluated to minimize ion coalescence. AGC of 1e5, 2e5, 5e5 and 1e6 were tested at a resolving power of 500K. No coalescence was observed at AGC of 1e5 while peaks started to merge together at AGC of 2e5 and unresolved at AGC of 1e6 (**Figure S5**). Thus, resolving power of 500K with AGC target of 1e5 was used for subsequent experiments.

Although MS/MS fragmentation is not required for relative quantification with mass defect based strategy, it is essential for N-glycan structural elucidation. The CID-MS/MS spectrum in **Figure 4A** demonstrated the fragmentation pattern for H_5N_4FS labeled with *md*SUGAR at reducing end. Due to higher proton affinity of tertiary amine in *md*SUGAR structure, labeled N-glycans tend to generate *y* ions with *md*SUGAR attached to the fragments. With additional labeling at terminal sialic acid, significant improvement of *b* ion (with red dots) production was observed from precursor glycan shown in **Figure 4B** which facilitated confident structural annotation.

Performance of quantitative analysis with triplex *md*SUGAR tags was evaluated by labeling N-glycans at known ratios. N-glycans released from BTG were aliquoted with 1:1:1, 1:5:10 and 10:5:1 ratios in triplicate and labeled with triplex *md*SUGAR tags respectively. Peak areas for each N-glycan were used to calculate the experimental ratios. **Figure 5A** illustrated the experimental ratios with triplex *md*SUGAR against theoretical ratio of 1:5:10. Representative extracted ion chromatograms are shown in **Figures 5B-C** for two types of N-glycans. The representative spectra for 1:1:1 are shown in **Figures S3-4**. For quantitative analysis with known ratios, less than 20% relative errors were observed, demonstrating the accurate quantification performance with *md*SUGAR tags. In addition, no obvious LC retention time shift was noticed for triplex *md*SUGAR labeled N-glycans with HILIC column separation.

To further evaluate whether additional periodate oxidation would cause quantification bias, N-glycans released from the same amount of BTG and human serum protein (HSP) were labeled with *md*SUGAR via either reducing end only labeling or reducing end and sialic acid labeling. With limited resolving power of 500K available, larger N-glycans labeled with triplex *md*SUGAR at only reducing end were unable to be fully separated (**Figure S2**), leading to inaccurate quantification. Thus, only duplex *md*SUGAR-L and H were used for the evaluation. The N-glycan

ratios of BTG/HSP from different labeling strategies were plotted in **Figure 6**. In total, 53 N-glycans were identified and quantified with reducing end strategy while 95 N-glycans were detected with the periodate oxidation strategy. The full list of quantified N-glycans by different strategies were summarized in **Table S1**. Most N-glycans showed significant abundance in BTG. However, several N-glycans including H_4N_4F and H_4N_5F were present at higher abundance in HSP. Although some N-glycans, for example H_7N_2 and H_5N_4FS, showed slight variation using different labeling strategies which might be due to the altered efficiency of enzymatic digestion caused by incomplete denaturation and disulfide bond reduction. Most N-glycans exhibited comparable quantification results with both labeling strategies, demonstrating the excellent quantification performance by incorporating additional periodate oxidation with *md*SUGAR tags.

Conclusions

In summary, triplex *md*SUGAR tags with periodate oxidation were developed in this study for quantitative glycomics. Periodate oxidation at polyhydroxy chain provided additional conjugation sites to extend mass difference and minimize instrument requirement. Stepwise reductive amination was utilized to enable complete labeling with hydrazide group of *md*SUGAR tags. The *md*SUGAR labeled glycans demonstrated improved fragmentation pattern to generate abundant *b* ions for structural elucidation. Triplex *md*SUGAR tags also showed accurate quantification with broad dynamic range. Moreover, *md*SUGAR tags with periodate oxidation strategy was applied to compare the abundance of N-glycans released from BTG and HSP. The successful development of *md*SUGAR approach offers a platform for N-glycan analysis to generate multiple conjugation sites, especially for mass defect labeling with limited resolving

power instrument. In conclusion, we anticipate that the triplex *md*SUGAR with periodate oxidation strategy can be widely applied in multiple biological and clinical studies.

Acknowledgements

This research was support for this research was provided in part by the NIH grants U01CA231081, R01 DK071801, R21 AG055377, RF1 AG052324, and a Robert Draper Technology Innovation Fund grant with funding provided by the Wisconsin Alumni Research Foundation (WARF). The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531) 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 Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

References

1. Varki, A.; Gagneux, P., Biological Functions of Glycans. In *Essentials of Glycobiology*, rd; Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor (NY), 2015; pp 77-88.

2. Varki, A.; Kannagi, R.; Toole, B.; Stanley, P., Glycosylation Changes in Cancer. In *Essentials of Glycobiology*, rd; Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor (NY), 2015; pp 597-609.

3. Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M., Glycan labeling strategies and their use in identification and quantification. *Anal Bioanal Chem* **2010,** *397* (8), 3457-81.

4. Hahne, H.; Neubert, P.; Kuhn, K.; Etienne, C.; Bomgarden, R.; Rogers, J. C.; Kuster, B., Carbonyl-reactive tandem mass tags for the proteome-wide quantification of N-linked glycans. *Anal Chem* **2012,** *84* (8), 3716-24.

5. Yang, S.; Yuan, W.; Yang, W.; Zhou, J.; Harlan, R.; Edwards, J.; Li, S.; Zhang, H., Glycan analysis by isobaric aldehyde reactive tags and mass spectrometry. *Anal Chem* **2013,** *85* (17), 8188-95.

6. Yang, S.; Wang, M.; Chen, L.; Yin, B.; Song, G.; Turko, I. V.; Phinney, K. W.; Betenbaugh, M. J.; Zhang, H.; Li, S., QUANTITY: An Isobaric Tag for Quantitative Glycomics. *Sci Rep* **2015,** *5*, 17585.

7. Feng, Y.; Chen, B.; Yu, Q.; Zhong, X.; Frost, D. C.; Ikonomidou, C.; Li, L., Isobaric Multiplex Labeling Reagents for Carbonyl-Containing Compound (SUGAR) Tags: A Probe for Quantitative Glycomic Analysis. *Anal Chem* **2019,** *91* (4), 3141-3146.

8. Chen, B.; Zhong, X.; Feng, Y.; Snovida, S.; Xu, M.; Rogers, J.; Li, L., Targeted MultiNotch MS(3) Approach for Relative Quantification of N-Glycans Using Multiplexed Carbonyl-Reactive Isobaric Tags. *Anal Chem* **2018,** *90* (2), 1129-1135.

9. Walker, S. H.; Budhathoki-Uprety, J.; Novak, B. M.; Muddiman, D. C., Stable-isotope labeled hydrophobic hydrazide reagents for the relative quantification of N-linked glycans by electrospray ionization mass spectrometry. *Anal Chem* **2011,** *83* (17), 6738-45.

10. Walker, S. H.; Taylor, A. D.; Muddiman, D. C., Individuality Normalization when Labeling with Isotopic Glycan Hydrazide Tags (INLIGHT): a novel glycan-relative quantification strategy. *J Am Soc Mass Spectrom* **2013,** *24* (9), 1376-84.

11. Hecht, E. S.; McCord, J. P.; Muddiman, D. C., Definitive Screening Design Optimization of Mass Spectrometry Parameters for Sensitive Comparison of Filter and Solid Phase Extraction Purified, INLIGHT Plasma N-Glycans. *Anal Chem* **2015,** *87* (14), 7305-12.

12. Chen, Z.; Zhong, X.; Tie, C.; Chen, B.; Zhang, X.; Li, L., Development of a hydrophilic interaction liquid chromatography coupled with matrix-assisted laser desorption/ionization-mass spectrometric imaging platform for N-glycan relative quantitation using stable-isotope labeled hydrazide reagents. *Anal Bioanal Chem* **2017,** *409* (18), 4437-4447.

13. Wei, L.; Cai, Y.; Yang, L.; Zhang, Y.; Lu, H., Duplex Stable Isotope Labeling (DuSIL) for Simultaneous Quantitation and Distinction of Sialylated and Neutral N-Glycans by MALDI-MS. *Anal Chem* **2018,** *90* (17), 10442-10449.

14. Atwood, J. A., 3rd; Cheng, L.; Alvarez-Manilla, G.; Warren, N. L.; York, W. S.; Orlando, R., Quantitation by isobaric labeling: applications to glycomics. *J Proteome Res* **2008,** *7* (1), 367-74.

15. Chen, B.; Feng, Y.; Frost, D. C.; Zhong, X.; Buchberger, A. R.; Johnson, J.; Xu, M.; Kim, M.; Puccetti, D.; Diamond, C.; Ikonomidou, C.; Li, L., Quantitative Glycomic Analysis by Mass-Defect-Based Dimethyl Pyrimidinyl Ornithine (DiPyrO) Tags and High-Resolution Mass Spectrometry. *Anal Chem* **2018,** *90* (13), 7817-7823.

16. De Bank, P. A.; Kellam, B.; Kendall, D. A.; Shakesheff, K. M., Surface engineering of living myoblasts via selective periodate oxidation. *Biotechnol Bioeng* **2003,** *81* (7), 800-8.

17. Zeng, Y.; Ramya, T. N.; Dirksen, A.; Dawson, P. E.; Paulson, J. C., High-efficiency labeling of sialylated glycoproteins on living cells. *Nat Methods* **2009,** *6* (3), 207-9.

18. Barrientos, R. C.; Zhang, Q., Isobaric Labeling of Intact Gangliosides toward Multiplexed LC-MS/MS-Based Quantitative Analysis. *Anal Chem* **2018,** *90* (4), 2578-2586.

19. Abdul Rahman, S.; Bergstrom, E.; Watson, C. J.; Wilson, K. M.; Ashford, D. A.; Thomas, J. R.; Ungar, D.; Thomas-Oates, J. E., Filter-aided N-glycan separation (FANGS): a convenient sample preparation method for mass spectrometric N-glycan profiling. *J Proteome Res* **2014,** *13* (3), 1167-76.

20. Gorshkov, M. V.; Fornelli, L.; Tsybin, Y. O., Observation of ion coalescence in Orbitrap Fourier transform mass spectrometry. *Rapid Commun Mass Spectrom* **2012,** *26* (15), 1711-7.

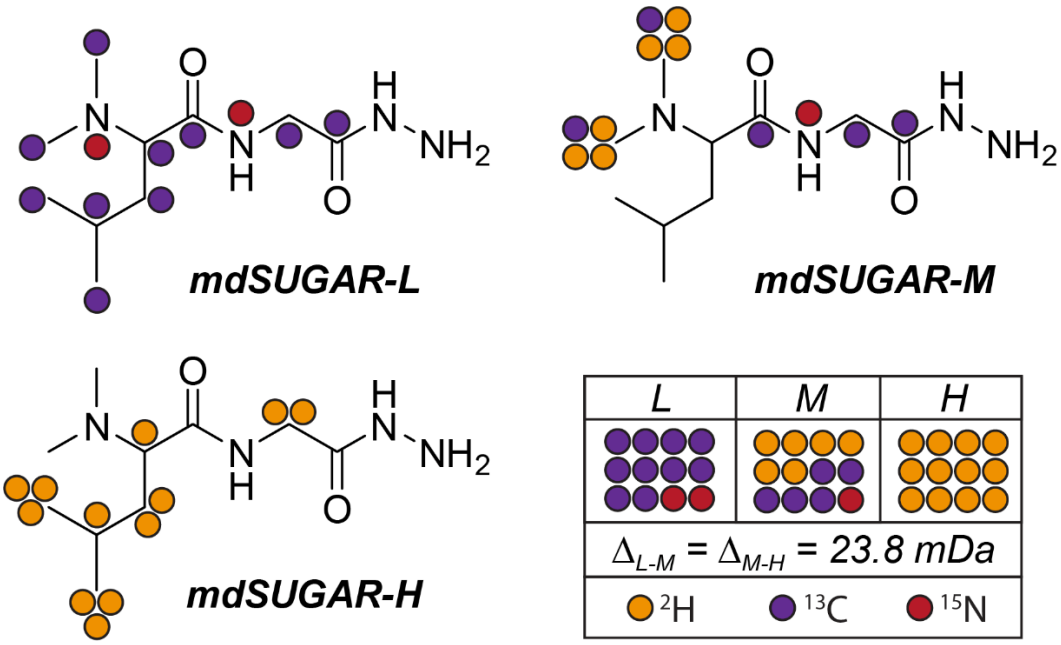


Figure 1. Structure and isotope configurations of triplex *md*SUGAR tags. Orange dot: ^{2}H, purple dot: ^{13}C, red dot: ^{15}N.

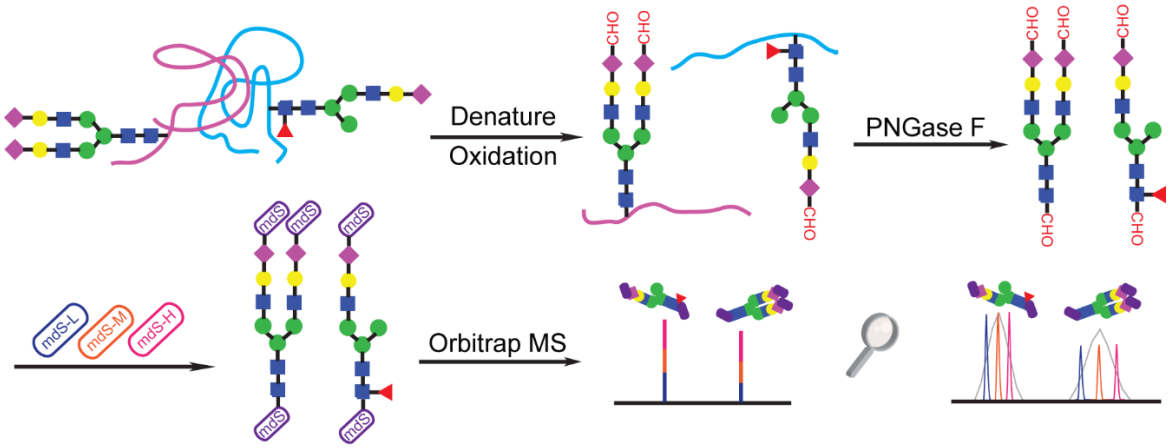


Figure 2. Workflow of quantitative glycomics analysis by triplex *md*SUGAR.

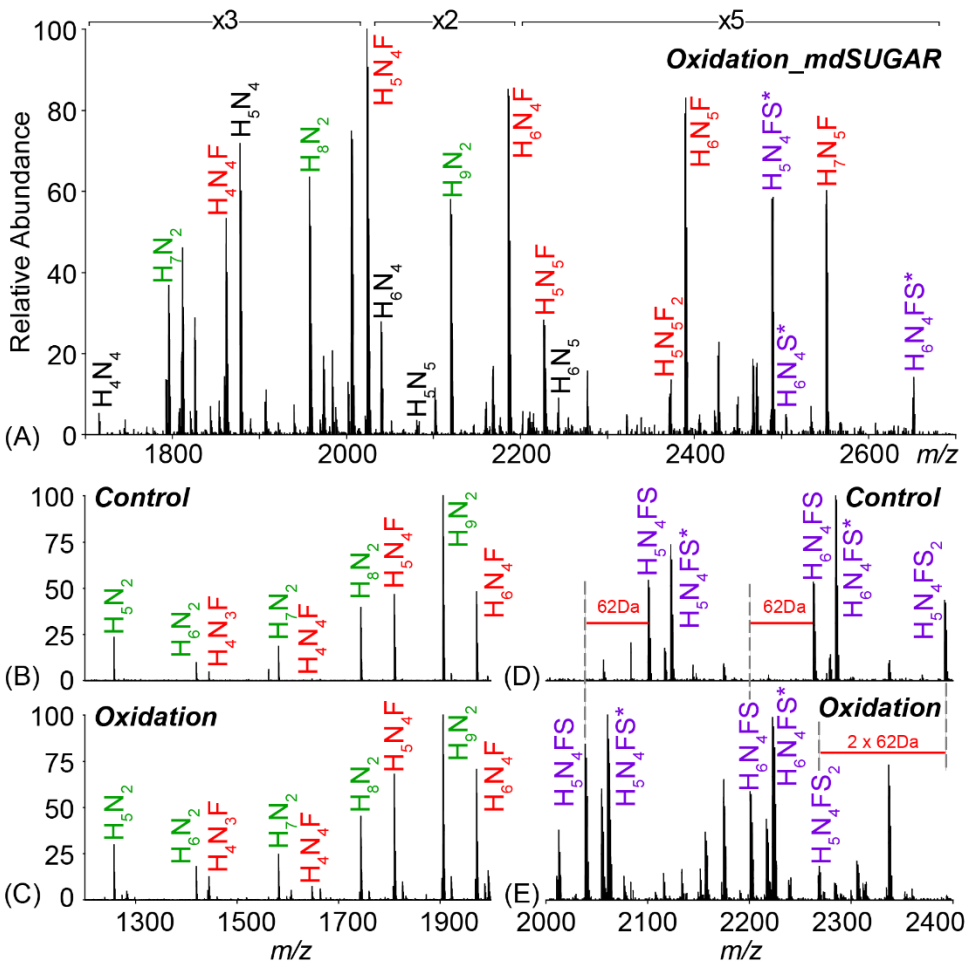


Figure 3. MALDI-MS spectra of N-glycans (Na+ adduct) released from BTG: mdSUGAR labeled with periodate oxidation (A). High mannose (green label) and fucosylated (red label) N-glycans showed almost identical spectral profile prior to (B) and after (C) periodate oxidation due to lack of sialic acid. Highlighted acidic N-glycans (purple labels) showed mass loss before (D) and after (E) periodate oxidation. Mass shift of 62 Da was illustrated for each sialic acid after periodate oxidation. Star represents N-glycans as [M-H+2Na]+.

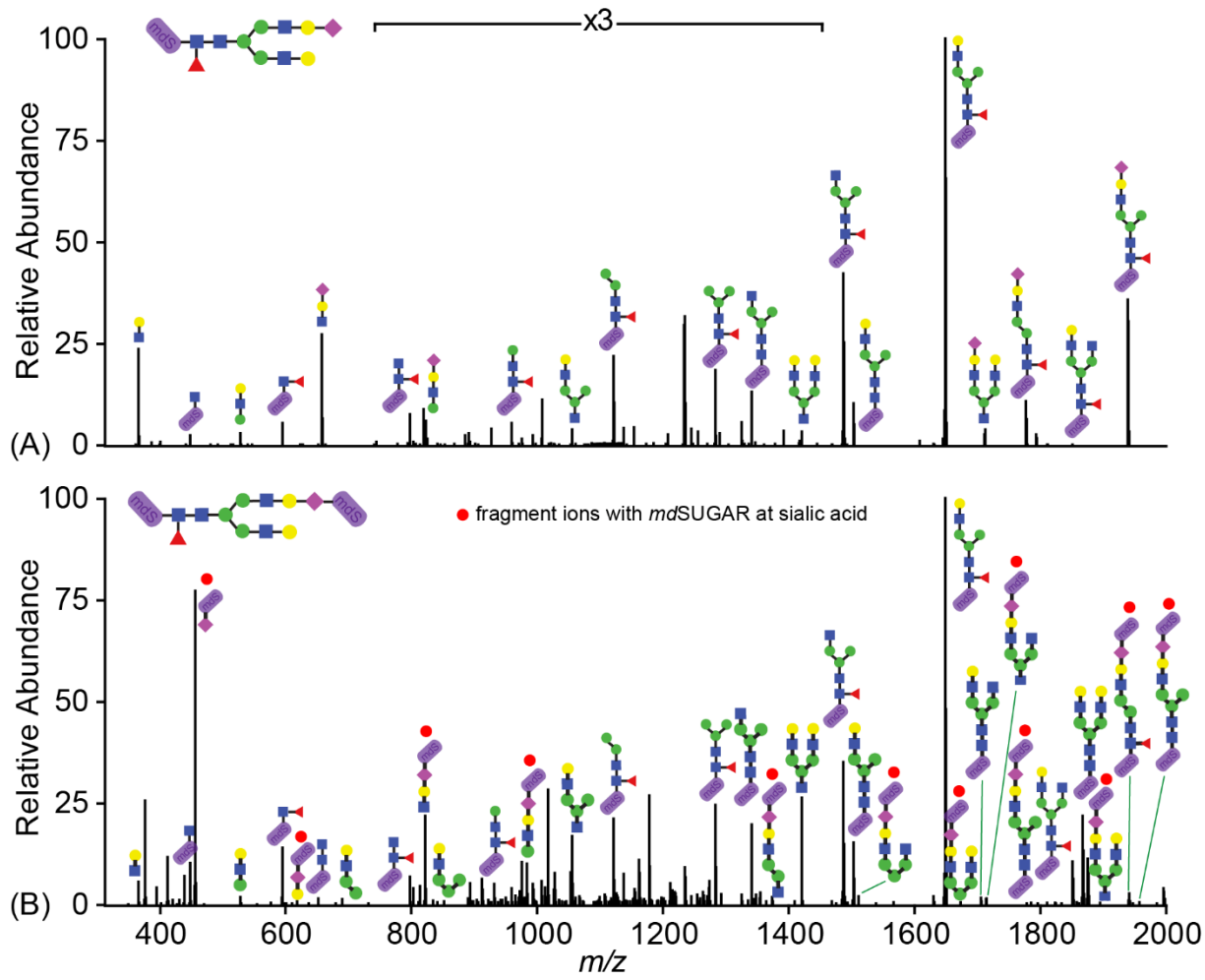


Figure 4. ESI-MS/MS fragmentation comparison of mdSUGAR-labeled H5N4FS N-glycan at only reducing end (A) or reducing end and terminal sialic acid (B). Fragments with red dots are fragment ions with mdSUGAR at sialic acid generated from labeled H5N4FS with periodate oxidation.

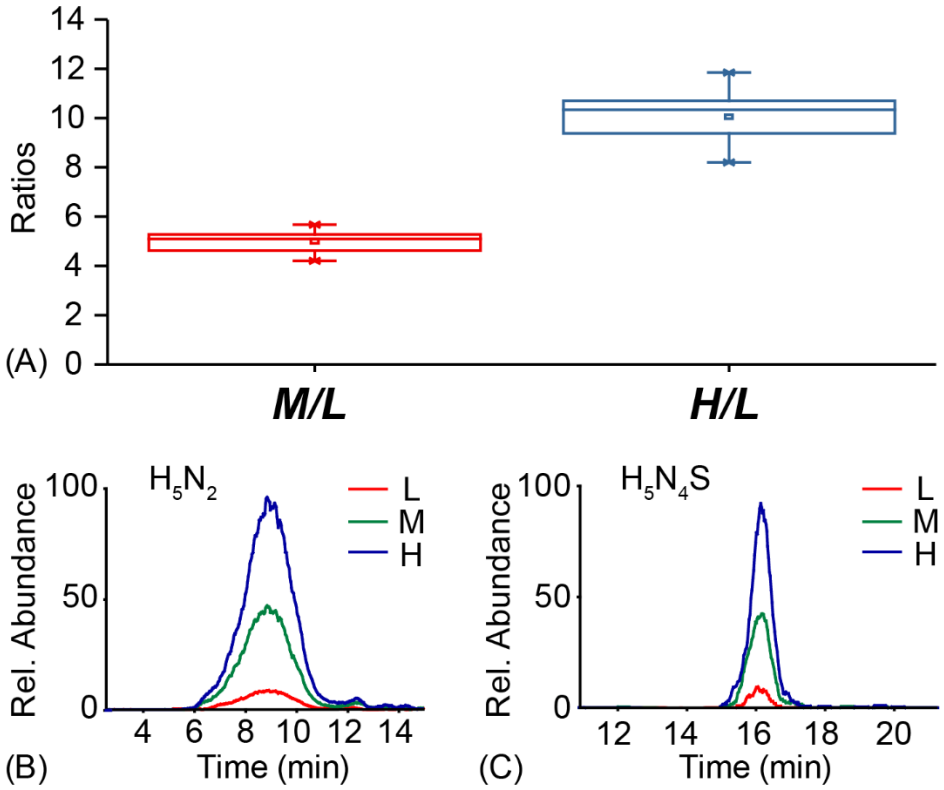


Figure 5. Relative quantification performance of triplex *md*SUGAR-labeled N-glycans released from BTG. Labeled N-glycans were mixed at ratios of 1:5:10 and analyzed in triplicate. The peak areas of M/L and H/L are plotted. Box plots show the median (line), the 25th and 75th percentile (box), and the 5th and 95th percentile (whiskers) (A). Representative extracted ion chromatograms for *md*SUGAR-labeled N-glycans H_5N_2 (B) and H_5N_4S (C) are shown.

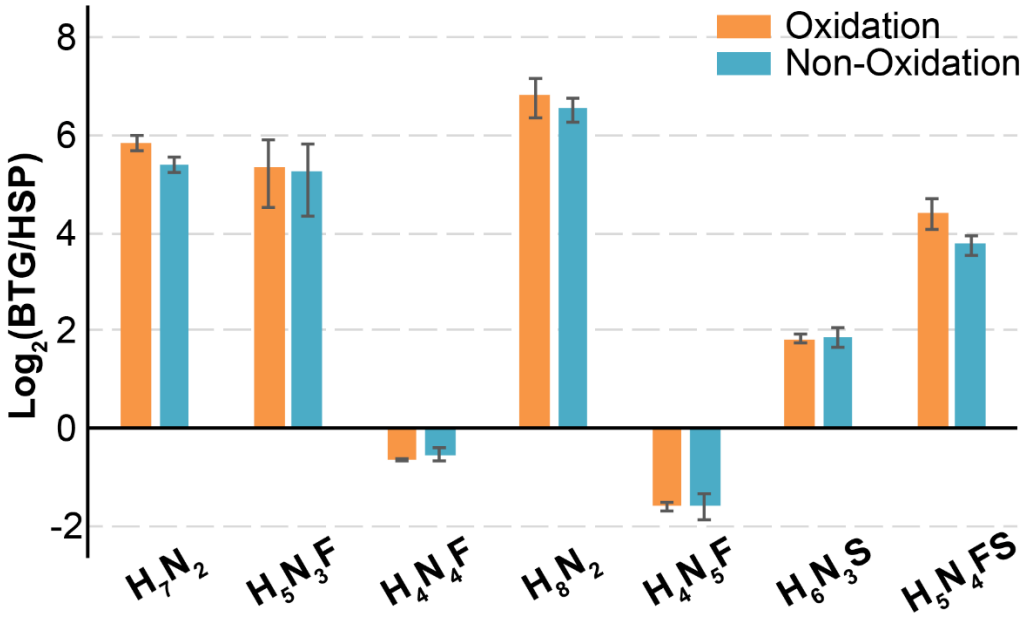
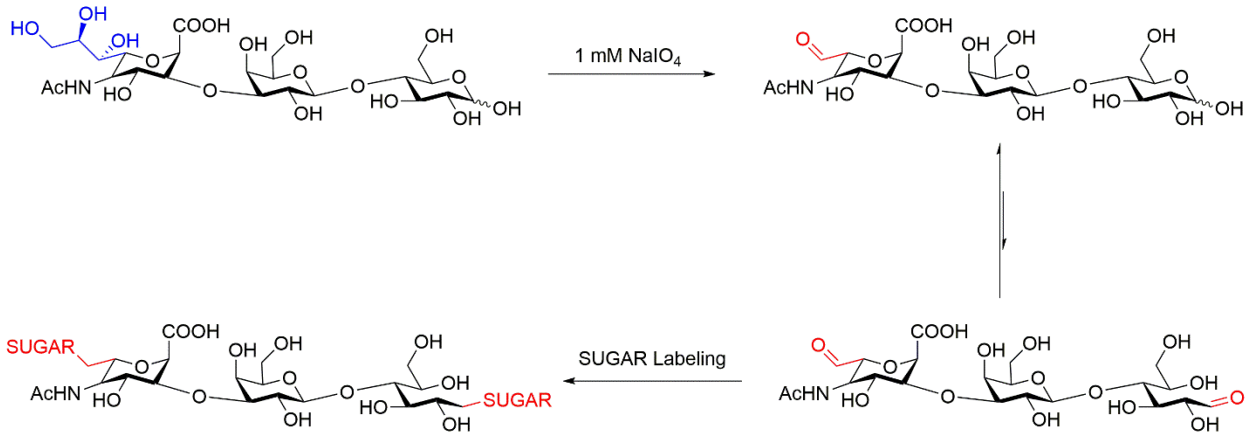


Figure 6. Selected N-glycan relative quantification released from equal amount of BTG and HSP. Ratios represent peak areas for *md*SUGAR-labeled N-glycans. Error bars represent the standard deviation of three replicates.

Supplemental Information



Scheme S1. Periodate oxidation and *md*SUGAR labeling for quantitative glycomics.

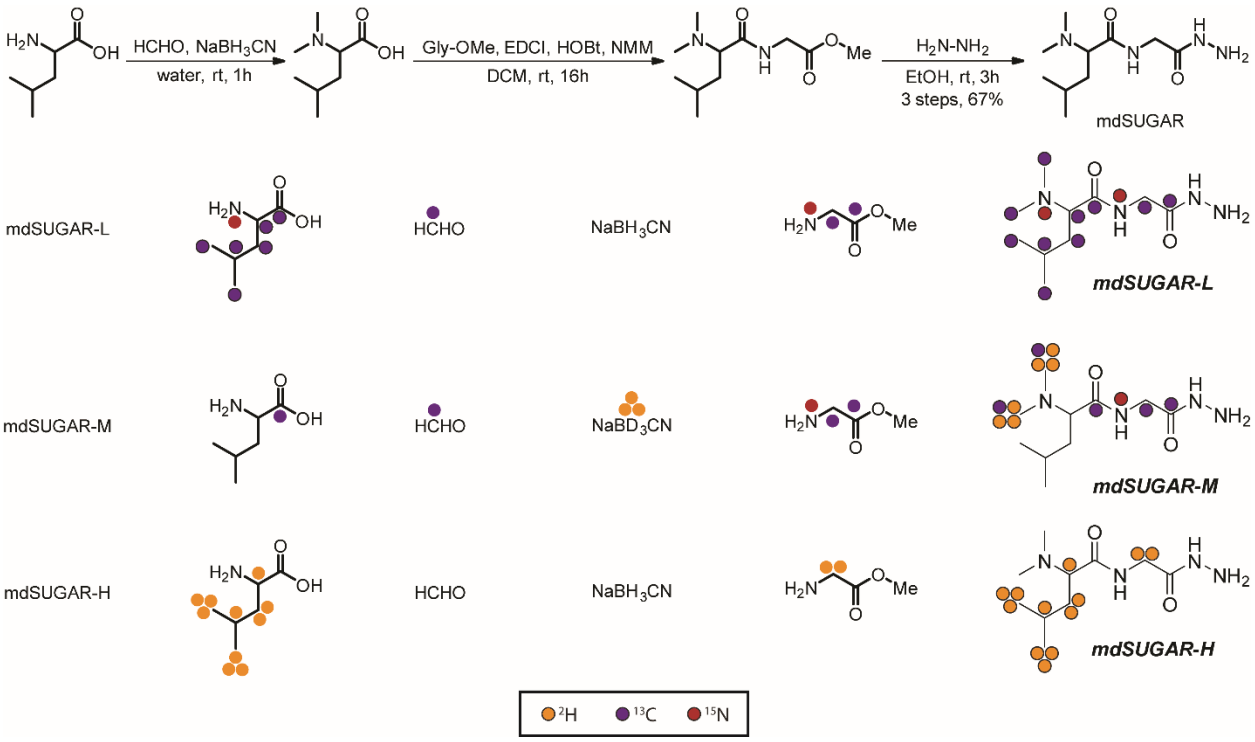


Figure S1. Synthetic route and starting materials for triplex *md*SUGAR tags.

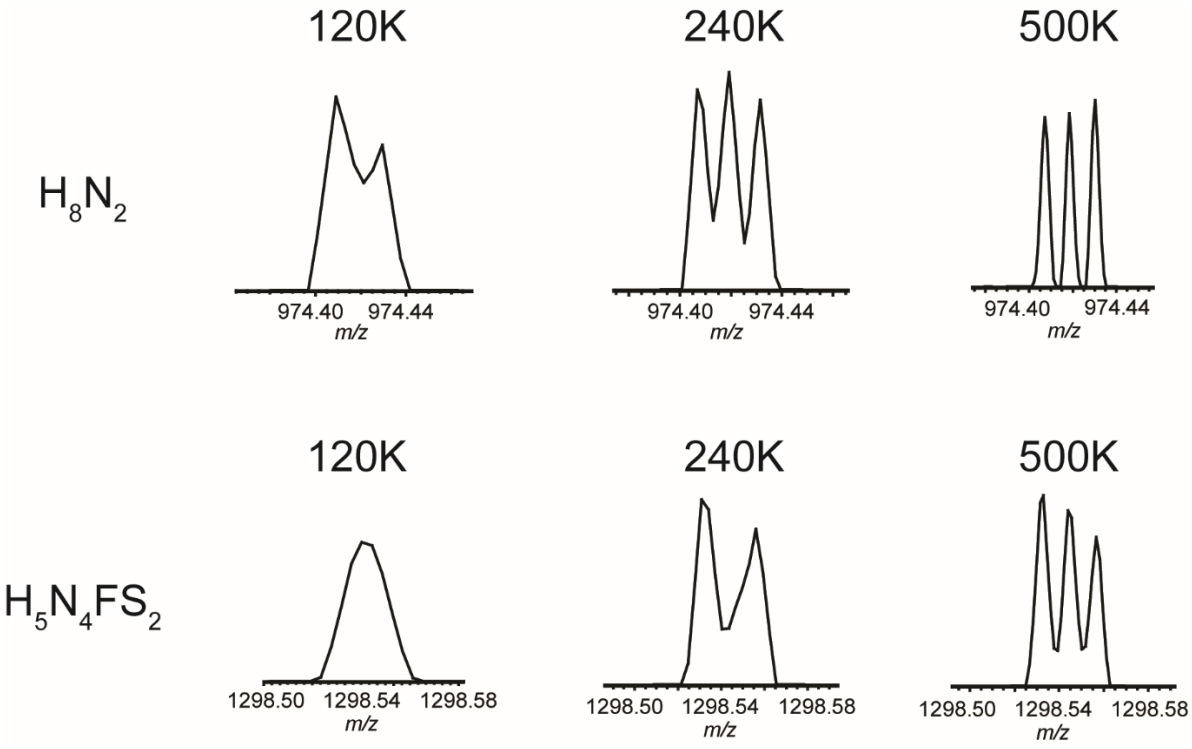


Figure S2a. Resolving power optimization with mdSUGAR tags labeled at reducing end only.

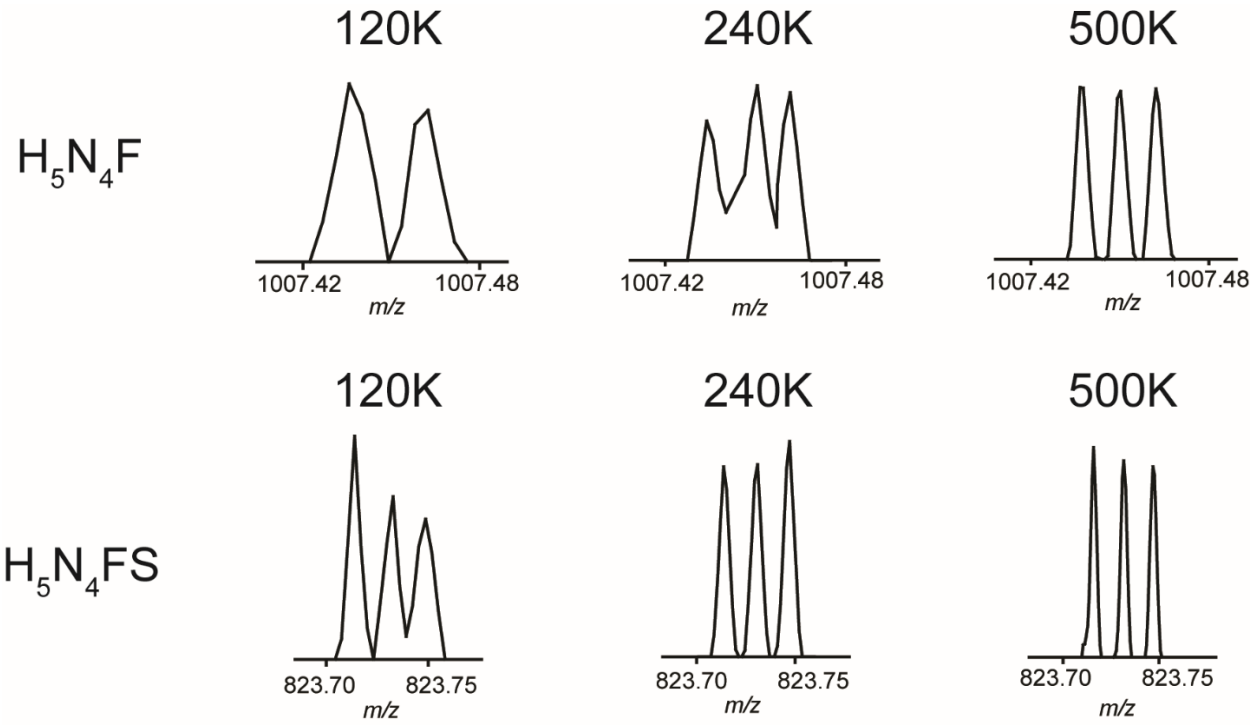


Figure S2b. Resolving power optimization with mdSUGAR tags labeled at reducing end and sialic acid after periodate oxidation.

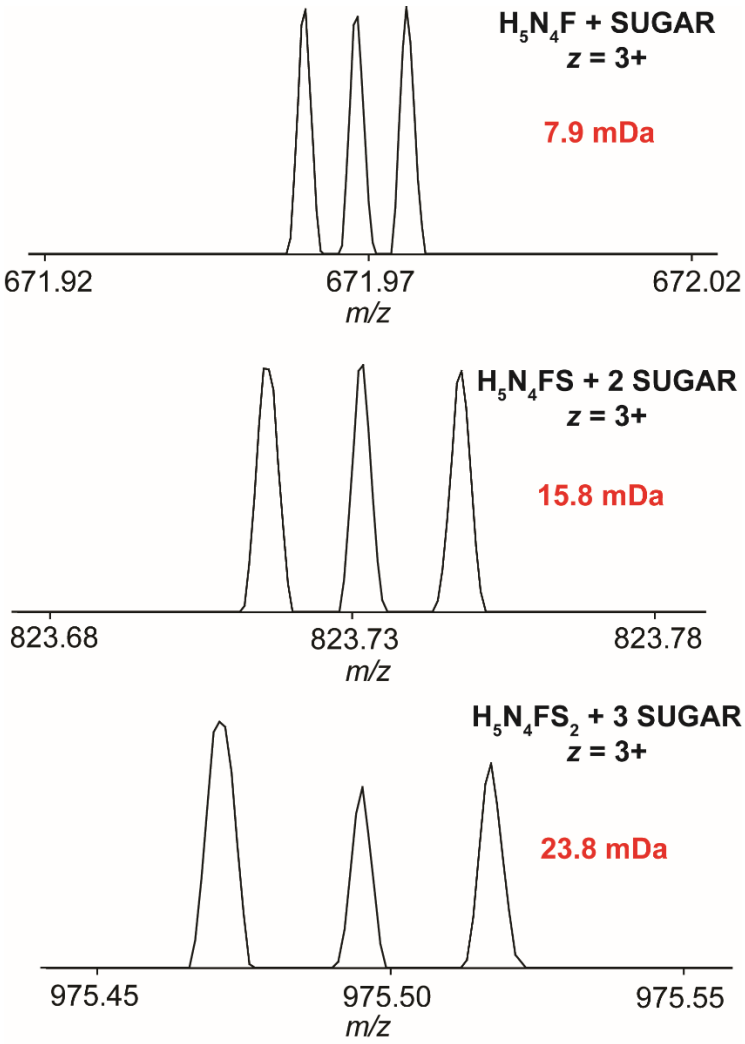


Figure S3. Extended mass difference was observed with *md*SUGAR labeling after periodate oxidation at resolving power of 500K.

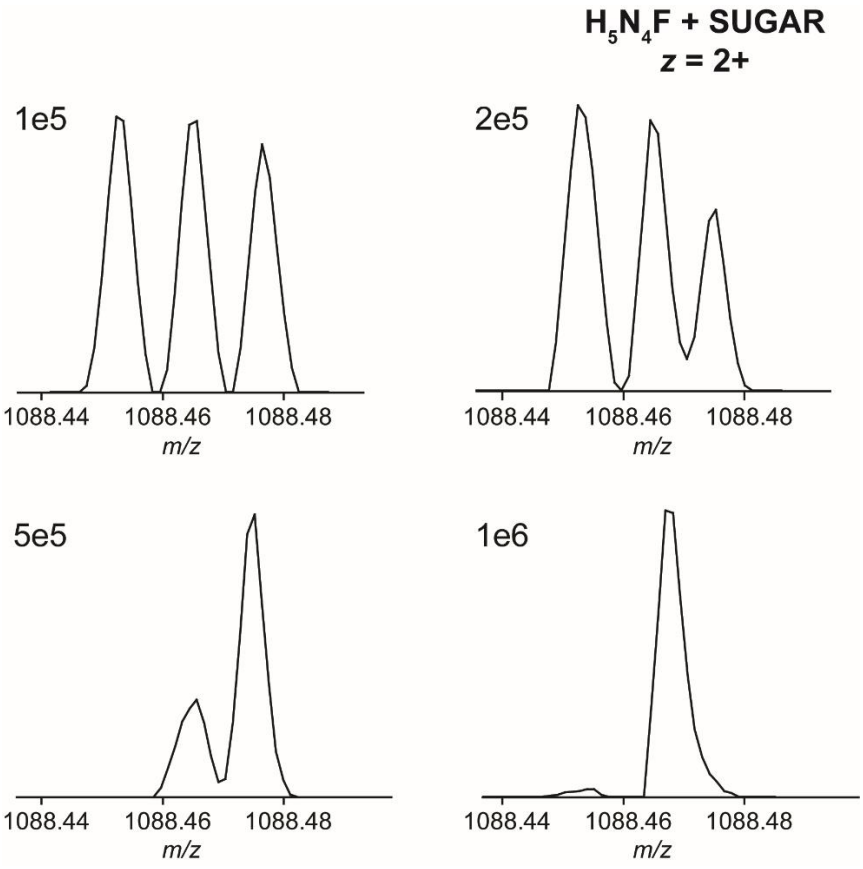


Figure S4. AGC target optimization at resolving power at 500K.

Table S1. Full list of quantified N-glycans by reducing end labeling or reducing end and sialic labeling with mdSUGAR tags.

	BTG_Re	BTG_Ox	HSP_Re	HSP_Ox
H_3N_2	√	√	√	√
H_3N_2F	√	—	—	—
H_4N_2	√	√	√	√
H_3N_3	√	√	√	√
H_4N_2F	√	√	√	√
H_5N_2	√	√	√	√
H_3N_3F	√	√	√	√
H_4N_3	√	√	√	√
H_3N_4	√	√	√	√
$H_4N_2F_2$	—	√	—	—
H_5N_2F	√	√	—	—
H_6N_2	√	√	√	√
H_4N_3F	√	√	√	√
H_5N_3	√	√	√	√
H_3N_4F	√	√	√	√
H_4N_4	√	√	√	√
$H_4N_2F_3$	—	√	—	—
H_3N_5	√	√	√	—
H_6N_2F	—	√	—	√
H_7N_2	√	√	√	√
H_3N_3S	√	√	√	√
H_5N_3F	√	√	√	√
H_6N_3	√	√	√	√
H_4N_4F	√	√	√	√
H_5N_4	√	√	√	√
H_3N_5F	√	√	√	√
H_4N_5	√	√	√	√
H_3N_3FS	√	√	√	—
H_8N_2	√	√	√	√
H_4N_3S	√	√	√	√
H_6N_3F	—	√	—	—
H_5N_4F	√	√	√	√
H_6N_4	√	√	√	√
H_4N_5F	√	√	√	√
H_5N_5	√	√	—	√
H_4N_3FS	√	√	√	—
H_9N_2	√	—	—	—
H_4N_6	—	√	—	—
H_5N_3S	√	√	√	—
$H_4N_4F_3$	—	√	—	—
$H_5N_4F_2$	—	√	—	—
H_4N_4S	√	√	√	—
H_6N_4F	√	—	—	—
$H_4N_5F_2$	—	√	—	—
H_5N_5F	√	√	—	√

$H_6N_3F_3$	—	√	—	√
H_5N_3FS	√	√	√	√
$H_{10}N_2$	—	√	—	—
H_6N_3S	√	√	√	√
H_3N_7F	—	√	—	√
$H_5N_4F_3$	—	√	—	√
H_4N_4FS	√	√	√	√
H_4N_7	—	√	—	√
H_5N_4S	√	√	√	√
$H_5N_5F_2$	—	√	—	√
H_4N_5S	√	√	√	√
H_6N_5F	—	√	—	√
H_8N_2S	—	√	—	√
H_6N_3FS	√	√	√	√
H_6N_6	—	√	—	√
$H_3N_7F_2$	—	√	—	√
$H_4N_4F_2S$	—	√	—	√
H_4N_7F	—	√	—	—
H_5N_4FS	√	√	√	√
H_6N_4S	√	√	√	√
$H_4N_5F_4$	—	√	—	√
$H_5N_5F_3$	—	√	—	√
H_4N_5FS	√	√	√	—
H_5N_5S	√	√	√	√
H_7N_5F	—	√	—	—
H_4N_6S	—	√	—	—
$H_5N_3S_2$	√	—	√	—
H_6N_6F	√	√	—	√
H_7N_6	—	√	—	√
$H_5N_4F_2S$	√	—	√	—
$H_3N_8F_2$	—	√	—	√
H_4N_8F	√	√	—	√
H_5N_5FS	√	√	√	√
$H_7N_5F_2$	—	√	—	√
H_6N_5S	√	√	√	√
$H_5N_6F_3$	—	√	—	√
H_7N_6F	—	√	—	—
H_8N_6	—	√	—	√
$H_5N_4F_3S$	—	√	—	—
$H_7N_4F_4$	—	√	—	—
$H_6N_4F_2$	—	√	—	√
H_5N_4S	—	√	—	√
H_6N_7F	—	√	—	√
$H_8N_4F_3$	—	√	—	√
H_7N_7	—	√	—	—
H_5N_8F	—	√	—	√
H_6N_5FS	—	√	—	√
H_7N_5S	—	√	—	—
H_6N_6S	—	√	—	√

$H_5N_4FS_2$	—	√	—	√
$H_6N_4S_2$	—	√	—	—
$H_5N_5S_2$	—	√	—	√
H_4N_8FS	—	√	—	√
$H_5N_5FS_2$	—	√	—	√
$H_6N_5FS_2$	—	√	—	√

Chapter 3

DiLeuPMP: a multiplexed isobaric labeling method for quantitative analysis of O-glycans

Adapted from: **Miyang Li**[#], Tingjia Gu[#], Xiaorong Lin, Lingjun Li. DiLeuPMP: a multiplexed isobaric labeling method for quantitative analysis of O-glycans. *Anal. Chem.* **2021**, 93, 9845−9852. (Author contribution: study was designed by M Li; experiment was performed by M. Li., T. Gu; data was analyzed by M, Li., X. Lin; manuscript was written by M. Li and edited by T, Gu., X. Lin, L. Li.)

Abstract

As one of the most important post-translational modifications, glycosylation plays a pivotal role in many essential physiological functions, including cell recognition, signaling, and immune response. Thus, various qualitative and quantitative analytical strategies for glycomic profiling have been developed in recent decades. However, while extensive efforts have been devoted to the analysis of N-glycans, high throughput quantitative analysis of O-glycans is often overlooked and underexplored. This is partially due to the lack of universal enzyme for the release of O-glycans from protein backbone. Furthermore, the traditional chemical releasing method suffers from severe side reactions and involves tedious sample preparation procedures. Here, a multiplexed isobaric labeling method enabled by *N*, *N*-dimethyl leucine containing pyrazolone analog (DiLeuPMP) is introduced. This method combines the release and labeling of O-glycans in a one-pot reaction and achieves accurate MS^2-based relative quantification with the ability to process four samples at a time. The method has been applied on core-1 O-glycan standard and three glycoproteins first and the results demonstrated its validity. Then complex biological specimen, the human serum sample was analyzed. Overall, this method provides an effective and reliable approach for the profiling and high-throughput quantitative analysis of O-glycans in complex samples.

Introduction

Protein glycosylation is a pivotal post-translational modification (PTM) and plays essential roles in many biological processes such as immune response, receptor recognition, cellular communication, and embryonic development[1-3]. The important biological functions and ubiquitous existence of glycosylated proteins in nature have made pressing needs for developing

simple and sensitive platforms capable of analyzing glycans. Typically, glycosylation is classified into two types, N-linked and O-linked glycosylation, based on the attachment between oligosaccharides and amino acid residues. N-glycans conjugate to asparagine (N) residues in the consensus peptide motif of Asn-X-Ser/Thr (where X is any amino acid except proline) with a core structure (GlcNAc2Man3), while O-glycans are conjugated to serine (S) or threonine (T) without a clear consensus motif and consistent core structure.

For a typical glycomics analysis, cleavage of glycans from the peptide backbone is commonly the first step. For N-glycans, they can be readily released by enzymes such as PNGase F and PNGase A[4]. In contrast, the release and recovery of O-glycans have remained a very challenging problem. Due to lack of a common core structure, O-glycans are typically categorized into eight subclasses (core-1 to core-8)[5]. Although endo-α-N-acetylgalactosaminidase (O-glycanase) is reported to be able to cleave the core-1 type O-glycans (Gal-GalNAc), there is no universal O-glycosidase that enables removal of all O-glycans. Thus, the chemical method so-called β-elimination, is often employed for releasing O-glycans in practice. β-elimination is performed under alkali conditions, and the strong base removes the β-proton on amino acid residue where an elimination reaction is triggered making for an unsaturated amino acid (from serine or threonine residues) together with the released O-glycans. However, this process can proceed stepwise and remove one saccharide residue at a time from the polysaccharide backbone, which results in the degradation of released glycans (**Figure S1**). This side reaction is known as 'peeling'[6-8]. Previously, scientists found that adding blocking reagent 1-phenyl-3-methyl-5-pyrazolone (PMP) during β-elimination can suppress side reactions since PMP derivatization of the released O-glycans proceeds faster than peeling degradation and thus is able to competitively inhibit peeling process[9-11]. Mass difference-based isotopic labeling strategy using light and heavy

isotopic PMP reagent was also reported to achieve quantitative glycomics analysis[11, 12]. However, the multiplexing capacity was limited to dual-plex due to the increase in spectral complexity and limited commercially available isotopic PMP reagents.

Mass spectrometry (MS) has emerged as a prominent analytical tool for the O-glycomics analysis[13, 14]. Currently, developed methods have largely improved the recovery and facilitated the profiling of O-glycans, but most of them suffer from low throughput and limited capacity for quantitative analyses. Furthermore, the procedures are usually time-consuming and tedious. Isobaric labeling is a popular approach in MS to achieve high throughput analysis based on MS^2 fragmentation and reporter ion quantification. Our laboratory has previously developed a set of custom isobaric labeling reagents, *N*, *N*-dimethyl leucine (DiLeu), for large scale quantitative proteomics and peptidomics[15]. Very recently, the general architecture of DiLeu was borrowed and extended for the development of a novel set of isobaric multiplex reagent for carbonyl-containing compound (SUGAR) tags capable of quantitative N-glycomics[4]. The SUGAR tags employed reductive amination chemistry on the reducing end of released N-glycans, thus are not applicable for direct analysis of O-glycans. Herein, we combined the isobaric labeling strategy employed by DiLeu with PMP-aided β-elimination to develop a novel chemical tool, 4-plex *N*, *N*-dimethyl leucine containing pyrazolone analog (DiLeuPMP), which provides the first isobaric tag aiming for high throughput quantitative O-glycan analysis and facilitates the application of O-glycomics.

Materials and Methods

Materials and Reagents

Acetic acid (AA), acetonitrile (ACN), *N*, *N*-dimethylformamide (DMF), tetrahydrofuran (THF), carbonyldiimidazole (CDI), tris-(2-carboxyethyl)phosphine (TCEP), formic acid (FA), methanol (MeOH), chloroform ($CHCl_3$), 20% ammonium hydroxide (NH_4OH) and water (HPLC grade) were purchased from Fisher Scientific (Pittsburgh, PA). 2-(4-aminophenyl)-5-methyl-2,4-dihydro-3H-pyrazole-3-one was purchased from Matrix Scientific (Columbia, SC). Core-1 O-glycan standard, bovine fetuin, porcine stomach mucin (PSM) and bovine submaxillary mucin (BSM) were purchased from Sigma-Aldrich (St. Louis, MO). Single donor healthy human serum was purchased from Innovative Research Inc. (Novi, MI). PNGase F was purchased from Promega (Madison, WI). Microcon-10kDa centrifugal filters (10K MWCO) were purchased from Merck Millipore Ltd. (Darmstadt, Germany). Sep-Pak C18 Cartridges were purchased from Waters Corporation (Milford, MA). Pierce™ C18 Tips, 10 ul bed were purchased from Thermo Fisher Scientific (Waltham, MA). Ethylene Bridged Hybrid C18 packing material (1.7 µm) was purchased from PolyLC Inc. (Columbia, MD). Fused silica capillary tubing (inner diameter 75 µm, outer diameter 375 µm) was purchased from Polymicro Technologies (Phoenix, AZ). All reagents were used without additional purification.

Synthesis of DiLeuPMP tags

L-Leucine and sodium cyanoborohydride ($NaBH_3CN$) were suspended in MeOH and the mixture was cooled in an ice-water bath. Then, formaldehyde (CH_2O, 37% w/w) was added dropwise, and the mixture was stirred in an ice-water bath for 30 min. The product was purified by flash column chromatography (MeOH/DCM) and dried in vacuo to produce white solid *N*, *N*-dimethyl leucine (DiLeu). The DiLeu reagent was then dissolved in anhydrous THF with 1.2 molar excess of carbonyldiimidazole (CDI) at room temperature for 30 min and then equal molar of 2-(4-aminophenyl)-5-methyl-2,4-dihydro-3*H*-pyrazole-3-one was added. The reaction was heated to

70°C for 18hr and the non-isotopic version of DiLeuPMP was purified by flash column chromatography (MeOH/DCM) as brown color oil. The isobaric version of 4-plex DiLeuPMP reagents were synthesized following the same procedures with different isotopic starting reagents (**Figure S2**). For the channels requiring ^{18}O exchange, isotopic leucine was dissolved in 1 N HCl $H_2{}^{18}O$ solution (pH 1) and stirred on a hot plate at 65 °C for 4 h. Following evaporation of HCl from the solution in vacuo, trace amounts of acid were removed with StratoSpheres PL-HCO_3 MP resin (Agilent Technologies) to obtain ^{18}O L-leucine in free base form. The purity was tested using HPLC and structures were confirmed by NMR and/or MS (**Figures S3 – S5**).

DiLeuPMP labeling on core-1 O-glycan standard

The core-1 O-glycan standard was first dissolved in deionized water to make a stock solution at 1 mg/mL. 0.5 M DiLeuPMP was prepared in MeOH/NH4OH (50/50 v/v) at final concentration of 10% NH4OH. 10 μL of core-1 standard and DiLeuPMP were equally mixed, and the reaction was stood at 70 °C for 1h. Then the mixture was desalted using Pierce™ C18 tips, and the eluted fractions were dried in vacuo, reconstituted in 20 μL of 5% ACN, and analyzed by MALDI-MS or LC-MS/MS immediately.

DiLeuPMP labeling on standard glycoproteins and human serum

The release of O-glycans from glycoproteins was performed according to the published procedure with moderate modifications16. Briefly, 100 μg of standard glycoproteins were dissolved in 10% NH4OH/DMF (50/50 v/v), and DiLeuPMP was added to make the final concentration at 0.5 M. After incubating at 70°C for 24h, the pH of the solution was adjusted to neutral by acetic acids followed by liquid extraction using chloroform for three times. Organic phase was discarded and remaining aqueous phase was desalted using Sep-Pak C18 Cartridge. The

eluted fractions were dried in vacuo and reconstituted in 20 μL of 5% ACN, analyzed by MALDI-MS or LC-MS/MS. For 4-plex DiLeuPMP labeled samples, they were combined firstly before extraction.

In terms of human serum sample preparation, we followed FASP protocol with minor modifications.17 In brief, 100 μL healthy human serum was mixed with 300 μL 0.1% SDS and 100 mM TCEP in 10 mM sodium phosphate (pH 7.5) at 60 °C for 1 h. After the sample was cooled down, 10 kDa MWCO was used for buffer exchange. 400 μL of 10 mM sodium phosphate was used to rinse the filter three times. 8U of PNGase F in 100 μL TEAB buffer was added into the sample and the sample was incubated at 37°C for 18 h to remove N-glycans. 10 kDa MWCO filters were used to separate proteins and glycans. After the proteins were recollected, ZIC-HILIC were applied for O-glycoprotein/peptide enrichment18. HILIC beads were first activated with 200 μL of elution buffer (0.1% TFA, 99.9% H2O) for 30 min and then washed with binding buffer (0.1% TFA, 19.9% H2O, 80% ACN) twice. Proteins were dissolved in 300 μL of binding buffer and mixed with 7 mg activated ZIC-HILIC resin at a 1:50 peptide-to-material mass ratio in a microcentrifuge tube. The tube was shaken over a vortex mixer for 1 h and the supernatant was removed by centrifugation. The beads were washed with 70 μL binding buffer three times and glycoprotein/peptides were eluted with 70 μL elution buffer. Same labeling procedure with DiLeuPMP was conducted subsequently as mentioned above.

Matrix assisted Laser Desorption/Ionization (MALDI)-MS analysis

Samples were prepared by premixing 1 μL of DiLeuPMP-labeled O-glycans with 1 μL 2,5-dihydroxybenzoic acid (DHB) matrix (100 mg/mL in 5% *N*, *N*-dimethylaniline, 47.5% MeOH and 47.5% water), and 1 μL of each matrix/sample mixture was spotted onto the MALDI target plate. A MALDI-LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, Bremen, Germany) was used

for MALDI-MS analysis. Ionization was performed using laser energy of 17 µJ. Spectra were acquired in the Orbitrap mass analyzer within a mass range of *m/z* 1,000-4,000 at a mass resolution of 60 K (at *m/z* 400).

LC-MS/MS analysis

A self-fabricated nano-C18 column (15 cm, 75 µm i.d., 1.7 µm Ethylene Bridged Hybrid C18 packing material) was used for glycan separation. A Dionex Ultimate 3000 nanoLC system was coupled to a Q Exactive HF Hybrid Quadrupole Orbitrap Mass Spectrometer (Thermo Scientific, Bremen, Germany) for all LC-MS/MS analyses. Mobile phase A was water with 5% DMSO, and mobile phase B was ACN with 5% DMSO. The flow rate was set at 0.3 µL/min, and the injection volume was 2 µL. The following gradient was used (time, % mobile phase B) unless otherwise specified: (0 min, 5%), (28 min, 5%), (38 min, 9%), (128 min, 37%), (133 min, 95%), (143 min, 95%), (148 min, 5%), (158 min, 5%).

The following mass spectrometer parameters were used for all data acquisition. Samples were ionized in positive ion mode with a spray voltage of 3 kV. S-lens radio frequency (RF) level was set to be 30, and capillary temperature was set to be 300 °C. Full MS scans were acquired at *m/z* 350-2000 with resolving power of 120 K. Maximum injection time of 100 ms, automatic gain control (AGC) target value of 5×10^5, and 1 microscan were used for full MS scans. Top 15 data-dependent MS^2 analysis was performed at a resolving power of 15 K with higher-energy collisional dissociation (HCD) operating with normalized collision energy (NCE) of 25. The first fixed mass sets to 100 *m/z* in order to obtain reporter ions. The isolation window was set as 2.5 *m/z* and the dynamic exclusion of acquired precursors was set to 15 sec with a ± 20 ppm tolerance.

O-glycan data analysis

The raw data was compared against an in-house database including the most possible combinations of O-glycan units (Hexose (H), HexNAc (N), Fucose (F), and NeuAc (S)). DiLeuPMP-labeled O-glycans were identified by accurate mass matching in full MS with a mass tolerance of 10 ppm and fragmentation matching in MS/MS spectra assisted by GlycoWorkbench. Peak areas of reporter ions for DiLeuPMP-labeled glycans were used for relative quantification. Microsoft Excel and Origin were used for calculations and statistical analyses.

Results and Discussion

Numerous challenges exist for O-glycan analysis. In contrast to N-glycosylation, O-glycosylation lacks a known amino acid consensus sequence and Mucin-type O-glycosylation is usually highly heterogeneous. However, among the various challenges, the major difficulty lies in the lack of universal enzymatic tools for release of O-glycans from proteins. Therefore, the chemical method is currently the most effective and reliable way of acquiring released O-glycans, where β-elimination is the most common one. However, β-elimination demands alkaline conditions and usually results in degradation which was referred as "peeling". Previously, researchers found that adding PMP can efficiently block the reducing end of released glycans and prevent them from sequential peeling degradation. The derivatization reaction itself is a Michael addition involving a two-step labeling process in which Michael donor molecules of the labeling reagent are formed and consecutively added to the reducing end of the glycan with a stoichiometry of two label molecules per glycan[19]. The use of PMP and related tags was restricted in fluorescence detection initially[20] and was recently employed in the field of mass spectrometry[10, 21, 22].

O-glycans are known to be associated with many critical biological functions. Abnormal O-glycosylation has been implicated in a variety of diseases, including familial tumoral calcinosis[23], Tn syndrome[24], and IgA nephropathy[25]. Thus, global profiling and quantitative analysis of O-glycans is crucial to understand the structures and functions of O-glycosylation in the development of these diseases. The current analytical tool suffers from low throughput and tedious sample preparation procedures. Mass spectrometry (MS)-based, isobaric labeling strategy for relative quantification allows for parallel multiplexing of experiments, which provides an opportunity to address these limitations for O-glycan analysis. Generally, multiple samples are chemically labeled with isobaric chemical tag variants and each variant has the same chemical structure and nominal mass which is indistinguishable in full MS spectrum. However, each variant is fragmented to produce a unique "reporter ion" during tandem MS and the signal intensity can be used for relative quantification. Commercially available reagents such as isobaric tags for relative and absolute quantification (iTRAQ) and tandem mass tags (TMT) are widely used in quantitative proteomics[26]. In our lab, we previously developed DiLeu isobaric tags for cost-effective proteomics[27] and isobaric multiplex reagents for carbonyl-containing compound (SUGAR) tags for N-glycan analysis[4, 28].

By combining the PMP labeling and DiLeu isobaric labeling strategies together, we designed and synthesized a set of 4-plex DiLeuPMP tags to enable high-throughput quantitative O-glycomics (**Figure 1**). The reactive site of Michael donor PMP is on the five-member ring. Hence, we incorporated DiLeu reporter ion onto the benzene moiety. DiLeuPMP was created in two-step synthesis with an overall yield of 70%. The 4-plex isobaric version of DiLeuPMP tags were synthesized using heavy isotope-coded starting materials accordingly (**Figure S2**). The isotope purities of 4-plex DiLeuPMP were over 98% (**Figure S3**). As two label molecules were

added per glycan, the chromatographic properties of the derivatized oligosaccharides were largely affected by the tags, thus we avoided incorporating deuterium atom into 4-plex DiLeuPMP structures, using only ^{13}C, ^{15}N, and ^{18}O, to prevent possible retention time shift.

Releasing and labeling of O-glycans was performed in a one-pot manner. The general workflow is illustrated in **Figure 2**. After *in situ* releasing and labeling of O-glycans during β-elimination, the pH of the system was adjusted to neutral, and chloroform was used for liquid-liquid extraction. The organic layer was discarded to get rid of excess tags and the remaining aqueous phase was desalted using Sep-Pak C18 cartridges. The elute was dried and reconstituted for MALDI-MS or LC-MS/MS analysis. Compared to developed methods[29, 30], we further lowered the alkalinity of the system, choosing 5% NH_4OH instead of 100 mM NaOH or 50% hydrazine and found that in 5% NH_4OH, the releasing of O-glycans was still efficient and the peeling degradation can be effectively suppressed.

Commercially available core-1 O-glycan standard was used for method validation. As shown in **Figure 3**, after labeling, the signal corresponding to bis-DiLeuPMP labeled glycan (*m/z* 1026.5495) was abundant, while mono-DiLeuPMP labeled (*m/z* 696.3434) and peeling product (*m/z* 864.4965) were also observed. However, they were less than 5% of overall signal intensity, which indicated satisfactory labeling efficiency of DiLeuPMP and its efficient peeling suppression. It is worth mentioning that because two tags are added per glycan and it is the tag that commonly carries charges, on ESI-MS spectrum, the dominant peak corresponding to glycan conjugates carries +2 charges (*m/z* 513.7766 $[M+2H]^{2+}$, *m/z* 524.7674 $[M+H+Na]^{2+}$) while +1 charged signals are also noticeable (*m/z* 1026.5466 $[M+H]^{+}$, *m/z* 1048.5283 $[M+Na]^{+}$). The elevated charge states facilitate the fragmentation process and produce abundant fragment ions for structural elucidation.

Various model glycoproteins were used to further evaluate the feasibility of the described method for O-glycan analysis on glycoproteins. Bovine fetuin, bovine submaxillary mucin (BSM) and porcine stomach mucin (PSM) were chosen as model glycoproteins as the O-glycans on these glycoproteins are well studied[4, 28]. Bovine fetuin, which contains both N- and O-linked glycans, has a relatively simple O-glycosylation pattern. BSM belongs to the class of salivary glycoproteins, and it consists of a long protein chain with numerous disaccharide and oligosaccharide side chains, some of which are known to be sialylated. PSM is primarily composed of carbohydrate units with protein contributing to only 20% of the molecular weight. Following the one-pot procedure described above, the O-glycans were released from these glycoproteins and labeled by DiLeuPMP. We set a control group using PMP and compared our results with previously established methods. As shown in **Figure 4**, for MALDI-MS profiling, 3 O-glycans were found on bovine fetuin, 14 O-glycans were identified on BSM and 26 O-glycans were identified on PSM using DiLeuPMP labeling. Our results showed great overlap between PMP control as expected and the identified O-glycans are similar to those previously reported[31, 32] (**Figure S6, Table S1**). After coupling with LC separation, a higher number of identifications and coverage were achieved in LC-MS/MS results (**Table S2**). The distinct profiling, high coverage and good reproducibility of identified glycans between PMP and DiLeuPMP demonstrated the consistency and robustness of our new method compared with existing methods.

The quantification performance of the 4-plex DiLeuPMP tags was evaluated next by labeling O-glycans on PSM at known ratios. Denatured PSM was aliquoted into four equal portions in triplicate and labeled with 4-plex DiLeuPMP tags respectively, and then they were mixed with known ratios at 1:1:5:10 before performing liquid extraction and subsequent desalting. The combined desalted samples were then vacuum dried and reconstituted before performing LC-

MS/MS analysis. The intensities of reporter ions in MS/MS spectra for each conjugate were used to calculate the experimental ratios. In **Figure 5A**, experimental ratios of 4-plex DiLeuPMP-labeled O-glycans are plotted against theoretical ratios 1:1:5:10. Representative MS^2 spectra of low mass range containing reporter ions are shown in **Figure 5B**. For all three known ratios, less than 15% relative errors were observed with standard deviations of 0.15, 0.19, and 0.23, demonstrating that the DiLeuPMP quantification approach offers an accurate tool for quantitative analysis of O-glycans in a high-throughput manner. Moreover, no retention time shift was observed for 4-plex DiLeuPMP tag labeled O-glycans on the C18 column, as shown in **Figure 5C**.

Lastly, we applied our DiLeuPMP labeling strategy to the O-glycan analysis of healthy human serum for exploration in real biological samples. Since numerous clinically relevant analyses utilize human blood serum or plasma for both routine clinical measurements and potential disease biomarker discovery, the ability to analyze real biological specimen will surely expand the scope of usage of DiLeuPMP tags. The healthy human serum was collected in a FDA-approved collection center. The sample preparation was followed by a previously reported method with moderate modifications[33]. Briefly, the proteins were extracted and denatured firstly, and N-glycans were removed prior to O-glycoprotein/peptides enrichment. The enriched samples were then aliquoted into four equal portions before being labeled by 4-plex DiLeuPMP. After the derivatization procedure, four samples were equally mixed followed by cleanup procedure prior to LC-MS analysis. The data analysis revealed that, 12 O-glycans could be reliably recorded in a profile with a quantitative ratio error of less than 15% (**Table S3**). We anticipate that these are among the most abundant O-glycans in human serum, and it is consistent with previous studies[33, 34]. Furthermore, it is noted that the fragment ions from the precursors are abundant, which aids manual confirmation and structural elucidation of O-glycans (**Figure 5D and Figure S7**). In

summary, our results shed light on the potential application of the DiLeuPMP labeling strategy for high-throughput quantitative O-glycomic analysis of clinical samples.

Conclusions

In summary, 4-plex DiLeuPMP isobaric tags were developed in this study for quantitative O-glycomics. This is the first isobaric tag designed and developed for high-throughput quantitative analysis of O-glycans. By combining traditional PMP-aided β-elimination method and DiLeu enabled isobaric labeling strategy together, DiLeuPMP achieved releasing and labeling of O-glycans simultaneously as well as enabling high throughput MS2-based quantitative analysis. The labeling efficiency of DiLeuPMP is high and the peeling degradation is suppressed in a decreased alkali condition. The labeling pattern of two tags per glycan elevated the charges of labeled glycan conjugates and thus facilitated the fragmentation during MS/MS and produced abundant fragment ions for structural elucidation and reporter ions for quantification. We also demonstrated accurate relative quantification across a magnitude dynamic range using standard glycoproteins with these novel isobaric tags and applied the method onto analysis of human serum sample, suggesting the potential applications for large-scale analyses of biological and clinical specimens. Overall, the successful development of DiLeuPMP tags offers a powerful chemical tool for glycomics study in many biological and clinical applications and makes parallel profiling and quantitative analysis of O-glycans readily implemented. In conclusion, we anticipate that the novel DiLeuPMP labeling approach can be widely applied in a variety of biomedical research areas.

Acknowledgements

This research was support for this research was provided in part by the NIH grants U01CA231081, R01DK071801, RF1AG052324, and P41GM108538. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531) 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 Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

References

1. Dell, A.; Morris, H. R., Glycoprotein structure determination mass spectrometry. *Science* **2001,** *291* (5512), 2351-2356.

2. Helenius, A.; Aebi, M., Intracellular functions of N-linked glycans. *Science* **2001,** *291* (5512), 2364-2369.

3. Marth, J. D.; Grewal, P. K., Mammalian glycosylation in immunity. *Nat. Rev. Immunol.* **2008,** *8* (11), 874-887.

4. Feng, Y.; Li, M. Y.; Lin, Y. Y.; Chen, B. M.; Li, L. J., Multiplex Quantitative Glycomics Enabled by Periodate Oxidation and Triplex Mass Defect Isobaric Multiplex Reagents for Carbonyl-Containing Compound Tags. *Anal. Chem.* **2019,** *91* (18), 11932-11937.

5. Jensen, P. H.; Kolarich, D.; Packer, N. H., Mucin-type O-glycosylation - putting the pieces together. *FEBS J.* **2010,** *277* (1), 81-94.

6. Mulagapati, S.; Koppolu, V.; Raju, T. S., Decoding of O-Linked Glycosylation by Mass Spectrometry. *Biochemistry* **2017,** *56* (9), 1218-1226.

7. Huang, Y. P.; Mechref, Y.; Novotny, M. V., Microscale nonreductive release of O-linked glycans for subsequent analysis through MALDI mass spectrometry and capillary electrophoresis. *Anal. Chem.* **2001,** *73* (24), 6063-6069.

8. Goetz, J. A.; Novotny, M. V.; Mechref, Y., Enzymatic/Chemical Release of O-Glycans Allowing MS Analysis at High Sensitivity. *Anal. Chem.* **2009,** *81* (23), 9546-9552.

9. Furukawa, J.; Fujitani, N.; Araki, K.; Takegawa, Y.; Kodama, K.; Shinohara, Y., A Versatile Method for Analysis of Serine/Threonine Posttranslational Modifications by beta-Elimination in the Presence of Pyrazolone Analogues. *Anal. Chem.* **2011,** *83* (23), 9060-9067.

10. Wang, C. J.; Fan, W. C.; Zhang, P.; Wang, Z. F.; Huang, L. J., One-pot nonreductive O-glycan release and labeling with 1-phenyl-3-methyl-5-pyrazolone followed by ESI-MS analysis. *Proteomics* **2011,** *11* (21), 4229-4242.

11. Sic, S.; Maier, N. M.; Rizzi, A. M., Quantitative fingerprinting of O-linked glycans released from proteins using isotopic coded labeling with deuterated 1-phenyl-3-methyl-5-pyrazolone. *J. Chromatogr. A* **2015,** *1408*, 93-100.

12. Zhang, P.; Zhang, Y.; Xue, X. D.; Wang, C. J.; Wang, Z. F.; Huang, L. J., Relative quantitation of glycans using stable isotopic labels 1-(d0/d5) phenyl-3-methyl-5-pyrazolone by mass spectrometry. *Anal. Biochem.* **2011,** *418* (1), 1-9.

13. Zauner, G.; Kozak, R. P.; Gardner, R. A.; Fernandes, D. L.; Deelder, A. M.; Wuhrer, M., Protein O-glycosylation analysis. *Biol. Chem.* **2012,** *393* (8), 687-708.

14. Wilkinson, H.; Saldova, R., Current Methods for the Characterization of O-Glycans. *J. Proteome Res.* **2020,** *19* (10), 3890-3905.

15. Xiang, F.; Ye, H.; Chen, R. B.; Fu, Q.; Li, L. J., N,N-Dimethyl Leucines as Novel Isobaric Tandem Mass Tags for Quantitative Proteomics and Peptidomics. *Anal. Chem.* **2010,** *82* (7), 2817-2825.

16. Karlsson, N. G.; Packer, N. H., Analysis of O-linked reducing oligosaccharides released by an in-line flow system. *Anal. Biochem.* **2002,** *305* (2), 173-85.

17. Ni, M. W.; Wang, L.; Chen, W.; Mou, H. Z.; Zhou, J.; Zheng, Z. G., Modified filter-aided sample preparation (FASP) method increases peptide and protein identifications for shotgun proteomics. *Rapid Commun. Mass Spectrom.* **2017,** *31* (2), 171-178.

18. 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* (9), 1751-1764.

19. Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M., Glycan labeling strategies and their use in identification and quantification. *Anal. Bioanal. Chem.* **2010,** *397* (8), 3457-3481.

20. Goso, Y.; Sugaya, T.; Ishihara, K.; Kurihara, M., Comparison of Methods to Release Mucin-Type O-Glycans for Glycomic Analysis. *Anal. Chem.* **2017,** *89* (17), 8870-8876.

21. Wang, C. J.; Lu, Y.; Han, J. L.; Jin, W. J.; Li, L. M.; Zhang, Y.; Song, X. Z.; Huang, L. J.; Wang, Z. F., Simultaneous Release and Labeling of O- and N-Glycans Allowing for Rapid Glycomic Analysis by Online LC-UV-ESI-MS/MS. *J. Proteome Res.* **2018,** *17* (7), 2345-2357.

22. Yuan, J. B.; Wang, C. J.; Sun, Y. J.; Huang, L. J.; Wang, Z. F., Nonreductive chemical release of intact N-glycans for subsequent labeling and analysis by mass spectrometry. *Anal. Biochem.* **2014,** *462*, 1-9.

23. Ichikawa, S.; Guigonis, V.; Imel, E. A.; Courouble, M.; Heissat, S.; Henley, J. D.; Sorenson, A. H.; Petit, B.; Lienhardt, A.; Econs, M. J., Novel GALNT3 mutations causing hyperostosis-hyperphosphatemia syndrome result in low intact fibroblast growth factor 23 concentrations. *J. Clin. Endocrinol. Metab.* **2007,** *92* (5), 1943-1947.

24. Ju, T. Z.; Cummings, R. D., Protein glycosylation - Chaperone mutation in Tn syndrome. *Nature* **2005,** *437* (7063), 1252-1252.

25. Hiki, Y.; Kokubo, T.; Iwase, H.; Masaki, Y.; Sano, T.; Tanaka, A.; Toma, K.; Hotta, K.; Kobayashi, Y., Underglycosylation of IgA1 hinge plays a certain role for its glomerular deposition in IgA nephropathy. *J. Am. Soc. Nephrol.* **1999,** *10* (4), 760-769.

26. Moulder, R.; Bhosale, S. D.; Goodlett, D. R.; Lahesmaa, R., Analysis of the plasma proteome using iTRAQ and TMT-based Isobaric labeling. *Mass Spectrom. Rev.* **2018,** *37* (5), 583-606.

27. Greer, T.; Hao, L.; Nechyporenko, A.; Lee, S.; Vezina, C. M.; Ricke, W. A.; Marker, P. C.; Bjorling, D. E.; Bushman, W.; Li, L. J., Custom 4-Plex DiLeu Isobaric Labels Enable Relative Quantification of Urinary Proteins in Men with Lower Urinary Tract Symptoms (LUTS). *Plos One* **2015,** *10* (8).

28. Feng, Y.; Chen, B. M.; Yu, Q. Y.; Zhong, X. F.; Frost, D. C.; Ikonomidou, C.; Li, L. J., Isobaric Multiplex Labeling Reagents for Carbonyl-Containing Compound (SUGAR) Tags: A Probe for Quantitative Glycomic Analysis. *Anal. Chem.* **2019,** *91* (4), 3141-3146.

29. Kozak, R. P.; Royle, L.; Gardner, R. A.; Fernandes, D. L.; Wuhrer, M., Suppression of peeling during the release of O-glycans by hydrazinolysis. *Anal. Biochem.* **2012,** *423* (1), 119-128.

30. Xu, G. G.; Amicucci, M. J.; Cheng, Z.; Galermo, A. G.; Lebrilla, C. B., Revisiting monosaccharide analysis - quantitation of a comprehensive set of monosaccharides using dynamic multiple reaction monitoring. *Analyst* **2018,** *143* (1), 200-207.

31. Miura, Y.; Kato, K.; Takegawa, Y.; Kurogochi, M.; Furukawa, J.; Shinohara, Y.; Nagahori, N.; Amano, M.; Hinou, H.; Nishimura, S. I., Glycoblotting-Assisted O-Glycomics: Ammonium Carbamate Allows for Highly Efficient O-Glycan Release from Glycoproteins. *Anal. Chem.* **2010,** *82* (24), 10021-10029.

32. Yamada, K.; Hirabayashi, J.; Kakehi, K., Analysis of O-Glycans as 9-Fluorenylmethyl Derivatives and Its Application to the Studies on Glycan Array. *Anal. Chem.* **2013,** *85* (6), 3325-3333.

33. Gizaw, S. T.; Gaunitz, S.; Novotny, M. V., Highly Sensitive O-Glycan Profiling for Human Serum Proteins Reveals Gender-Dependent Changes in Colorectal Cancer Patients. *Anal. Chem.* **2019,** *91* (9), 6180-6189.

34. Williams, T. I.; Saggese, D. A.; Muddiman, D. C., Studying O-linked protein glycosylations in human plasma. *J. Proteome Res.* **2008,** *7* (6), 2562-8.

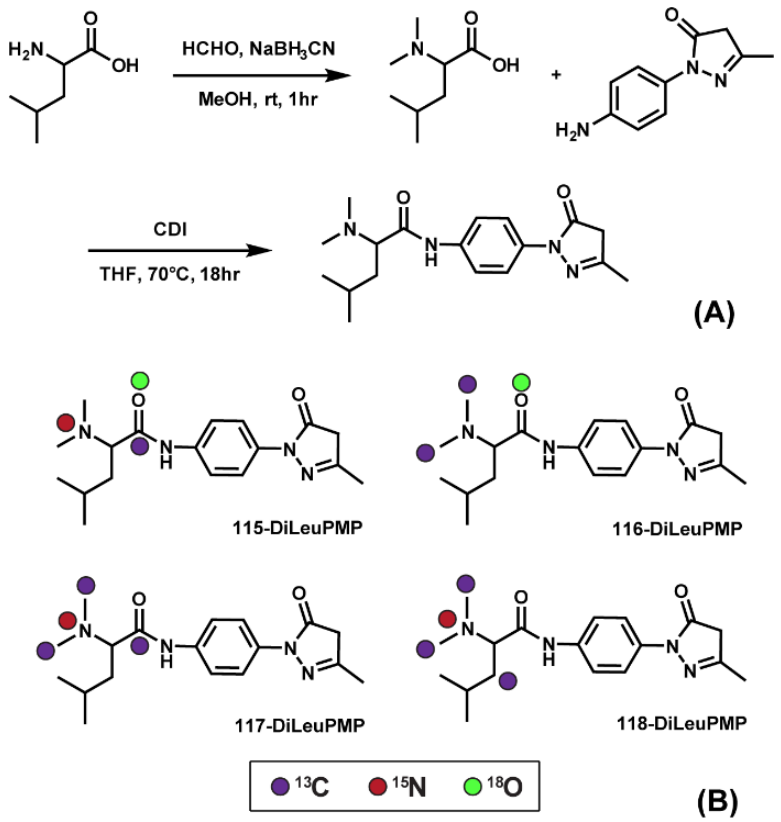


Figure 1. (A) Structure and synthetic routes of DiLeuPMP tag; (B) Isotopic configurations of 4-plex DiLeuPMP tags.

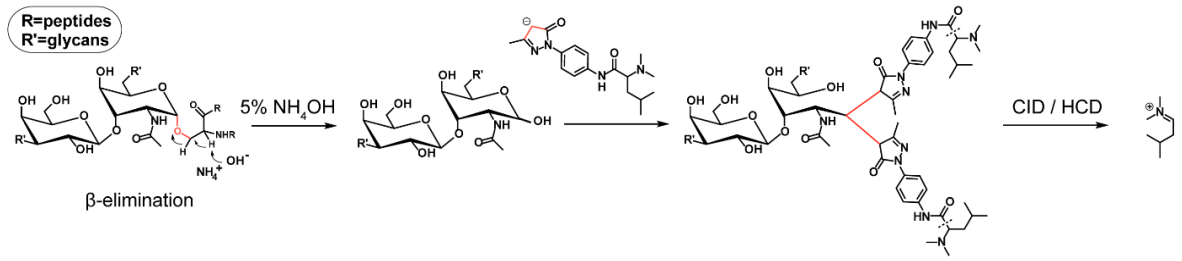


Figure 2. Scheme of released O-glycan labeled by DiLeuPMP at 2:1 (tag to glycan) ratio.

*the reaction site is highlighted

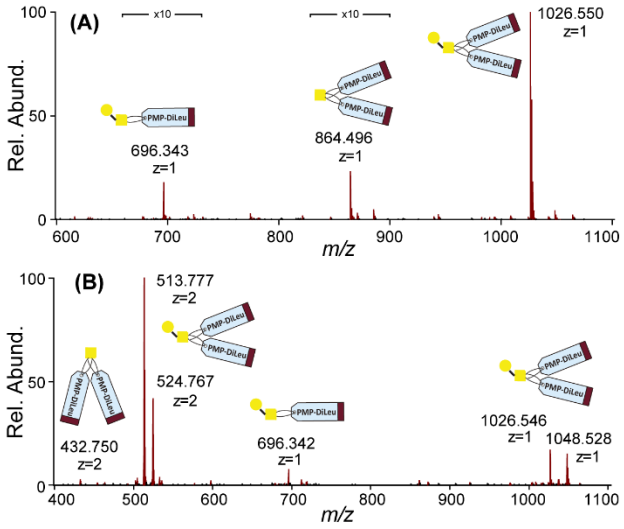


Figure 3. (A). MALDI-MS spectrum of DiLeuPMP labeled core-1 O-glycan standard; (B) ESI-MS spectrum of DiLeuPMP labeled core-1 O-glycan standard.

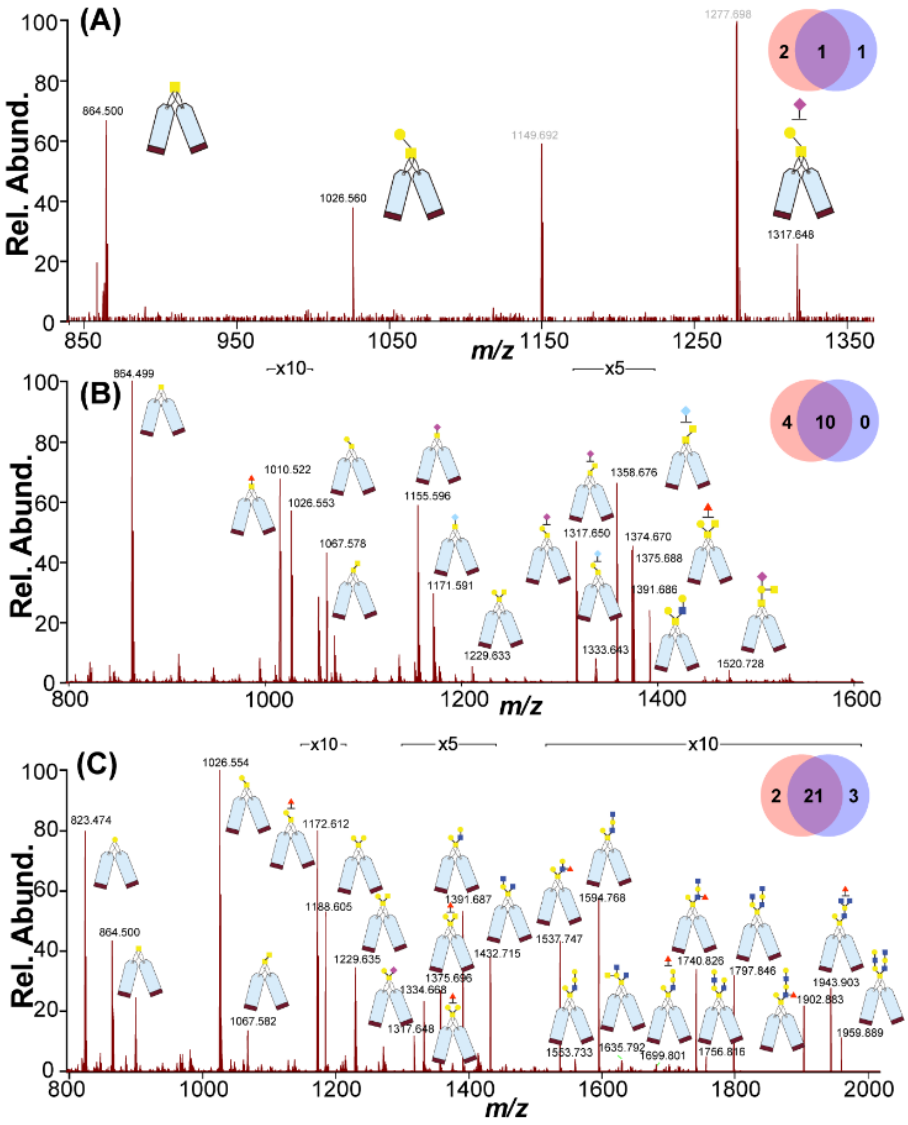


Figure 4. MALDI-MS profiling of DiLeuPMP labeled O-glycans released from standard glycoproteins (A) Bovine fetuin; (B) Bovine Submaxillary Mucin (BSM); (C) Porcine Stomach Mucin (PSM). The inset Venn diagrams above each spectrum show the numbers of identified glycans from PMP control (blue purple, right) and DiLeuPMP (orange red, left).

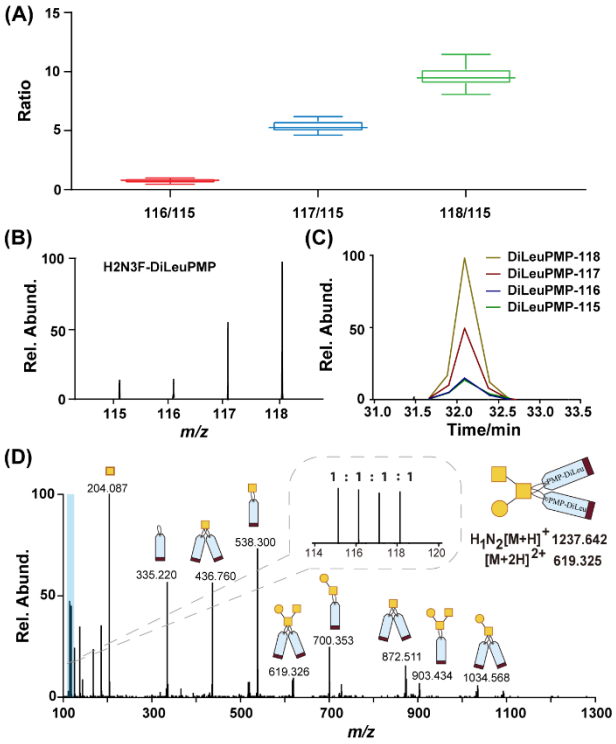


Figure 5. (A) Relative quantification performance of 4-plex DiLeuPMP labeled O-glycans released from glycoprotein PSM. Labeled O-glycans were mixed at ratios of 1:1:5:10 and analyzed in triplicates. Box plots show the median (line), the 25th and 75th percentile (box), and the 5th and 95th percentile (whiskers); (B) Representative MS spectrum of reporter ion region for 4-plex DiLeuPMP labeled H_2N_3F; (C) LC retention times of reporter ions generated from 4-plex DiLeuPMP labeled H_2N_3F; (D) ESI-MS/MS fragmentation of DiLeuPMP labeled H_1N_2 from healthy human serum sample showing efficient fragmentation and abundant reporter ions in the zoomed inset allowing quantification. Note: H-hexose, N-N-acetylhexoamine, F-fucose.

Supplemental Information

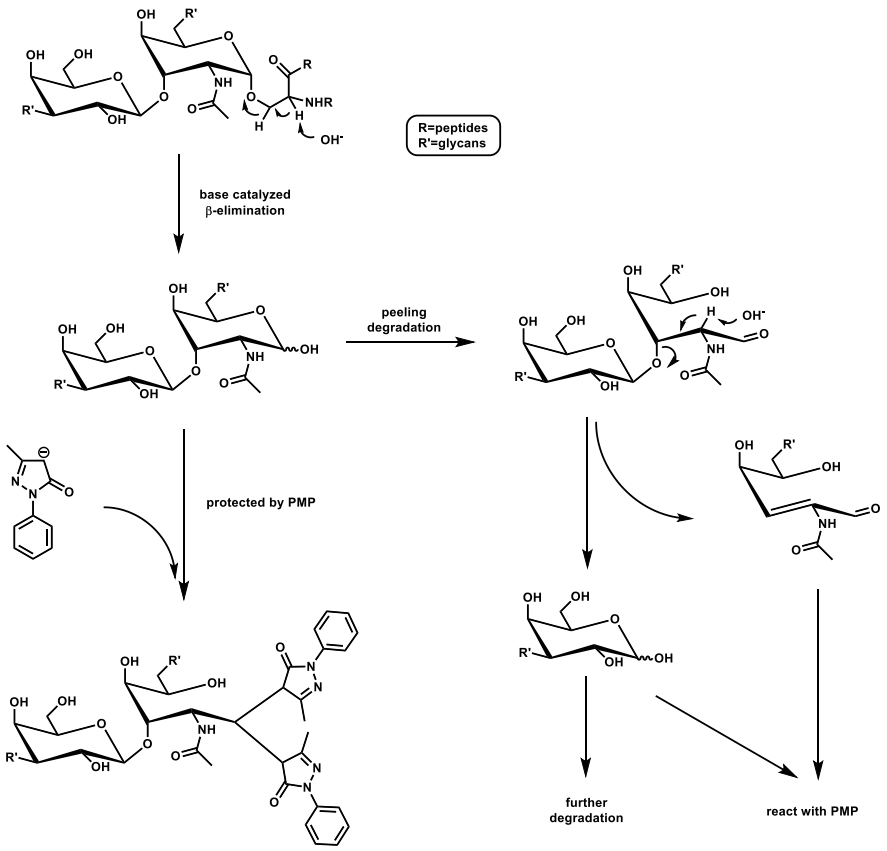


Figure S1. β-elimination causing 'peeling' degradation.

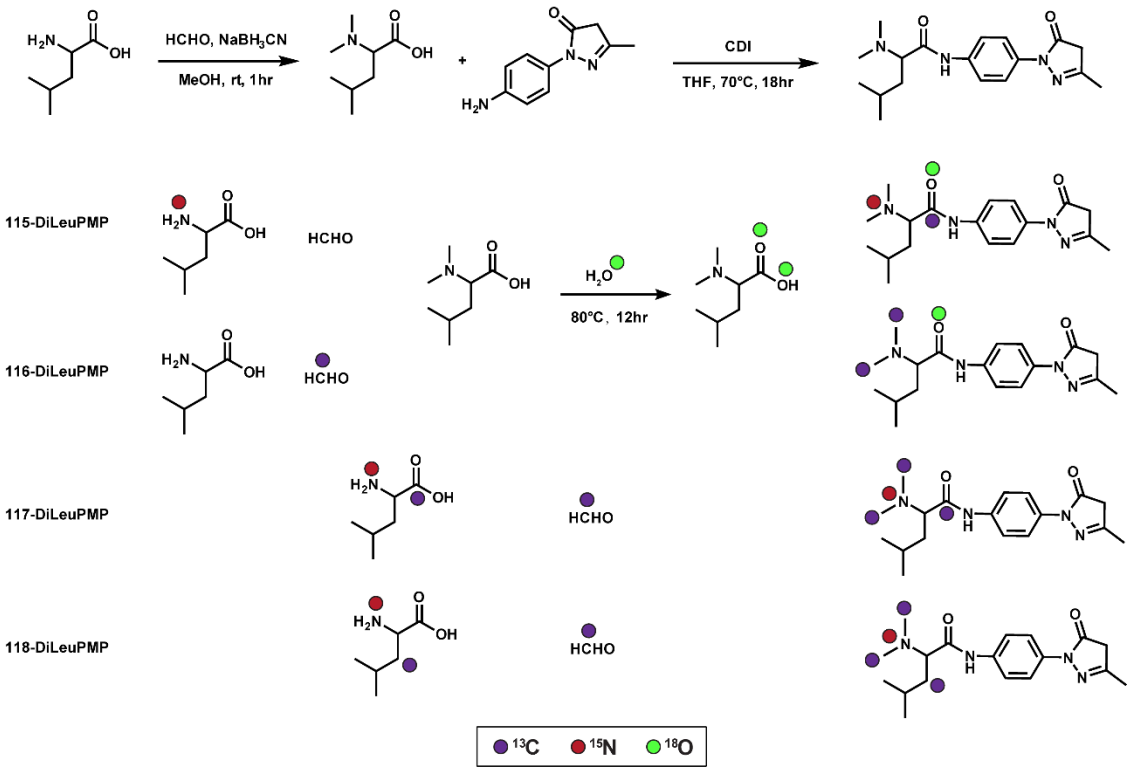


Figure S2. Synthetic route and starting isotopic materials for 4-plex DiLeuPMP tags.

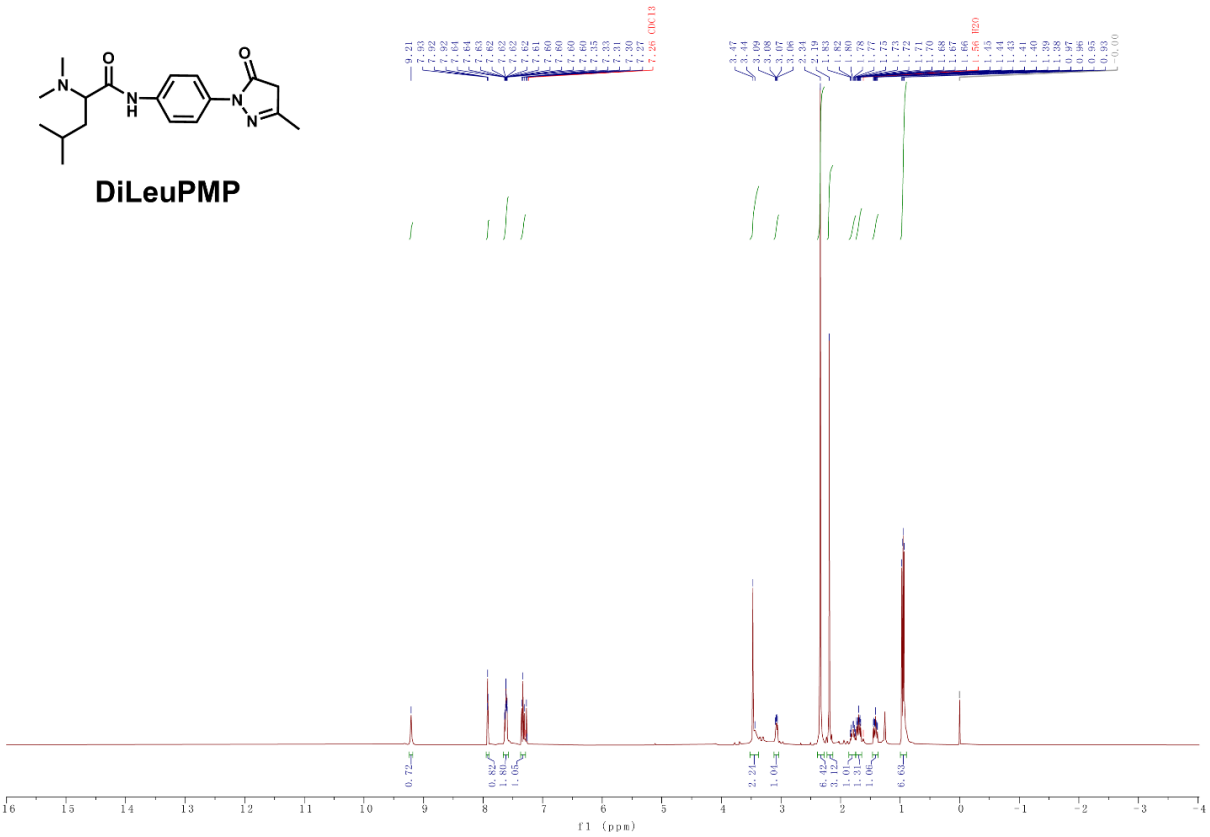


Figure S3. ^{1}H NMR of DiLeuPMP tag.

^{1}H NMR (400 MHz, $CDCl_3$) 9.21 (s, NH), 7.93 (s, 1H), 7.64 (m, 2H), 7.35 (m, 1H), 3.47 (s, 2H), 3.09 (q, 1H), 2.34 (s, 6H), 2.19 (s, 3H), 1.83 (m, 1H), 1.75 (t, 1H), 1.45 (m, 1H), 0.97 (d, 6H).

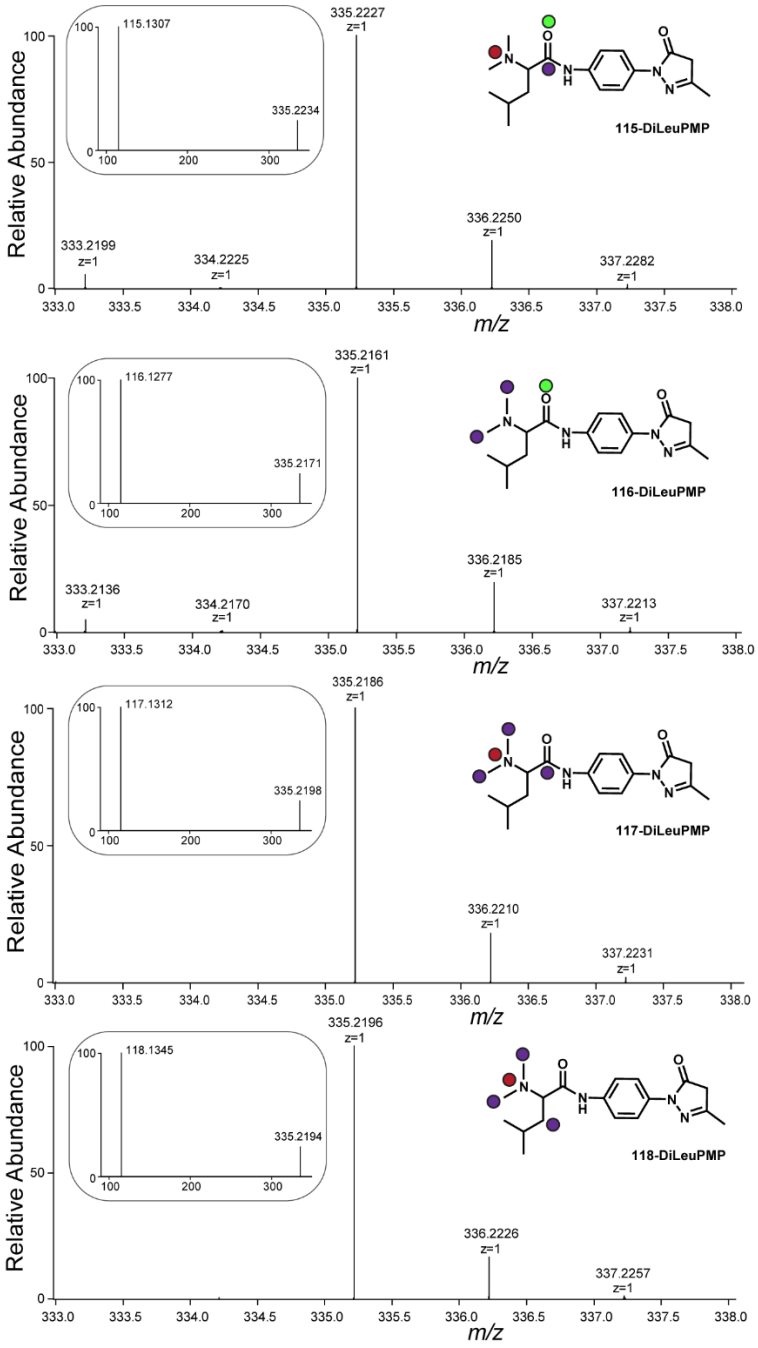


Figure S4. Isotope purities of 4-plex DiLeuPMP. The peaks at *m/z* 333 in 115 and 116 channels are due to insufficient ^{18}O exchange that are less than 5% of abundance. Insets are MS/MS of precursor ions highlighting the distinct reporter ions for 4-plex channel (@HCD25).

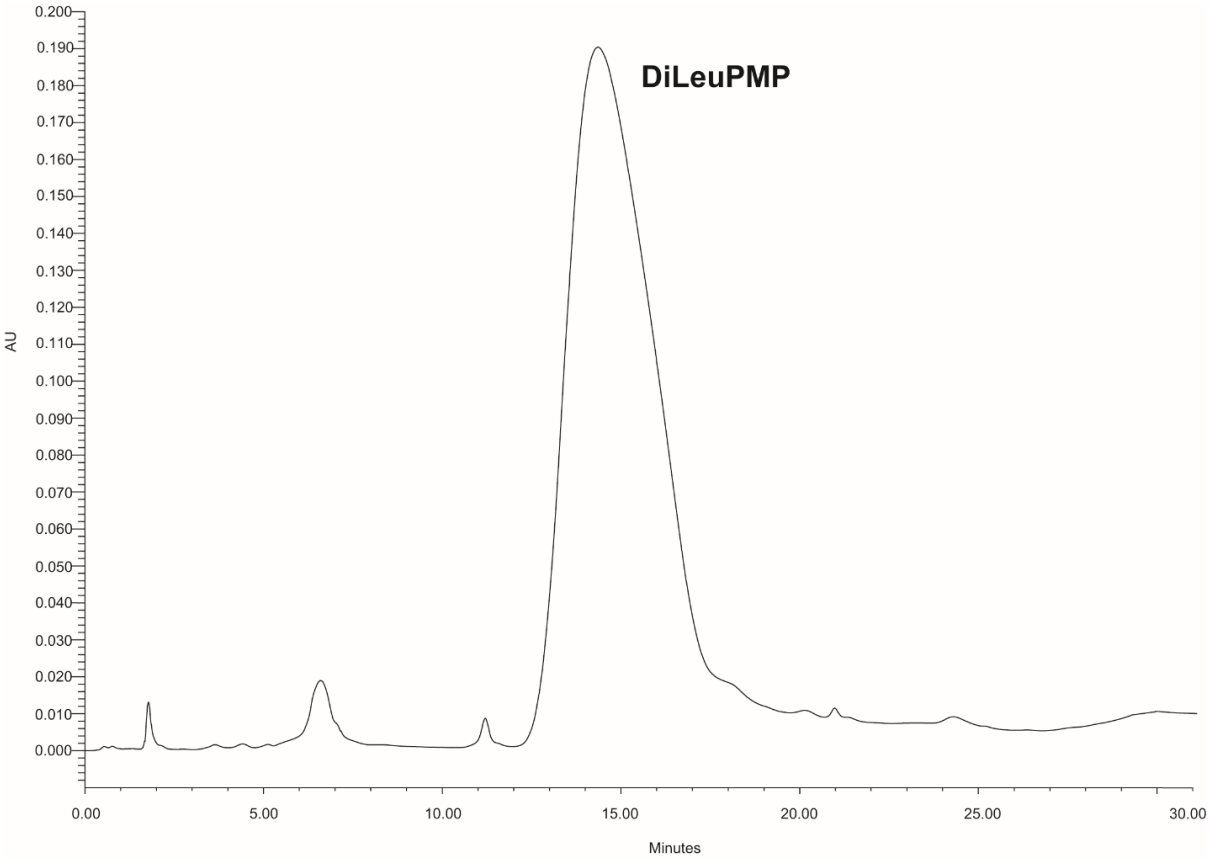


Figure S5. HPLC graph of DiLeuPMP.

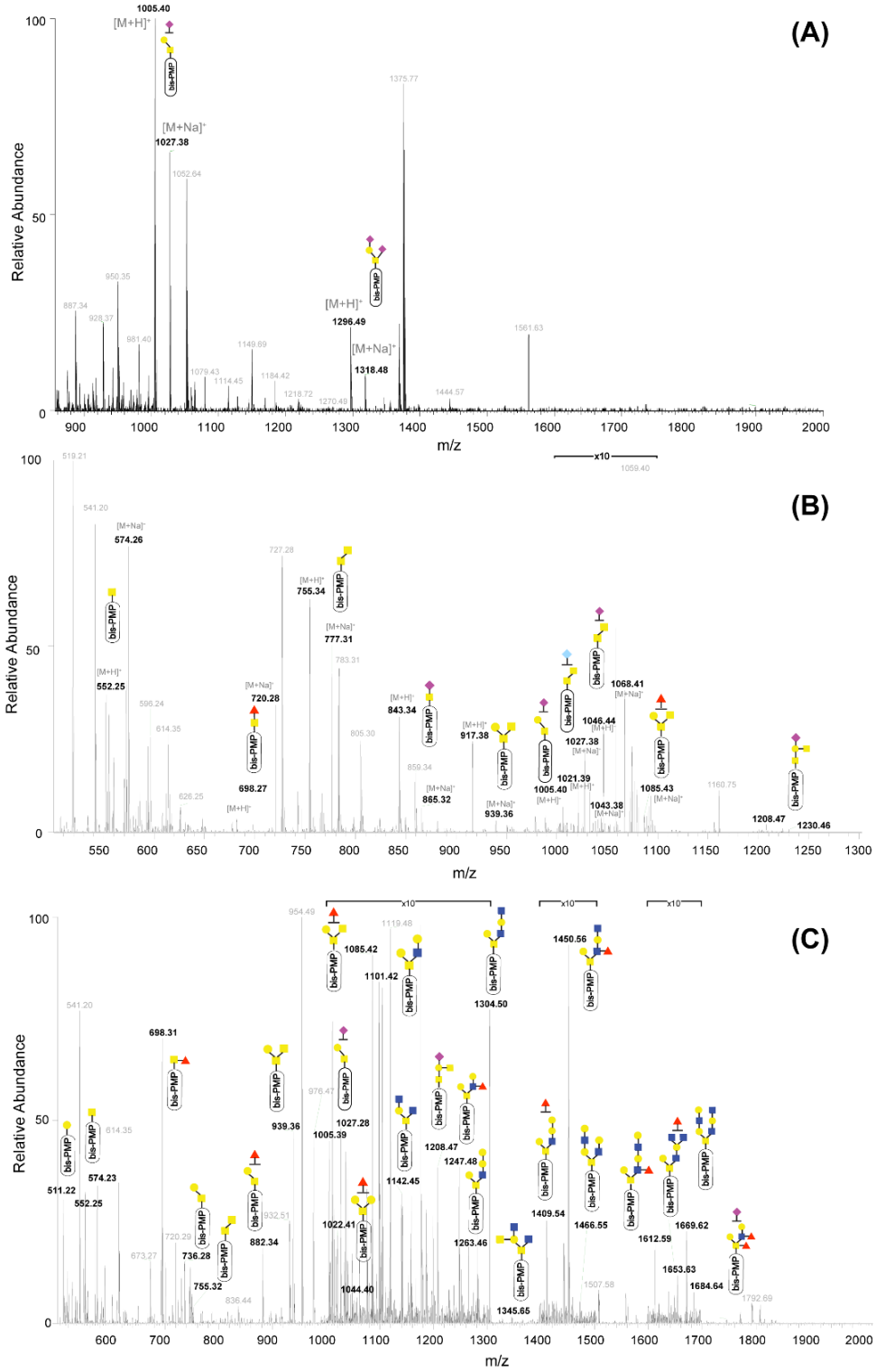


Figure S6. MALDI-MS profiling of PMP labeled O-glycans released from standard glycoproteins (A) Bovine fetuin; (B) Bovine Submaxillary Mucin (BSM); (C) Porcine Stomach Mucin (PSM).

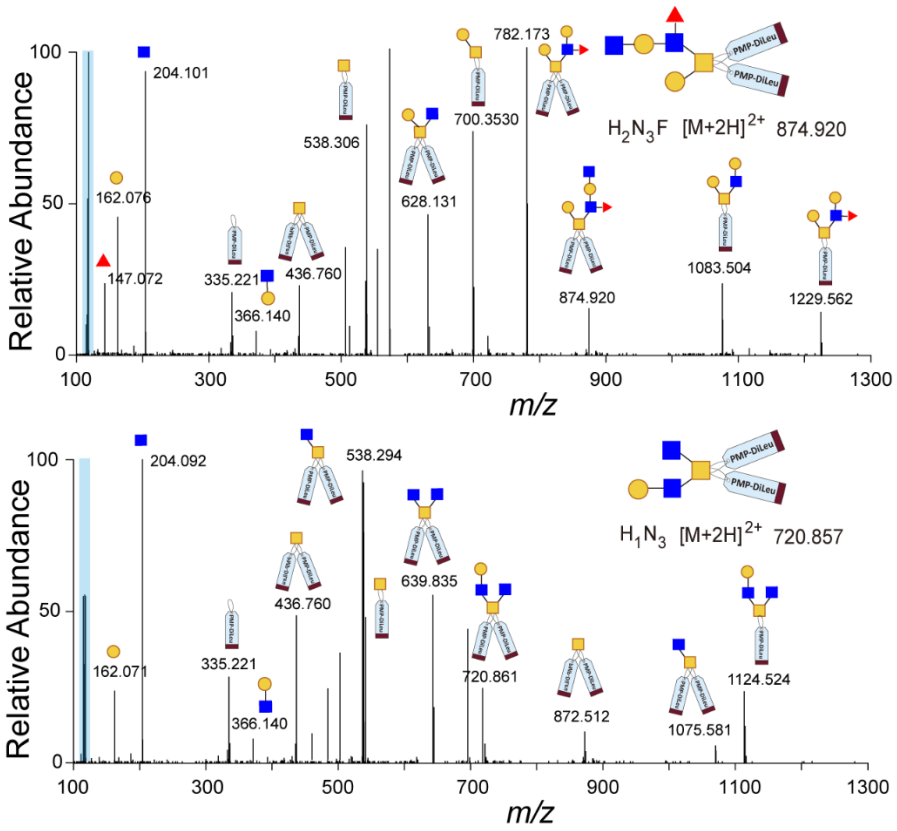


Figure S7. Representative MS/MS of O-glycans quantified in PSM, H2N3F (top) and in human serum, H1N3 (bottom).

Table S1. Observed O-glycans in MALDI-MS from glycoproteins using two methods, DiLeuPMP and PMP. (H: hexose, N: N-acetylhexosamine, F: fucose; S: sialic acid)

	H	N	F	S(NeuAc)	S(NeuGc)	DiLeuPMP	PMP
O-glycans of Bovine Fetuin							
1	0	1	0	0	0	√	-
2	1	1	0	0	0	√	-
3	1	1	0	1	0	√	√
4	1	1	0	2	0	-	√
O-glycans of Bovine Submaxillary Mucin (BSM)							
1	0	1	0	0	0	√	√
2	0	1	1	0	0	√	√
3	1	1	0	0	0	√	-
4	0	2	0	0	0	√	√
5	0	1	0	1	0	√	√
6	0	1	0	0	1	√	-
7	1	2	0	0	0	√	√
8	1	1	0	1	0	√	√
9	1	1	0	0	1	√	-
10	0	2	0	1	0	√	√
11	0	2	0	0	1	√	√
12	1	2	1	0	0	√	√
13	2	2	0	0	0	√	-
14	1	2	0	1	0	√	√
O-glycans of Porcine Stomach Mucin (PSM)							
1	1	0	0	0	0	√	√
2	0	1	0	0	0	√	√
3	1	1	0	0	0	√	√
4	0	2	0	0	0	√	√
5	0	1	1	0	0	-	√
6	1	1	1	0	0	√	√
7	2	1	0	0	0	√	-
8	1	2	0	0	0	√	√
9	1	2	0	1	0	-	√
10	1	1	0	1	0	√	√
11	2	1	1	0	0	√	√
12	1	2	1	0	0	√	√
13	2	2	0	0	0	√	√
14	1	3	0	0	0	√	√
15	2	2	1	0	0	√	√
16	3	2	0	0	0	√	√
17	1	4	0	0	0	√	√
18	3	2	1	0	0	√	√
19	2	3	0	0	0	√	√
20	2	3	1	0	0	√	√
21	3	3	0	0	0	√	√
22	2	4	0	0	0	√	-
23	3	3	1	0	0	√	√
24	2	4	1	0	0	√	√
25	3	4	0	0	0	√	√
26	2	2	2	1	0	-	√

Chapter 4

Selective Enrichment of Sialylglycopeptides Enabled by Click Chemistry and Dynamic Covalent Exchange

Adapted from: **Miyang Li**, Junfeng Huang, Min Ma, Xudong Shi, Lingjun Li. Selective Enrichment of Sialylglycopeptides Enabled by Click Chemistry and Dynamic Covalent Exchange *Anal. Chem.* **2022**, 94, 18, 6681–6688.

Abstract

Despite the important roles of protein sialylation in biological processes such as cellular interaction and cancer progression, simple and effective methods for the analysis of intact sialylglycopeptides (SGPs) are still limited. Analyses of low-abundance SGPs typically require efficient enrichment prior to comprehensive liquid chromatography-mass spectrometry (LC-MS)-based analysis. Here, a novel workflow combining mild periodate oxidation, hydrazide chemistry, copper-catalyzed azide/alkyne cycloaddition (CuAAC) click chemistry, and dynamic covalent exchange, has been developed for selective enrichment of SGPs. The intact SGPs could be separated easily from protein tryptic digests and the signature ions were produced during LC-MS/MS for unambiguous identification. The structure of the signature ions and corresponding dynamic covalent exchange was confirmed by using isotopic reagent. Under the optimized condition, over 70% enrichment efficiency of SGPs was achieved on bovine fetuin digests, and the method was successfully applied to complex biological samples, such as a mouse lung tissue extract. The high enrichment efficiency, good reproducibility, and easily adopted procedure without the need to generate specialized materials make this method a promising tool for broad applications in SGPs analysis.

Introduction

Glycosylation is one of the most important post-translational modifications (PTM) and plays essential roles in many biological processes such as receptor recognition, cellular communication, and immune response [1-3]. Protein sialylation, where the terminal of glycan moieties on a protein are capped with sialic acids (SA), is a key form of glycosylation and is commonly found on cell-surface proteins.[4] Aberrant sialylation closely correlates with many

pathological changes such as neurological diseases and malignant tumors[5-8]. Mass spectrometry (MS)-based glycoproteomics has gained popularity in characterizing glycopeptides[9]. However, direct analysis of intact sialylglycopeptides (SGPs) is still challenging, largely due to the low abundance of SGPs and the heterogeneity of peptide mixture in the protein digests. Therefore, selective enrichment of SGPs prior to MS analysis is essential for characterizing SGPs and understanding the vital roles of protein sialylation in cellular functions and disease progression[10].

Based on the physicochemical properties of SA, the enrichment methods for SGPs can be arbitrarily classified into two categories: physical adsorption and chemical derivatization. The former strategies utilize the hydrophilic or negatively charged properties of SA to selectively extract SGPs from a complex mixture, while the latter commonly modify the unique vicinal diol groups within the structure of SA to achieve selectivity. For example, lectin-based affinity chromatography, titanium dioxide (TiO_2) chromatography, and hydrophilic interaction liquid chromatography (HILIC) have been widely used to identify sialylated glycoproteins based on physical interactions between SGPs and packing materials, however, the limited specificity, low efficiency, and high cost restricted their applications[11-16]. Taking advantage of the negative charge of SA, ion-exchange chromatography was employed to enrich SGPs but the co-elution of non-glycopeptides with acidic amino acids reduced selectivity[16-19]. As a popular chemical method known as “reverse glycoblotting”, hydrazide chemistry promoted glycoproteomics long before it was applied specifically to SGPs analysis[20]. When enriching SGPs, sialylated glycoproteins were selectively periodate-oxidized, captured on hydrazide beads, trypsin digested, and released by acid hydrolysis of SAs[21]. Although exhibiting high SGPs enrichment selectivity, hydrazide chemistry-based methods lose the information of terminal SAs[22]. Recently, various novel solid-phase materials were developed to enrich SGPs exclusively, on account of in-depth interdisciplinarity

and a joint force of material science and analytical chemistry to solve problems[23-26]. Among them, an interesting study demonstrated the unique advantage of dynamic covalent chemistry based on Schiff-base hydrolysis in capturing SGPs[27]. Despite the excellent performance, the uniqueness of each material and complicated process for producing as well as characterizing them hindered their widespread applications in general analytical labs.

In this work, we developed a simple, general, and easy access method to selectively enrich SGPs using commercially available materials. By incorporating copper-catalyzed azide/alkyne cycloaddition (CuAAC), so-called click chemistry, into the scheme of hydrazide chemistry method, SGPs were extracted and enriched with high efficiency. In addition, under mild elution conditions, dynamic covalent exchange based on Schiff-base swap facilely released modified SGPs without affecting glycans and peptide structures, thus preserving information of SAs on the glycopeptides, overcoming the limitations of the traditional reverse glycoblotting method (**Figure 1**). The unique fragment ions produced in MS/MS were confirmed by isotopic labeling study and used for unambiguous identification of SGPs. The enrichment performance and reproducibility of this method were evaluated and validated using bovine fetuin digests. Furthermore, the new strategy was applied to the analysis of mouse lung tissue samples. These results suggest that our method has great potential to be widely applied for the enrichment of SGPs in biomedical research.

Materials and Methods

Materials and Reagents.

Acetonitrile (ACN), dichloromethane (DCM), dimethyl sulfoxide (DMSO), formic acid (FA), methanol (MeOH), thionyl chloride ($SOCl_2$), copper sulfate ($CuSO_4$), sodium hydroxide (NaOH),

sodium periodate ($NaIO_4$), sodium acetate, sodium ascorbate, hydrazine and water (HPLC grade) were purchased from Fisher Scientific (Pittsburgh, PA). Isotopic reagent hydrazine sulfate ($^{15}N_2H_4 \cdot H_2SO_4$), triethylammonium bicarbonate (TEAB) buffer, SGP standard, bovine fetuin, 4-pentynoic acid were purchased from Sigma-Aldrich (St. Louis, MO). Dde-Azide agarose, PEG-azide resin and tris-hydroxypropyltriazolylmethylamine (THPTA) were purchased from Click Chemistry Tools (Scottsdale, AZ). Maltooctaose $(Glc)_8$ was purchased from Biosynth Carbosynth. Trypsin was purchased from Promega (Madison, WI). Sep-Pak C18 Cartridges were purchased from Waters Corporation (Milford, MA). Ethylene Bridged Hybrid C18 packing material (1.7 µm) was purchased from PolyLC Inc. (Columbia, MD). Fused silica capillary tubing (inner diameter 75 µm, outer diameter 375 µm) was purchased from Polymicro Technologies (Phoenix, AZ). All reagents were used without additional purification.

Synthesis of Hydrazide-Alkyne tags (HA-tag)

HA-tag can be synthesized conveniently in-house or substituted by a commercially available reagent Alkyne Hydrazide from Lumiprobe Corporation (Hunt Valley, MD). The synthetic route has two steps (**Figure S1**). In brief, 4-pentynoic acid was dissolved in MeOH at 0 °C with the $SOCl_2$ slowly adding to the solution dropwise. After 4 h, the reaction mixture was dried on rotavapor, and the product was purified through column chromatography (MeOH/DCM=1:20). Then the acquired methyl ester was treated with 5% hydrazine in MeOH for 12 h and the reaction mixture was dried on rotavapor before being purified by column chromatography (MeOH/DCM=1:3). The final product HA-tag was acquired at 80% yield as brown color oil. The structure of HA-tag was confirmed by NMR and/or MS (**Figures S2**).

Development and validation of SGP enrichment using peptide standards.

The SGP standard was first dissolved in 0.1 M sodium acetate to make a stock solution at 1 mg/mL. Then, 100 µl stock solution was mixed with $NaIO_4$ to the final concentration of 1 mM. HA-tag was dissolved in DMSO at 1 M and added into the solution at 1:50 ratio. The resulting mixture was desalted using Sep-Pak C18 column to remove excess reagents, before spiking into a peptide mixture (Bruker, Daltonics). Then click chemistry was performed on either Dde-Azide agarose or PEG-azide resin. Briefly, the peptide mixture was incubated with solid phase material with 0.1 mM $CuSO_4$, 0.5 mM THPTA and 5 mM sodium ascorbate for 12 h at 4 °C with end-to-end rotation. The sample was then centrifuged, and the resin was washed with PBS buffer and water twice with vortexing to remove nonspecific binding peptides. 5% hydrazine was added to the resin for elution and incubated for 4 h before the supernatant was collected and dried prior to analysis. Samples were taken in each step for MALDI or ESI-MS/MS analysis. For the isotopic labeling study, a similar workflow was conducted on SGP standard or $(Glc)_8$ as described above, isotopic reagent hydrazine sulfate ($^{15}N_2H_4$ H_2SO_4) was treated with stoichiometric NaOH to free isotopic hydrazine ($^{15}N_2H_4$) first before adding to release SGP or exchange hydrazone.

Enrichment of SGPs in bovine fetuin digest and mouse lung tissue.

1 mg of bovine fetuin was dissolved with a 6 M urea aqueous solution containing 50 mM TEAB buffer (pH 8.0). The disulfide bonds in proteins were disrupted with 200 mM DTT and the resulting solution was incubated at 56 °C for 45 min. 100 mM IAA was added subsequently and the solution was stored in the dark for 30 min at room temperature. Then the urea solution was diluted to 1M and trypsin was used to digest proteins at a ratio of 1:30 (enzyme:protein). The digestion was performed at 37 °C for 16 hours. The mouse lung tissue was prepared from male adult C57BL/6 mice. The collected mouse lung tissue was cut into small pieces and washed with 150 ml ice-cold PBS for 3 times in a dish with an ice bath. Then the tissue pieces were

homogenized and lysed in lysis buffer consisting of 4% SDS, 65 mM DTT, 150 mM NaCl, and 25 mM Tris (pH 7.4) (1 protease inhibitor tablet (Roche, Mannheim, Germany) and 1 phosphatase inhibitor tablet (Roche, Mannheim, Germany) were added in every 10 mL of the lysis buffer). The lysed sample was sonicated with 60 W energy and 5s-on-15s-off cycle for 30 times on an ice bath, and centrifuged at 3000 g for 15 min. The supernatants were transferred into 5-fold volume of ice-cold precipitation buffer (acetone: ethanol: acetic acid=50: 50: 0.1). The sample was then put into -20 °C freezer for 12 h and then was centrifuged at 3000 g at 4 °C for 15 min. The precipitated protein pellet was washed with 10 ml ice-cold precipitation buffer twice and was put in hood for 15 min to evaporate the remaining buffer. Then the pellet was re-dissolved in 8M urea and 50 mM TEAB buffer. Protein concentration was measured by BCA assay. The tryptic digestion procedure was similar to the treatment of fetuin above. The digestion was stopped by adding TFA to the final concentration of 1%. Triplicate of samples were desalted with Sep-Pak C18 columns, and the desalted peptides were dried with speed vacuum before SGP enrichment. The enrichment procedures were similar as described above.

Matrix-assisted Laser Desorption/Ionization (MALDI)-MS analysis.

Samples were prepared by premixing 1 μL of peptide/peptide mixture with 1 μL 2,5-dihydroxybenzoic acid (DHB) matrix (100 mg/mL in 5% *N*, *N*-dimethylaniline, 47.5% MeOH, and 47.5% water), and 1 μL of each matrix/sample mixture was spotted onto the MALDI target plate. RapifleX MALDI-TOF/TOF mass spectrometer (Bruker, Daltonics, Germany) was used for MALDI-MS analysis. Ionization was performed under positive mode and the laser power was adjusted manually (60% – 80% laser power) before every start of a batch process to reach a sufficient signal intensity for the internal standard (~5E4 counts). Spectra were acquired within a mass range of *m/z* 1,000-4,000 after calibration using Peptide Calibration Standard IV (Bruker,

Daltonics, Germany). FlexControl (v4.0) and FlexAnalysis (v4.0) were used for MS acquisition and data analysis.

NanoLC−MS/MS Analysis.

A self-fabricated nano-C18 column (15 cm, 75 µm i.d., 1.7 µm Ethylene Bridged Hybrid C18 packing material) was used for separation. A Dionex Ultimate 3000 nanoLC system was coupled to an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Scientific, San Jose, CA) for all LC-MS/MS analyses. Mobile phase A was 0.1% FA with 5% DMSO, and mobile phase B was ACN with 0.1% FA and 5% DMSO. The flow rate was set at 0.3 µL/min, and the injection volume was 2 µL. The following gradient was used (time, % mobile phase B) unless otherwise specified: (0 min, 5%), (28 min, 5%), (38 min, 9%), (158 min, 37%), (163 min, 95%), (173 min, 95%), (178 min, 5%), (188 min, 5%).

The following mass spectrometer parameters were used for all data acquisition. Samples were ionized in positive ion mode with a spray voltage of 3 kV. S-lens radio frequency (RF) level was set to be 30, and the capillary temperature was set to be 300 °C. Full MS scans were acquired at *m/z* 350-2000 with a resolving power of 120 K. Maximum injection time of 100 ms, automatic gain control (AGC) target value of 5e5, and 1 microscan were used for full MS scans. Top 20 data-dependent MS^2 analysis was performed at a resolving power of 15 K with stepped higher-energy collisional dissociation (HCD) operating with normalized collision energy (NCE) of 25, 28, 31. The first fixed mass sets to 150 *m/z*. The isolation window was set as 2.0 *m/z* and the dynamic exclusion of acquired precursors was set to 15 sec with a ± 20 ppm tolerance.

Data Analysis.

Byonic software (Protein Metrics, San Carlos, CA) was used to analyze the acquired MS and MS/MS spectra of enriched SGPs. Raw files were searched against bovine fetuin or *Mus musculus*

protein database of reviewed (Swiss-Prot) sequences downloaded from Uniprot. Precursor ion mass tolerance of 10 ppm and fragment ion mass tolerance of 0.01 Da were selected. The oxidation of methionine (M) was set as variable modifications, and the carbamidomethylation of cysteine (C) was set as a fixed modification. Common N-linked glycopeptide searching used a mammalian N-glycan database that contains 309 glycans. In addition, the sialylated glycans were adjusted with corresponding trimmed mass (e.g., +225.0745 for one NeuAc or +241.0694 for one NeuGc) in the glycan database. Peptide identifications were filtered at two-dimensional false discovery rate (2D FDR) <1%, PEP 2D <0.05, |Log Prob| > 1, and Byonic Score > 100. Manual inspection of the MS/MS spectra of SGPs was performed to examine if Byonic identification results contained diagnostic ions. GO functional category enrichment was analyzed using DAVID bioinformatics resources[28] with an FDR cutoff of 0.05. Protein-protein interaction (PPI) enrichment analyses were generated using Metascape (version 3.5)[29].

Results and Discussion

Traditional hydrazide chemistry-based methods show high SGPs enrichment selectivity at the cost of losing the SA glycan information[22]. The good selectivity was achieved by specific modification of SAs, yet, the SAs were not retained in the final glycopeptides. To circumvent this limitation, we developed a novel enrichment method incorporating click chemistry and dynamic covalent exchange to achieve high selectivity while preserving SAs on the intact glycopeptides at the same time.

Development of SGPs Enrichment Method.

Previously, hydrazide chemistry has been applied to SGPs enrichment, which is based on the fact that the vicinal diol groups on SAs can be selectively oxidized by mild periodate oxidation

to an aldehyde group and then be captured onto hydrazide beads. After SGPs are enriched from the solid phase, taking advantage of that the glycosidic bond between the terminal sialic acid and the penultimate monosaccharide is sensitive to mild acid hydrolysis, the captured glycopeptides can be selectively released by acid hydrolysis. However, this method loses the information of terminal SAs[22]. Considering the decent selectivity of hydrazide chemistry, we adopted its chemical modification procedure, but instead of using hydrazide beads to conjugate modified SGPs, a hydrazide-alkyne tag (HA-tag) was used for probing SGPs while installing an alkyne handle at the same time. The alkyne handle provides a modification site for the subsequent copper-catalyzed azide/alkyne cycloaddition (CuAAC) enabled click chemistry, which is a specific and controllable bio-orthogonal reaction widely applied in the detection, localization and quantification of biomolecules[30, 31]. Its high yield and specificity allow it to inherit the selectivity of hydrazide chemistry, i.e., the selectivity towards SGPs. Our rationale is by extending modified moiety on SAs through click chemistry, a cleavable linker could be introduced which would be cleaved under a non-acidic condition after enrichment, so the labile SAs will be retained on glycopeptides. In our original design (**Figure S3**), an azide resin that contains a cleavable Dde linker was utilized. We envisioned that once alkyne modified SGPs are covalently captured to the resin, the resin can be washed with the highest stringency virtually eliminating any non-specifically bound peptides, and treatment of the resin with 5% hydrazine will break the Dde moiety and SGPs will be released without losing SAs.

We first used an SGP peptide standard to test our strategy. As shown in **Figure 2A and 2B**, after treating peptide standard with 1 mM $NaIO_4$ at 4 ℃ for 30 min, the complete oxidized form of SGP was observed on the MALDI-TOF MS spectrum. The newly generated carbonyl groups were then labeled with HA-tag at 1:50 ratio subsequently and this reaction also exhibited

complete derivatization efficiency (**Figure 2C, Figure S5a**). Since the SAs are usually labile in MALDI source[32], peaks corresponding to in-source fragmentation of one SA were noted, which validate the oxidation and HA-tag derivatization on SA from the other side. To evaluate the enrichment efficiency of our method, HA-tag labeled SGP standards were spiked into a peptide mixture (**Figure 2D and 2E, Table S1**) before performing click chemistry with Dde-Azide-resin and releasing by 5% hydrazine. Despite prominent signal with the clear background was acquired after enrichment, the mass of which did not match what we had anticipated (**Figure 2F**). What we expected to see is a peak at *m/z* 3217 (**Figure S5b**), however, the result we obtained had a peak at *m/z* 2857. Trimmed mass indicated unexpected bond break within the extended structure. When sorting possible fragments, we found the mass of the observed peak matched well to another hydrazone structure (**Figure S5c**). Therefore, we reasoned that there was a dynamic covalent switch occurred between the two Schiff-bases (**Figure S4**), as hydrazine seemed to have higher reactivity towards carbonyl group than hydrazide[33].

This hypothesis was then confirmed and validated by using isotopic hydrazine in the release step. It is noted that our unexpected discovery not only makes no difference in original design and rationale but also helps to reduce the cost, by substituting cleavable Dde-Azide agaroses with a normal PEG-azide resin in the workflow (**Figure 1**).

Isotopic Study Confirms Dynamic Covalent Exchange.

An interesting question arises: does hydrazine really induce Schiff-base swap? To answer this question, a glycan standard maltooctaose ($(Glc)_8$) with a reducing end was introduced first for hypothesis testing in solution phase. The reducing end of $(Glc)_8$ is a hemiacetal group and in aqueous solutions, it exists with an aldehyde form in equilibrium. When treating $(Glc)_8$ with HA-tag in 1:50 ratio, the formation of hydrazone between the open-ring form of $(Glc)_8$ and HA-tag

constantly consumed aldehyde species and shifted equilibrium. Eventually, almost all $(Glc)_8$ were converted to HA-tag labeled form (**Figure S6b**). This process was similar to the labeling step on SGPs after periodate oxidation. Then, without removing excess HA-tag in the solution, non-isotopic hydrazine N_2H_4 or isotopic hydrazine $^{15}N_2H_4$ were added in the same amount. We observed the disappearance of HA-tag labeled $(Glc)_8$ with emerging peaks corresponding to hydrazine labeled $(Glc)_8$. In addition, the sample treated with $^{15}N_2H_4$ clearly showed a 2 Da mass increase compared to the one treated with N_2H_4 (**Figure S6c, d**). This result shows that the carbonyl group (or aldehyde) prefers hydrazine rather than hydrazide when forming hydrazone, and through chemical equilibrium, hydrazine is able to convert hydrazide Schiff-base to hydrazone.

Further, we validated hydrazine-induced covalent exchange on SGP standard following the enrichment workflow mentioned above. Identical preparation steps were kept for two samples before they were released by either N_2H_4 or $^{15}N_2H_4$. As shown in **Figure 3A and 3B**, when enriched SGP was released by isotopic hydrazine $^{15}N_2H_4$, the signal at *m/z* 2860 has a 4 Da mass increase compared to the non-isotopic reagent treated sample. This mass increment corresponds to two SAs per SGP where each SA was exchanged by one molecule of hydrazine. Furthermore, in the typical MS/MS spectra of enriched SGP, we found the generation of specific glycan fragment ions as well as peptide fragments with and without glycan residues (**Figure 3C and 3D**). Fragments at *m/z* 244 and *m/z* 246 closely match to the proposed structures of modified SA (**Figure S5d, e**), and the fragmentation pattern is in agreement with a previous study[34]. These results demonstrated the facile swap of Schiff-base induced by hydrazine in the enrichment procedure and it seems likely that these signature signals in MS/MS spectra can be used for rapid identification of enriched SGPs in multiple reaction monitoring (MRM).

Moreover, we incubated tryptic digested peptide mixtures from HEK293T cell lines, a common quality control (QC) sample in our lab, in either water or 5% hydrazine for two hours and analyzed sequentially. Similar total ion chromatogram (TIC) and peptide identification indicates mild basic condition under 5% hydrazine for 2hr does not affect the integrity of the proteome (**Figure S7**).

SGP Enrichment from Bovine Fetuin and Mouse Lung Tissue.

The selectivity of the developed enrichment protocol for SGPs was evaluated by using the tryptic digestion of bovine fetuin as model protein samples. Fetuin is known as a SA-rich glycoprotein whose role as a carrier of bioactive molecules has been proposed based on observations that it binds and carries Ca^{2+} ion[35]. Triplicate parallel experiments of glycopeptide enrichment were performed by using 100 μg of bovine fetuin protein digest (based on Pierce Peptide Assay). Based on the similar optimized protocol described above for enriching SGP standard peptide, 389 SGPs were identified out of the total 545 peptides after enrichment while before enrichment, there were more than 1100 total peptides and SGPs constituted less than 40%. The sialylglycopeptide enrichment selectivity, which is defined as the ratio of the identified SGPs to the total number of peptides detected by the MS, of approximately 72% was obtained after enrichment, increased by almost 2-fold compared to prior-enrichment condition (**Figure 4A**, **Table S2**). In addition, 343, 354, and 337 glycoforms were identified in the 3 biological replicates respectively, and significant overlap (74%) of the identified glycopeptides among these replicates were obtained (**Figure 4B**, **Table S2**), indicating good reproducibility of the method. Moreover, the enrichment and release condition of our method was mild and intact structural information on the glycans could be retained in comparison with the hydrazide chemistry method (**Figure S8**).

To evaluate the enrichment selectivity and sensitivity, we used equal amounts of proteins but with different ratios of bovine serum albumin (BSA) / bovine fetuin, and performed the digestion and enrichment as mentioned above. BSA is a non-glycoprotein and provides an interference background for SGPs enrichment. Three prominent SGPs with the masses ranging from *m/z* 3000 to 4000 were monitored by MALDI-MS. Decreased signals and S/N corresponding to the enriched species were recorded with the increasing of BSA/bovine fetuin ratio, however, under BSA/bovine fetuin=100:1, the enriched SGPs were still observable (**Figure S9**).

In addition, the enrichment performance was compared with hydrazide chemistry. The hydrazide chemistry was performed according to the "reverse glycoblotting" method with slight modification[21]. In brief, hydrazide beads (Bio-Rad) were washed with PBS three times before mixing with sodium-periodate-oxidized bovine fetuin digests. 0.1 M formic acid was added onto the beads after the beads was washed thoroughly and incubated at 80°C for 1h. The released (glyco)peptides in supernatant were then collected and desalted before LC-MS/MS analysis. 367 glycopeptides were identified among 474 peptides in total, accounting for around 80% efficiency for glycopeptide enrichment (**Table S2**). However, it is noted that the SAs were hydrolyzed in the release step and therefore the intact SGPs among these enriched peptides were hardly determined. We arbitrarily assume that among these glycopeptides, if there is at least one of the SGP identified confidently by our method shared the same peptide sequences and glycosylation sites but with extended glycan structures, then this glycopeptide is counted for SGP. In this case, a sketchy comparison can be made: 280 SGPs were identified using both methods, and the overlap was close to 60%. The obvious advantage of our method is the SAs are preserved in intact peptides and therefore the SGPs can be determined with high confidence. However, it is worth mentioning that

careful control of oxidation condition is crucial for the success of our experiments as excessive oxidation may result in false discovery and sample loss.

The developed enrichment method was further applied to a much more complex biological sample, the mouse lung tissue extract. The enrichment was performed after the tissue was homogenized and the proteins were extracted and digested. A total of 1350 SGPs from 136 glycoproteins were successfully characterized with detailed glycan structure information as well as modification sites (**Table S3)**. The gene ontology (GO) annotation analysis of molecular function (MF), biological process (BP) and cellular component (CC) was conducted to gain a better functional understanding of the 136 identified glycoproteins. We noticed that these proteins are mainly involved in cell adhesion through binding to various molecules and they are widely distributed on the cell surface and extracellular exosome, which is in accordance with the common features of sialylated proteins (**Figure 4C**). Moreover, protein-protein interaction (PPI) enrichment analysis has been carried out using the Metascape platform. The Molecular Complex Detection (MCODE) algorithm has been applied to identify densely connected network components[36], and the MCODE networks identified have been gathered and are shown in **Figure 4D**. The PPI was mainly concentrated in the relevance among components of ECM-receptor interaction, lysosomal transport, and neutrophil degranulation.

Conclusions

In summary, a facile, selective, and easy access enrichment method of SGPs based on commercially available materials was developed. The intact SGPs could be simply enriched and separated from protein digests by hydrazide and click chemistry as well as hydrazine-induced dynamic covalent exchange. The detailed Schiff-base swap facilitated by hydrazine was confirmed

by isotopic study and the unique fragments produced in MS/MS can be used for unambiguously identification of SGPs. The method exhibited high enrichment efficiency and reproducibility and was successfully applied to enrichment of SGPs using peptide/protein standard and real biological samples. We anticipate that this new enrichment approach will be widely applicable for in-depth analysis of sialylation glycans or glycopeptides, which will aid in candidate biomarker discovery and medical diagnostics.

Acknowledgments

This work was funded in part by the National Institutes of Health (NIH) grants RF1 AG052324, U01CA231081, and R01 DK071801. The MALDI TOF/TOF RapifleX mass spectrometer was purchased through the support of an NIH shared instrument grant S10OD025084. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531) and Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. LL acknowledges NIH grant support R21AG065728 as well as 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. Dell, A.; Morris, H. R., Glycoprotein structure determination mass spectrometry. *Science* **2001,** *291* (5512), 2351-2356.

2. Ghazarian, H.; Idoni, B.; Oppenheimer, S. B., A glycobiology review: Carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochem.* **2011,** *113* (3), 236-247.

3. Ohtsubo, K.; Marth, J. D., Glycosylation in cellular mechanisms of health and disease. *Cell* **2006,** *126* (5), 855-867.

4. Kohnz, R. A.; Roberts, L. S.; DeTomaso, D.; Bideyan, L.; Yan, P.; Bandyopadhyay, S.; Goga, A.; Yosef, N.; Nomura, D. K., Protein Sialylation Regulates a Gene Expression Signature that Promotes Breast Cancer Cell Pathogenicity. *ACS Chem. Biol.* **2016,** *11* (8), 2131-2139.

5. Wasik, B. R.; Barnard, K. N.; Parrish, C. R., Effects of Sialic Acid Modifications on Virus Binding and Infection. *Trends Microbiol.* **2016,** *24* (12), 991-1001.

6. Wang, B., Molecular Mechanism Underlying Sialic Acid as an Essential Nutrient for Brain Development and Cognition. *Adv. Nutr.* **2012,** *3* (3), 465S-472S.

7. Pearce, O. M. T.; Laubli, H., Sialic acids in cancer biology and immunity. *Glycobiology* **2016,** *26* (2), 111-128.

8. Zhang, Z. J.; Wuhrer, M.; Holst, S., Serum sialylation changes in cancer. *Glycoconjugate J.* **2018,** *35* (2), 139-160.

9. Lu, H. J.; Zhang, Y.; Yang, P. Y., Advancements in mass spectrometry-based glycoproteomics and glycomics. *Natl. Sci. Rev.* **2016,** *3* (3), 345-364.

10. Zhang, Q. W.; Li, Z.; Wang, Y. W.; Zheng, Q.; Li, J. J., Mass spectrometry for protein sialoglycosylation. *Mass Spectrom. Rev.* **2018,** *37* (5), 652-680.

11. Zhao, J.; Simeone, D. M.; Heidt, D.; Anderson, M. A.; Lubman, D. M., Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: Application to pancreatic cancer serum. *J. Proteome Res.* **2006,** *5* (7), 1792-1802.

12. Lee, A.; Nakano, M.; Hincapie, M.; Kolarich, D.; Baker, M. S.; Hancock, W. S.; Packer, N. H., The Lectin Riddle: Glycoproteins Fractionated from Complex Mixtures Have Similar Glycomic Profiles. *OMICS* **2010,** *14* (4), 487-499.

13. Zhu, F. F.; Clemmer, D. E.; Trinidad, J. C., Characterization of lectin binding affinities via direct LC-MS profiling: implications for glycopeptide enrichment and separation strategies. *Analyst* **2017,** *142* (1), 65-74.

14. Palmisano, G.; Lendal, S. E.; Engholm-Keller, K.; Leth-Larsen, R.; Parker, B. L.; Larsen, M. R., Selective enrichment of sialic acid-containing glycopeptides using titanium dioxide chromatography with analysis by HILIC and mass spectrometry. *Nat. Protoc.* **2010,** *5* (12), 1974-1982.

15. Sun, N. R.; Deng, C. H.; Li, Y.; Zhang, X. M., Highly Selective Enrichment of N-Linked Glycan by Carbon-Functionalized Ordered Graphene/Mesoporous Silica Composites. *Anal. Chem.* **2014,** *86* (4), 2246-2250.

16. Sun, N. R.; Wang, J. W.; Yao, J. Z.; Deng, C. H., Hydrophilic Mesoporous Silica Materials for Highly Specific Enrichment of N-Linked Glycopeptide. *Anal. Chem.* **2017,** *89* (3), 1764-1771.

17. Chen, L. L.; Dong, X. F.; Cao, L. W.; Guo, Z. M.; Yu, L.; Zou, L. J.; Liang, X. M., Hydrophilic interaction/cation-exchange chromatography for glycopeptide enrichment by using a modified strong-cation exchange material. *Anal. Methods* **2013,** *5* (24), 6919-6924.

18. Lewandrowski, U.; Zahedi, R. P.; Moebius, J.; Walter, U.; Sickmann, A., Enhanced N-glycosylation site analysis of Sialoglycopeptides by strong cation exchange prefractionation applied to platelet plasma membranes. *Mol. Cell. Proteomics* **2007,** *6* (11), 1933-1941.

19. Caval, T.; Zhu, J.; Heck, A. J. R., Simply Extending the Mass Range in Electron Transfer Higher Energy Collisional Dissociation Increases Confidence in N-Glycopeptide Identification. *Anal. Chem.* **2019,** *91* (16), 10401-10406.

20. Kurogochi, M.; Amano, M.; Fumoto, M.; Takimoto, A.; Kondo, H.; Nishimura, S., Reverse glycoblotting allows rapid-enrichment glycoproteomics of biopharmaceuticals and disease-related biomarkers. *Angew. Chem., Int. Ed.* **2007,** *46* (46), 8808-8813.

21. Nilsson, J.; Ruetschi, U.; Halim, A.; Hesse, C.; Carlsohn, E.; Brinkmalm, G.; Larson, G., Enrichment of glycopeptides for glycan structure and attachment site identification. *Nat. Methods* **2009,** *6* (11), 809-U26.

22. Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R., Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **2003,** *21* (6), 660-666.

23. Dong, X. F.; Qin, H. Q.; Mao, J. W.; Yu, D. P.; Li, X. L.; Shen, A. J.; Yan, J. Y.; Yu, L.; Guo, Z. M.; Ye, M. L.; Zou, H. F.; Liang, X. M., In-Depth Analysis of Glycoprotein Sialylation in Serum Using a Dual-Functional Material with Superior Hydrophilicity and Switchable Surface Charge. *Anal. Chem.* **2017,** *89* (7), 3966-3972.

24. Zhang, M.; Liu, Y. J.; Zhang, D.; Chen, T. J.; Li, Z. L., Facile and Selective Enrichment of Intact Sialoglycopeptides Using Graphitic Carbon Nitride. *Anal. Chem.* **2017,** *89* (15), 8064-8069.

25. Wan, H. H.; Zhang, X. F.; Chen, C.; Li, X. L.; Liang, X. M., Selective enrichment of sialylated glycopeptides with mesoporous poly-melamine-formaldehyde (mPMF) material. *Anal. Bioanal. Chem.* **2020,** *412* (7), 1497-1508.

26. Sun, N.; Xiong, Y. T.; Qing, G. Y.; Zhao, Y. Y.; Li, X. L.; Liang, X. M., Selective enrichment of sialylated glycopeptides with a d-allose@SiO2 matrix. *RSC Adv.* **2018,** *8* (68), 38780-38786.

27. Xiong, Y. T.; Li, X. L.; Li, M. M.; Qin, H. J.; Chen, C.; Wang, D. D.; Wang, X.; Zheng, X. T.; Liu, Y. H.; Liang, X. M.; Qing, G. Y., What Is Hidden Behind Schiff Base Hydrolysis? Dynamic Covalent Chemistry for the Precise Capture of Sialylated Glycans. *J. Am. Chem. Soc.* **2020,** *142* (16), 7627-7637.

28. Huang, D. W.; Sherman, B. T.; Lempicki, R. A., Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009,** *4* (1), 44-57.

29. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A. H.; Tanaseichuk, O.; Benner, C.; Chanda, S. K., Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **2019,** *10* (1), 1523.

30. Chio, T. I.; Bane, S. L., Click Chemistry Conjugations. *Antibody-Drug Conjugates: Methods and Protocols* **2020,** *2078*, 83-97.

31. Meldal, M.; Diness, F., Recent Fascinating Aspects of the CuAAC Click Reaction. *Trends Chem.* **2020,** *2* (6), 569-584.

32. Nishikaze, T.; Kawabata, S.; Tanaka, K., In-Depth Structural Characterization of N-Linked Glycopeptides Using Complete Derivatization for Carboxyl Groups Followed by Positive- and Negative-Ion Tandem Mass Spectrometry. *Anal. Chem.* **2014,** *86* (11), 5360-5369.

33. Shigeri, Y.; Ikeda, S.; Yasuda, A.; Ando, M.; Sato, H.; Kinumi, T., Hydrazide and hydrazine reagents as reactive matrices for MALDI-MS to detect gaseous aldehydes. *J. Mass Spectrom.* **2014,** *49* (8), 742-749.

34. Kurogochi, M.; Matsushista, T.; Amano, M.; Furukawa, J.; Shinohara, Y.; Aoshima, M.; Nishimura, S. I., Sialic Acid-focused Quantitative Mouse Serum Glycoproteomics by Multiple Reaction Monitoring Assay. *Mol. Cell. Proteomics* **2010,** *9* (11), 2354-2368.

35. Wang, H. C.; Zhang, M. H.; Bianchi, M.; Sherry, B.; Sama, A.; Tracey, K. J., Fetuin (alpha(2)-HS-glycoprotein) opsonizes cationic macrophage-deactivating molecules. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95* (24), 14429-14434.

36. Bader, G. D.; Hogue, C. W., An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinf.* **2003,** *4*.

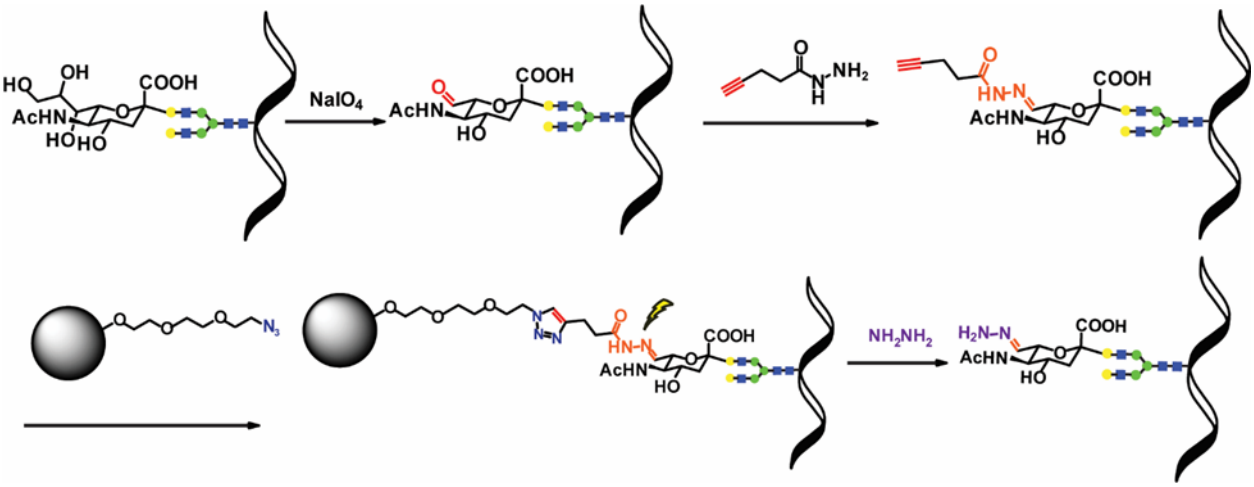


Figure 1. Scheme of sialylglycopeptide (SGP) enrichment enabled by periodate oxidation, click chemistry and dynamic covalent exchange; SGPs were first selectively oxidized to expose an aldehyde group, which was conjugated to an alkyne probe through the reaction with hydrazide, then click chemistry was performed to link the modified SGPs onto the azide resin, where the intact SGPs were covalently exchanged by hydrazine during the elution.

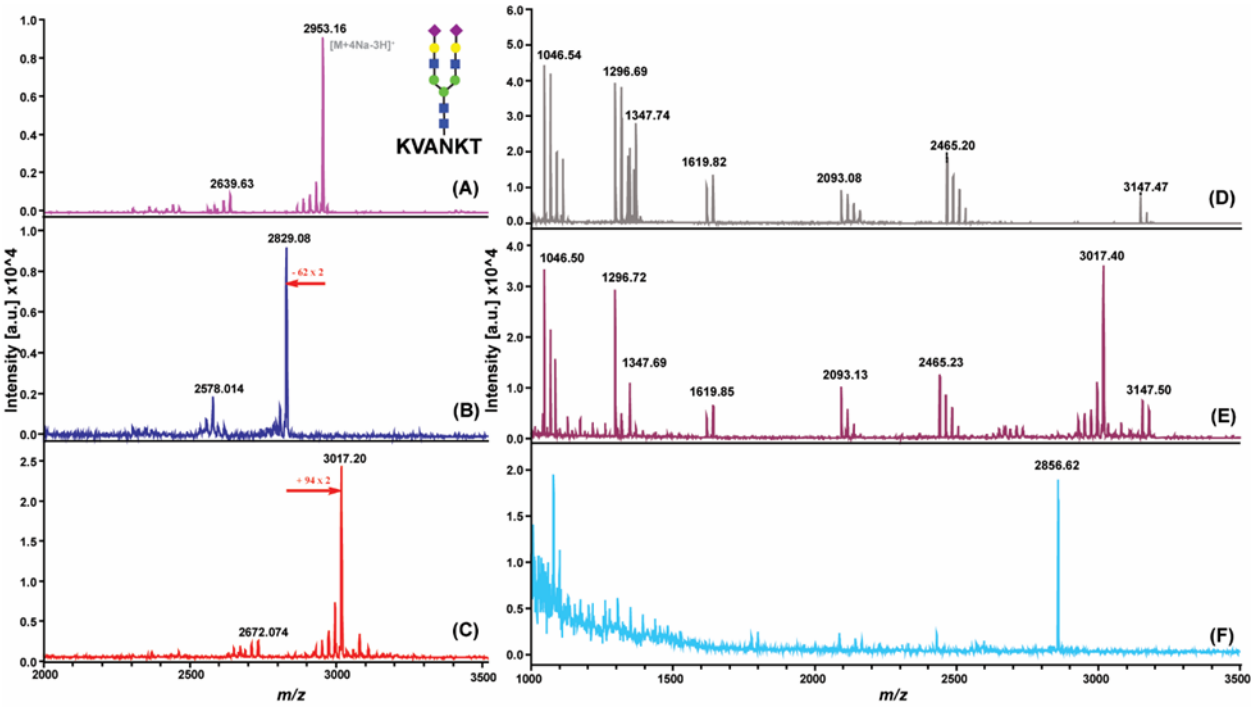


Figure 2. Establishment of SGP enrichment method using peptide standard; (A) MALDI-MS spectrum of SGP peptide standard native state; (B) after periodate oxidation; and (C) after HA-tag labeling; Enrichment efficiency test; (D) MALDI-MS spectrum of peptide mixture; (E) HA-tag labeled SGP standard spiked into peptide mixture; (F) after enrichment. (The most abundant peaks were labeled; multiple sodium adducts were observed for SGP standard peptide)

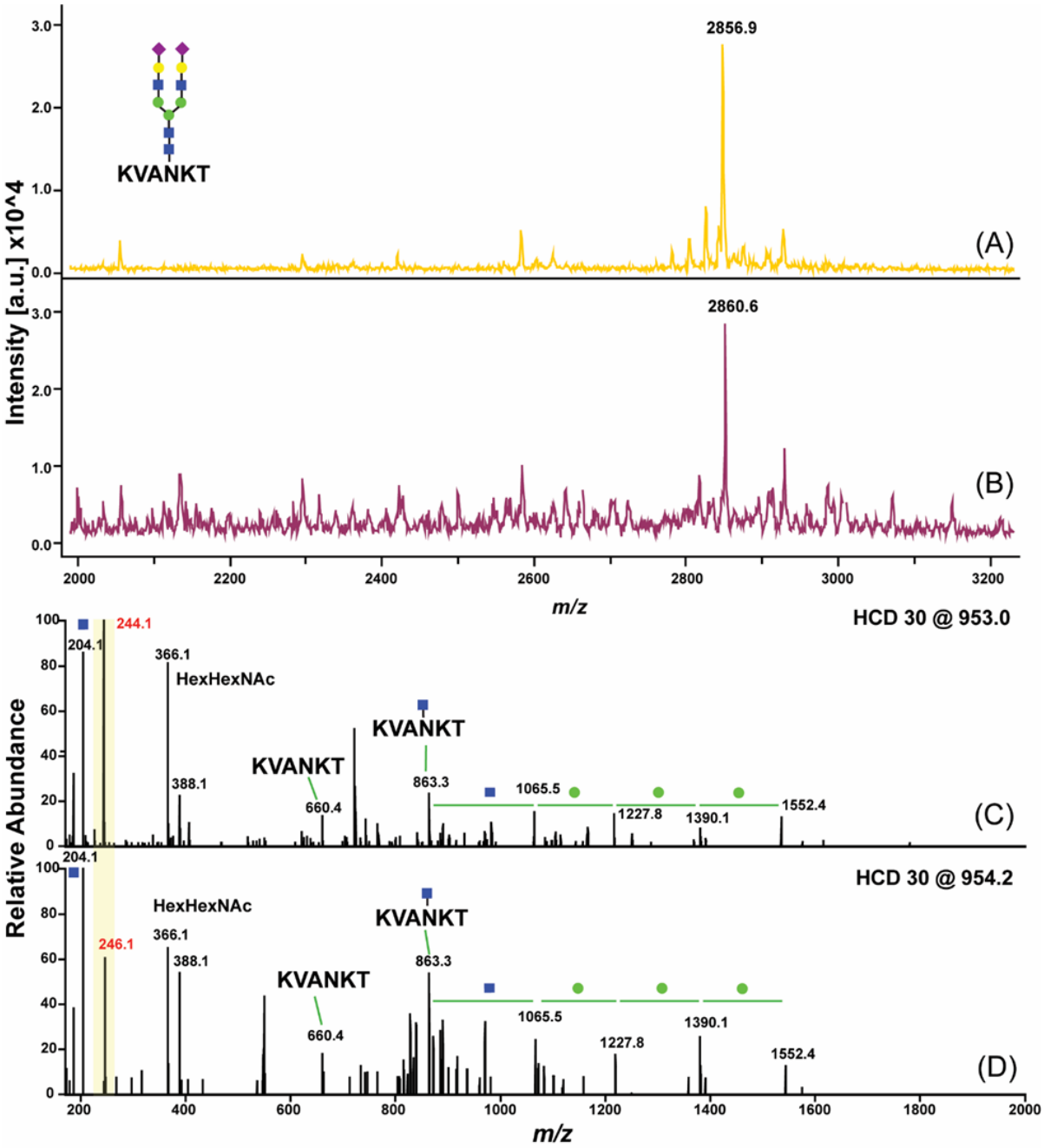


Figure 3. Isotopic study for hydrazine release. (A, C) MALDI-TOF and ESI-MS/MS spectra of SGP standard released by N_2H_4; (B, D) MALDI-TOF and ESI-MS/MS spectra of SGP standard released by $^{15}N_2H_4$.

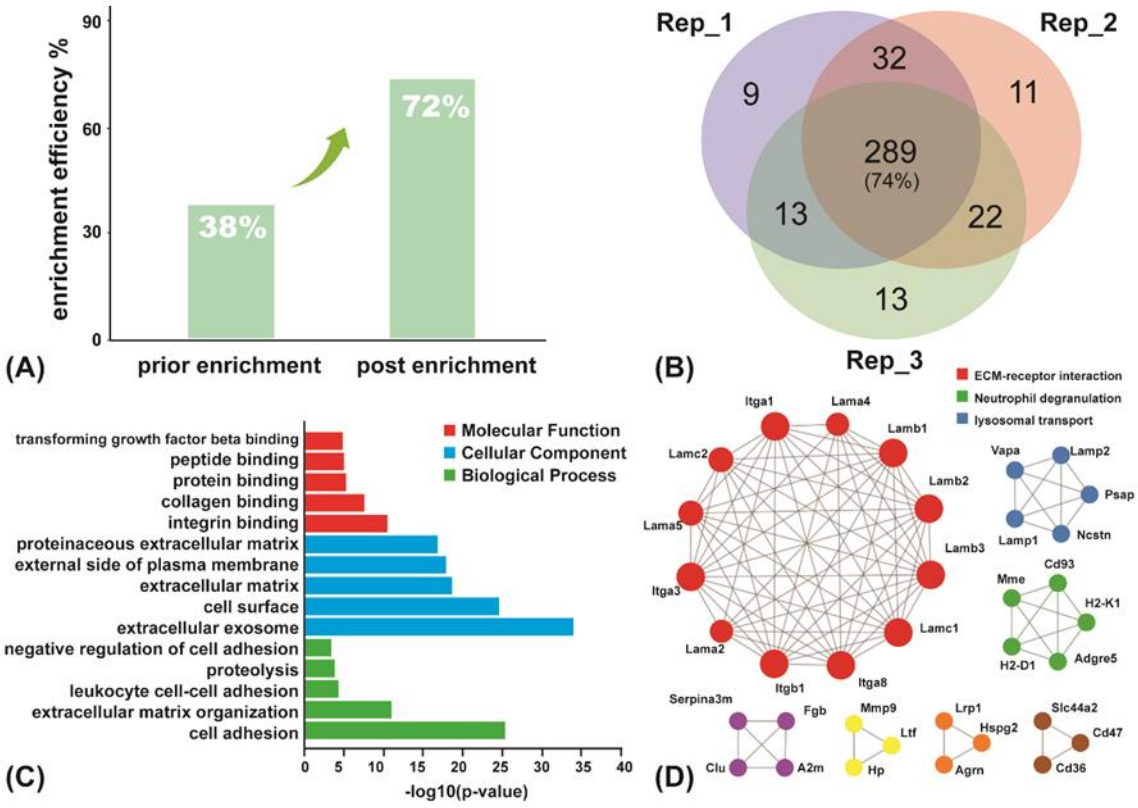


Figure 4. SGPs enrichment from digested bovine fetuin protein and proteins extracted from mouse lung tissue. (A) The SGPs enrichment efficiency on bovine fetuin; (B) Venn diagram of the identified SGPs from three replicated experiments; (C) GO analysis of SGPs enriched from mouse lung tissue. Top 5 most significant ($P<0.05$) categories in BP, MF and CC branches were plotted; (D) MCODE PPI network. Colors represent the different components of MCODE.

Supplemental Information

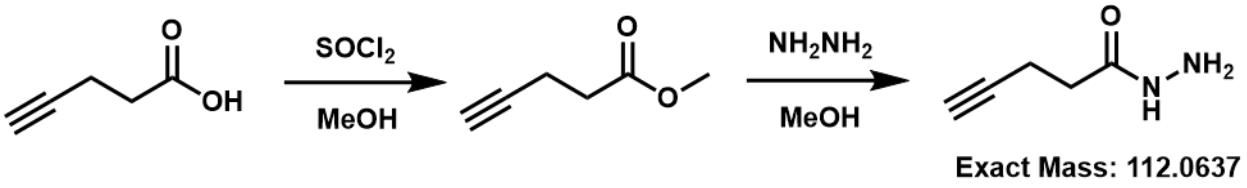


Figure S1. Synthetic route of Hydrazine-Alkyne (HA) tag.

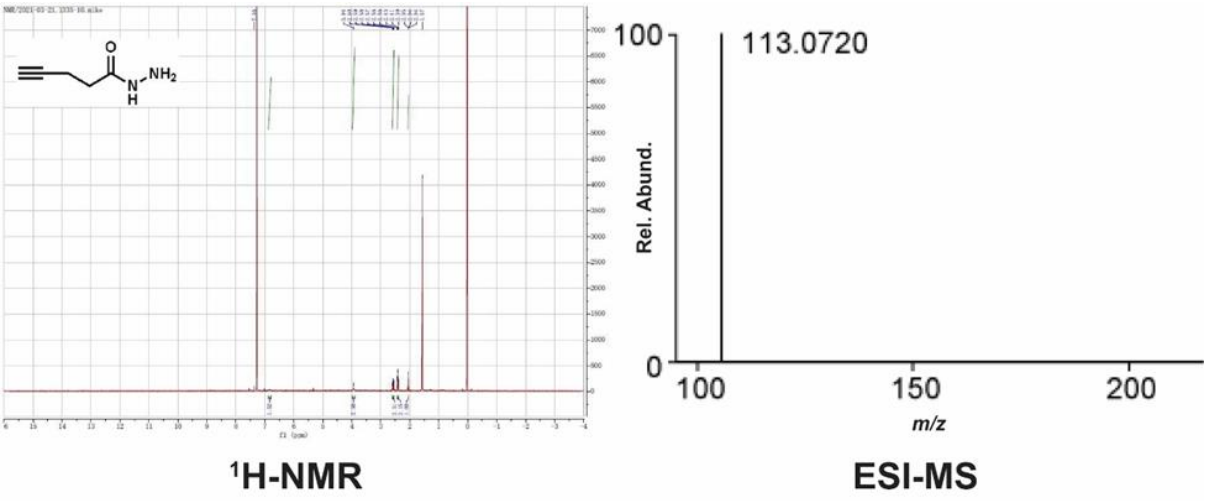


Figure S2. ^{1}H-NMR and high-resolution MS of Hydrazine-Alkyne (HA) tag.

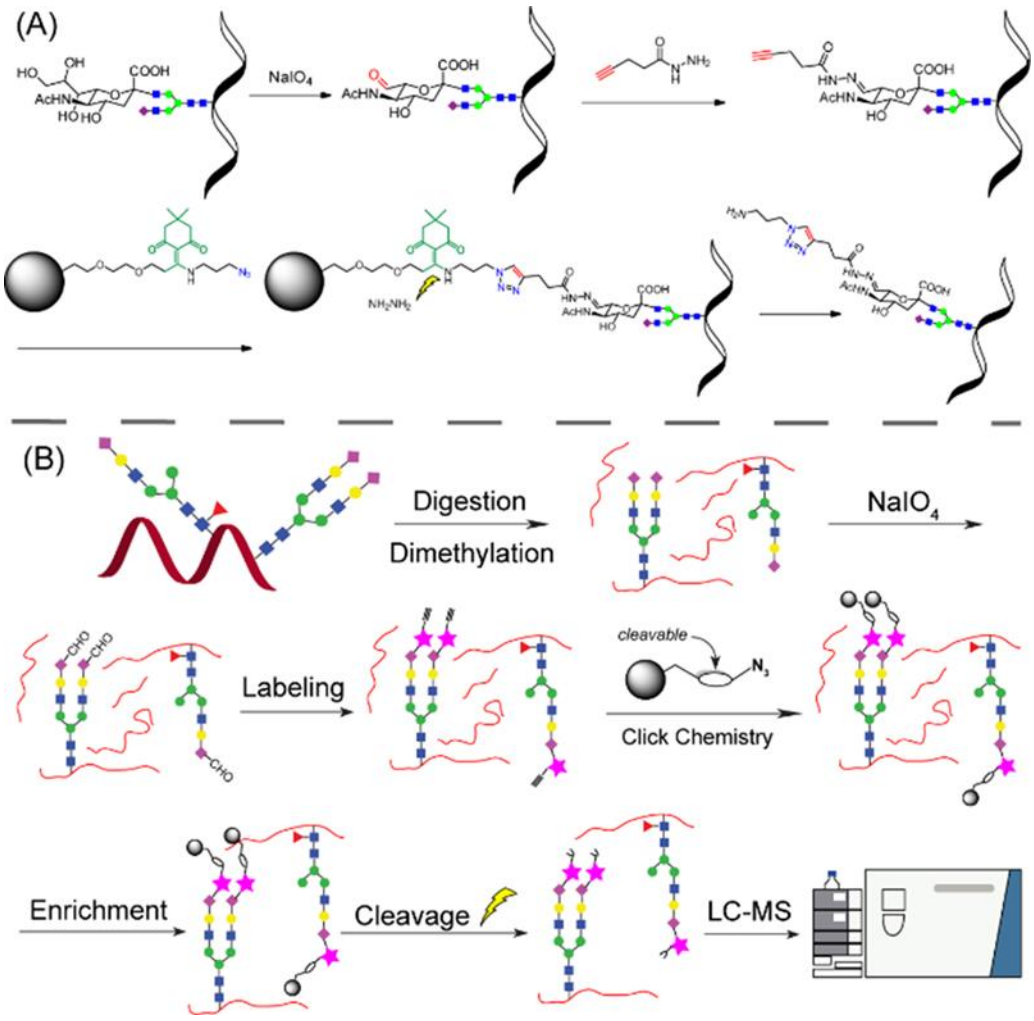


Figure S3. Original design: (A) Scheme of SGP enrichment enabled by periodate oxidation, click chemistry and hydrazine release. (B) Workflow of SGP enrichment sample preparation.

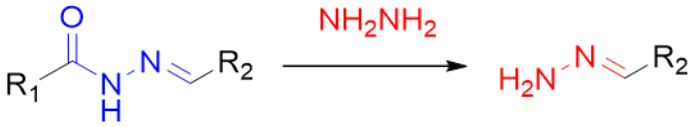


Figure S4. Scheme of Schiff-base dynamic covalent switch between hydrazones.

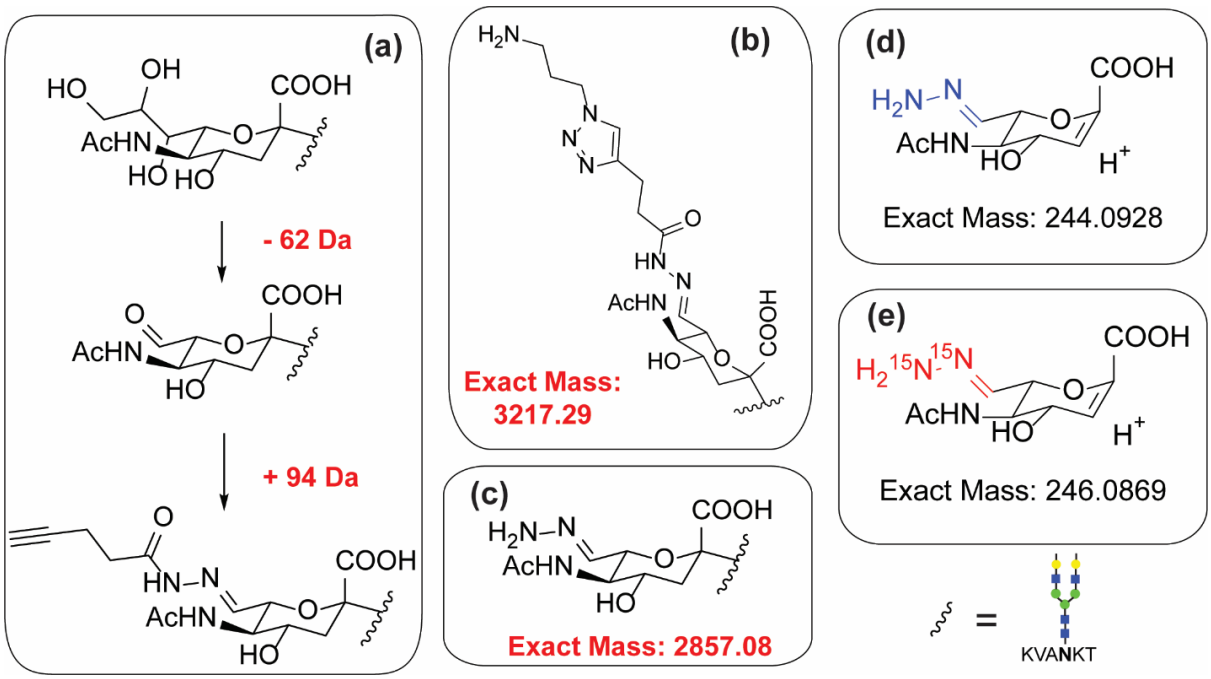


Figure S5. (a) Structures of SGP standard during the process of enrichment; (b) Structure of released SGP in original design; (c) Structure of released SGP in actual case; (d, e) Structures of sialic acid moiety released by N_2H_4 or $^{15}N_2H_4$.

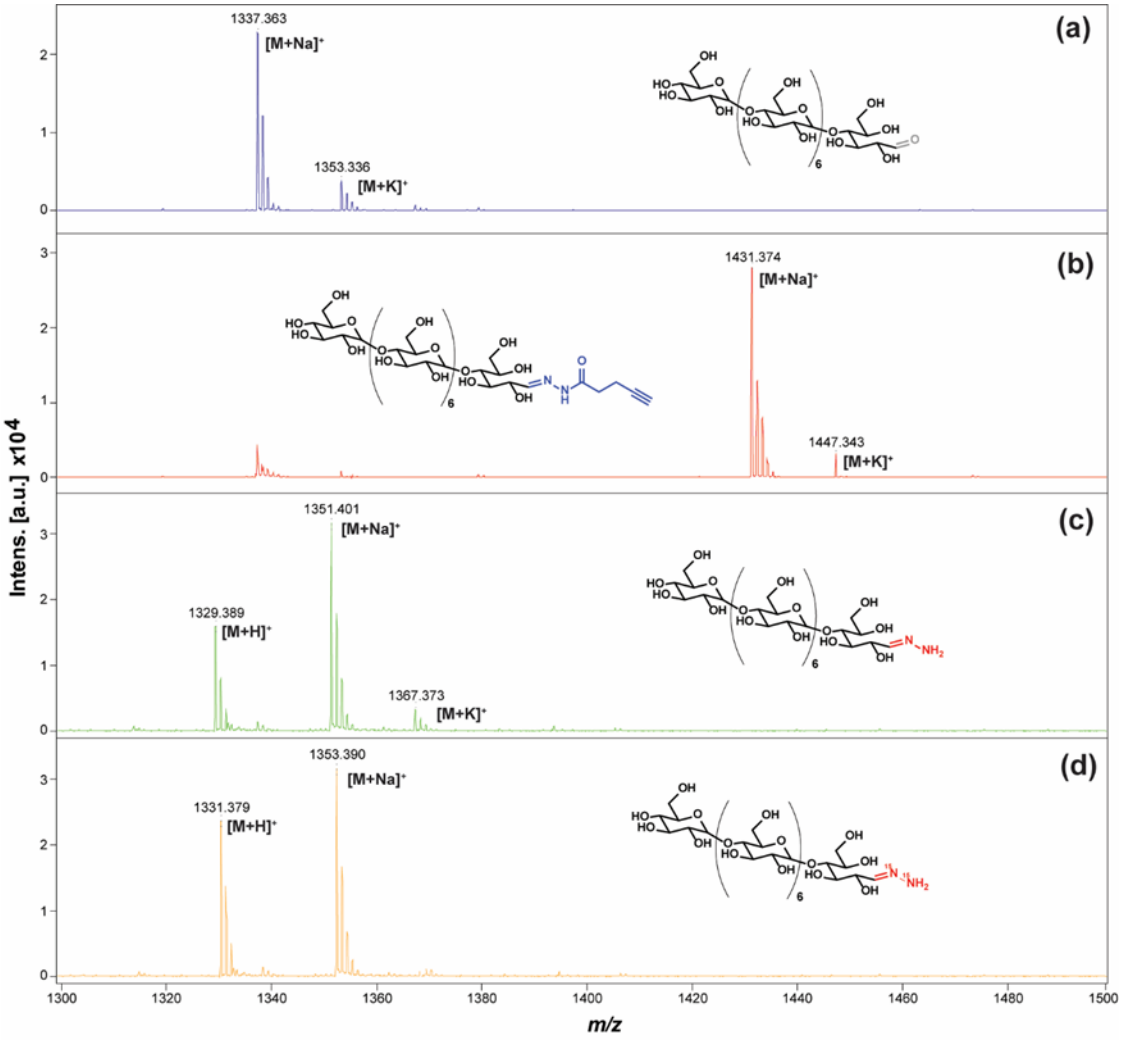


Figure S6. Isotopic study of hydrazine release using $(Glc)_8$; (a) Native $(Glc)_8$; (b) HA tag labeled $(Glc)_8$; (c, d) HA tag labeled $(Glc)_8$ treated with N_2H_4 or $^{15}N_2H_4$.

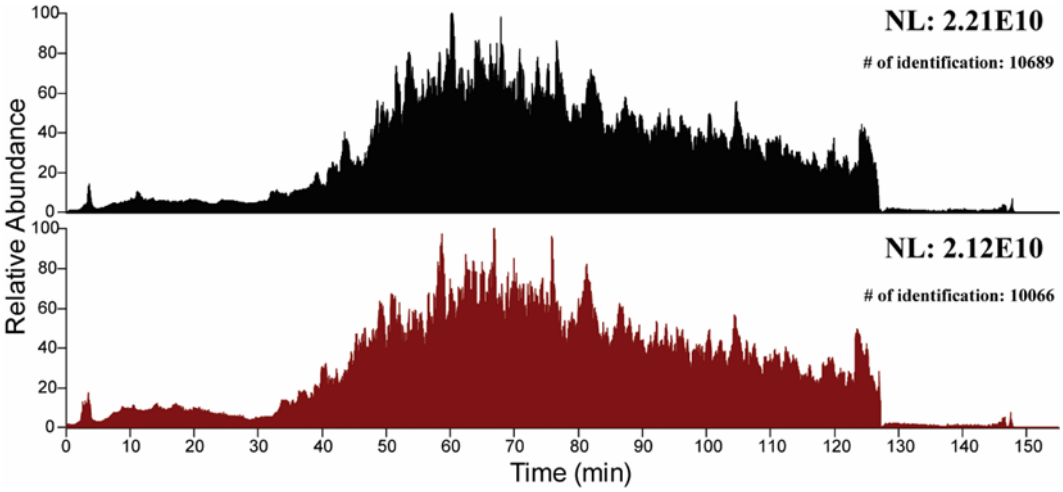


Figure S7. Incubation of the proteome from HEK293T cell line in either water (top) or 5% hydrazine (bottom) for 2h.

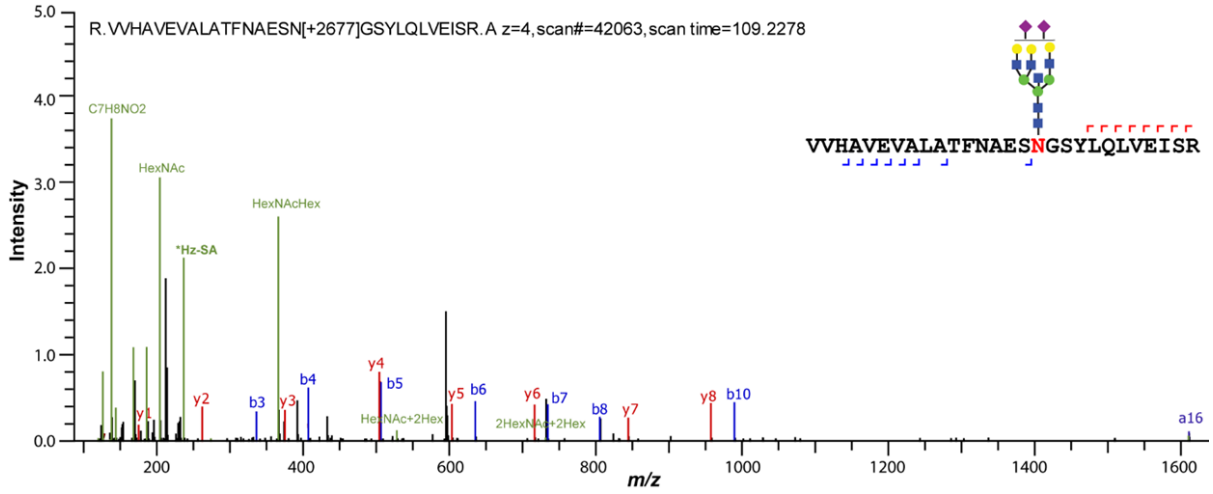


Figure S8. Representative MS/MS spectrum of an SGP enriched from mouse lung tissue extract.

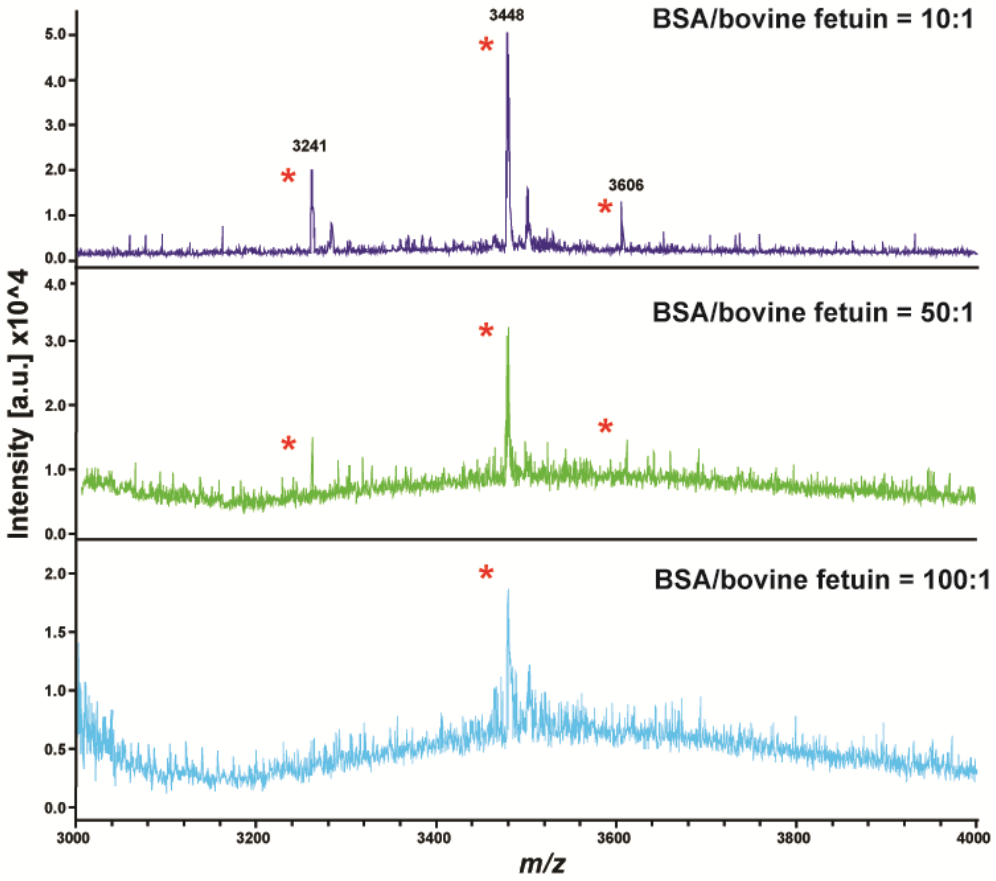


Figure S9. Enrichment selectivity and sensitivity evaluation at different BSA/bovine fetuin ratios assessed by MALDI-TOF MS. Mass spectral peaks labeled with red asterisks were major SGPs from glycoprotein bovine fetuin digest.

Peptides	*m/z* [M+H]$^+$	Sequences
Angiotensin_II	**1046.54180**	**DRVYIHPF**
Angiotensin_I	**1296.68480**	**DRVYIHPFHL**
Substance_P	**1347.73540**	**RPKPQQFFGLM**
Bombesin	**1619.82230**	**PryQRLGNGWAVGHLM**
ACTH_clip(1-17)	**2093.08620**	**SYSMEHFRWGKPVGKKR**
ACTH_clip(18-39)	**2465.19830**	**RPVKVYPNGAEDESAEAFPLEF**
Somatostatin	**3147.47100**	**SANSNPAMAPRERKAGCKNFFWKTFTSC**

Table S1. A list of peptides in peptide mixtures.

Chapter 5

Novel Isobaric Tagging Reagent Enabled Multiplex Quantitative Glycoproteomics via Electron-Transfer/Higher-Energy Collisional Dissociation (EThcD) Mass Spectrometry

Adapted from: **Miyang Li**[#], Xiaofang Zhong[#], Yu Feng, and Lingjun Li, Novel Isobaric Tagging Reagent Enabled Multiplex Quantitative Glycoproteomics via Electron-Transfer/Higher-Energy Collisional Dissociation (EThcD) Mass Spectrometry. *Ready for Submission.* (Author contribution: study was designed by M. Li., Y. Feng; experiment was performed by M. Li, X. Zhong; data was analyzed by M. Li., X. Zhong.; manuscript was written by M. Li and edited by X. Zhong, Y. Feng, L. Li.)

Abstract

Glycosylation, one of the most important post-translational modifications, plays a pivotal role in many essential physiological processes, including cellular signal transduction and molecular trafficking. Comprehensive site-specific and quantitative analysis is crucial for revealing the diverse functions and dynamics of glycosylation. To characterize intact glycopeptides, mass spectrometry (MS)-based glycoproteomics have employed versatile fragmentation methods, among which Electron-Transfer/Higher-Energy Collision Dissociation (EThcD) has gained great popularity. However, the inherent limitation of EThcD in fragmenting low-charge ions has prevented its widespread applications. Furthermore, there is a need to develop a high-throughput strategy for comparative glycoproteomics with a large cohort of samples. Herein, we developed isobaric *N*,*N*-dimethyl Leucine derivatized ethylenediamine (DiLeuEN) tags to increase the charge states of glycopeptides, thereby improving the fragmentation efficiency and allowing for in-depth intact glycopeptide analysis, especially sialoglycopeptides. Moreover, the unique reporter ions of DiLeuEN-labeled glycopeptides generated in tandem MS spectra enable relative quantification of up to four samples in a single analysis, which represents a new high-throughput method for quantitative glycoproteomics.

Introduction

Estimated to occur in over half of the proteins in eukaryotes, glycosylation is one of the most common and important post-translational modifications (PTMs), having major roles in various biological processes[1-3]. The aberrant glycosylation patterns are oftentimes associated with human diseases such as cancer and immune disorders[4-6]. It is therefore crucial to characterize

glycoproteins and measure their relative abundances in complex biological specimen to reveal their roles during disease progression. Mass spectrometry has been successfully employed in characterization and quantification of glycoproteins recently[7, 8]. However, there is still limited method available to simultaneously enable accurate mapping of glycosylation site and high-throughput quantitative analysis in glycoproteomics[9].

To comprehensively characterize glycopeptides, the ideal way is to preserve the glycan side chain on the peptide and acquire glycan and peptide fragmentation information simultaneously[10]. Traditional collisional activated fragmentation approaches (e.g., collision-induced dissociation (CID) and higher-energy collision dissociation (HCD)) preferentially cleave glycosidic bond prior to peptide backbone, resulting in the dominated saccharide peaks in the MS/MS spectrum[11-13]. With the availability of multiple collisional energies, stepped HCD is an emerging dissociation method for intact N-glycopeptide characterization, where peptide backbones and N-glycan moieties are selectively fragmented at high and low collisional energies, respectively, enabling more glycopeptide identifications with higher scores in comparison with the regular HCD methods[14, 15]. Electronic excitation method (e.g., electron transfer dissociation (ETD)) predominantly breaks the N-C_α bond along the peptide backbone to form *c*/*z* ions and preserves labile glycosidic bond, allowing for site-specific analysis and sequence elucidation[16, 17]. However, ETD is not efficient to fragment low-charge precursor ions, making these ions undergo a process so-called nondissociative electron transfer dissociation (ETnoD)[18, 19], in which the radical product ions are compact and yield no sequence information. This is particularly prominent in analysis of glycopeptides where the hydrophilic glycans are usually difficult to be protonated and sialoglycans even carry negative charges. Chemical derivatization can improve fragmentation efficiency, which addresses the sequence problem to some extent[20-23]. In addition, Heck and colleagues combined

the unique features of ETD and HCD to develop a hybrid fragmentation method termed EThcD[24]. In EThcD, more informative spectra can be acquired as both amide bond and glycosidic bond are cleaved. Over thousands of glycopeptides were identified using EThcD lately, though concurrent quantification remains a technological gap[25].

Isotopic labeling method is commonly employed in quantitative proteomics[26-28]. Isotope labels can be introduced into peptides metabolically or chemically[28]. The stable isotope-labeled peptide is chemically identical to its native counterpart and therefore the two peptides behave identically during chromatographic and/or mass spectrometric analysis. Given that a mass spectrometer can recognize the mass difference between the light and heavy forms of a labeled peptide, quantification is achieved by comparing their respective signal intensities[29]. However, isotopic labeling usually imparts several Da mass differences into analytes and increases spectral complexity and thereby limiting its throughput. Unlike isotopic labeling, isobaric labeling provides multiplexing capabilities to further expand throughput[30-33]. Isobaric tags incorporate same nominal mass onto the analytes from different samples which co-elute together, without increasing spectral complexity. Concurrent peptide identification and relative quantification are achievable based on peptide sequence-specific fragments and unique reporter ions in MS/MS[34]. To date, there is scarce report of using isobaric labeling to investigate glycopeptides comprehensively and quantitatively[35].

Herein, we developed a set of isobaric tags termed *N*,*N*-dimethyl Leucine derivatized ethylenediamine (DiLeuEN) to chemically modify glycopeptides. Through chemical derivatization, the charge states of labeled glycopeptides are increased, allowing for improved EThcD fragmentation efficiency, which can greatly facilitate accurate characterization of intact glycopeptides. Moreover, the *N*,*N*-dimethyl Leucine (DiLeu) reporter ions generated in tandem MS spectra can be utilized for relative quantification. We demonstrate the feasibility of DiLeuEN

in both glycoprotein standards and complex glycoproteomes and we anticipate that the DiLeuEN tags can serve as a useful tool for large-scale quantitative glycoproteomics and further promote the understanding of glycosylation and its role in disease diagnosis and progression.

Materials and Methods

Materials and Reagents.

Acetic acid (AA), acetonitrile (ACN), *N*-(3-Dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC), 1-Hydroxybenzotriazole hydrate (HOBt), (7-Azabenzotriazol-1-yloxy)-tripyrrolidinophosphonium hexafluorophosphate (PyAOP), sodium cyanoborohydride , L-Leucine, 1-^{13}C-Leucine, 3-^{13}C-Leucine, 1-^{13}C, ^{15}N-Leucine, N-Methylmorpholine (NMM), 2,5-dihydroxybenzoic acid (DHB), dithiotreitol (DTT), iodoacetamide (IAA), *N*, *N*-dimethylaniline (DMA), formaldehyde (CH_2O, 37% w/w), 13C-formaldehyde ($^{13}CH_2O$, 20% w/w), ^{18}O water, borane-pyridine, *N*-Boc-ethylenediamine, tris(hydroxymethyl)aminomethane (Tris-HCl), 0.5 N HCl in dioxane, dimethyl sulfoxide (DMSO), formic acid (FA), methanol (MeOH), dichloromethane (DCM), and water (HPLC grade) were purchased from Fisher Scientific (Pittsburgh, PA). Sialylglycopeptide (SGP), bovine fetuin were purchased from Sigma-Aldrich (St. Louis, MO). Pancreatic cancer cells (PANC1) were cultured in DMEM-F12 (ATCC) supplemented with 10% fetal bovine serum (Hyclone) and 1% antibiotic-antimycotic solution (Cellgro). MS-grade Trypsin was purchased from Promega (Madison, WI). Sep-Pak C18 and Oasis HLB Cartridge was purchased from Waters Corporation (Milford, MA). Pierce™ C18 Tips was purchased from Thermo Fisher Scientific (Waltham, MA). Ethylene Bridged Hybrid C18 and ZIC-HILIC packing material (1.7 μm) was purchased from PolyLC Inc. (Columbia, MD). Fused

silica capillary tubing (inner diameter 75 μm, outer diameter 375 μm) was purchased from Polymicro Technologies (Phoenix, AZ). All reagents were used without additional purification.

Synthesis of DiLeuEN tags.

The DiLeuEN tags were synthesized at overall yield of 76% in three steps. For the first step, L-Leucine and sodium cyanoborohydride were suspended in MeOH and the mixture was cooled in an ice-water bath. Then, formaldehyde (CH_2O, 37% w/w) was added dropwise, and the mixture was stirred in an ice-water bath for 1hr. The product was purified by flash column chromatography (MeOH/DCM) and dried in vacuo to produce white solid DiLeu. The DiLeu reagent was then dissolved in DCM with 1.2 molar excess EDC and HOBt at room temperature for 30 min and then equal molar of *N*-Boc-ethylenediamine was added. The reaction was stirred for 18 hr and the product was purified through column chromatography as colorless oil. Boc group was deprotected under acidic condition in 0.5 N HCl in dioxane. The non-isotopic version of DiLeuEN was acquired as light-yellow color oil through short column chromatography after solvent neutralization and evaporation. The isobaric version of 4-plex DiLeuEN was synthesized accordingly using different starting isotopic reagents. For the channels require ^{18}O exchange, isotopic leucine was dissolved in 1 N HCl $H_2{}^{18}O$ solution (pH 1) and stirred on a hot plate at 65 °C for 4 h. Following evaporation of HCl from the solution in vacuo, trace amounts of acid were removed with StratoSpheres PL-HCO_3 MP resin (Agilent Technologies) to obtain ^{18}O L-leucine in free base form. The structure was confirmed by NMR and/or MS (**Figure S1-2**).

Chemical labeling with standard glycopeptides and glycoprotein using DiLeuEN.

The amidation condition was employed for chemical derivatization of carboxylic acid with minor modification[36]. SGP standard was dissolved in deionized water to make a stock solution at

1 mg/ml. Bovine fetuin was treated with trypsin first following common proteomics procedures to make tryptic digest and aliquot to 1 mg/ml[37]. 100 µg of peptide samples were taken, 5 µl of formaldehyde and 10 µl of 1% borane-pyridine were added and the mixture was incubated for 1 hr at room temperature. 1 M DiLeuEN and 250 mM PyAOP were prepared in DMSO. Then the dimethylated peptides were vacuum dried and dissolved in DMSO/NMM (9:1), and DiLeuEN and PyAOP were added to make the final concentration of 0.5 M and 50 mM respectively. The mixture was vortexed and stood at room temperature for 2 hr before desalting using Pierce™ C18 tips or C18 cartridges, and the eluted fractions were dried in vacuo, reconstituted in 0.1% FA, and analyzed by MALDI-MS or LC-MS/MS immediately.

Enrichment, derivatization, and quantitative analysis of complex biological samples.

Panc 1 cell pellet was lysed with 8 M urea in 50 mM Tris-HCl supplemented with Phosphatase Inhibitor Cocktail (Thermo Scientific), which was followed by sonication for three cycles (15 s on, 15 s off) at 4 °C. After centrifugation at 14,000 × g for 15 min, the supernatants were collected, and the protein concentration were determined using Pierce BCA Protein Assay Kit (Thermo Scientific). All the proteins were reduced and alkylated with 10 mM DTT and 20 mM IAA at 37 °C for 60 min in the dark. Diluted the samples 8-fold with 50 mM Tris-HCl and digested using trypsin at an enzyme to protein ratio of 1:50 at 37 °C overnight. After digestion, the resulted peptides were desalted by C18 cartridges and dried in vacuo.

Peptides were reconstituted in loading buffer consisting of 80% ACN and 1% TFA to make a final concentration at around 1 mg/ml. Meanwhile, weighted out the ZIC-HILIC beads, and washed three times using the loading buffer. Loaded beads onto a pipette tip prepacked with cotton wool according to the peptide to beads ratio of 1:25. Samples were loaded three times followed by

three times wash with loading buffer. The retained glycopeptides were eluted with 0.1% TFA twice. The collected elutes were dried in a SpeedVac concentrator.

Enriched glycopeptides were aliquot into two parts in which one preserved in lyophilized states at -80 °C and the other incubated in 0.1% FA at 4 °C for two weeks. Then, same amount of two samples from each part were taken and labeled with 4-plex DiLeuEN respectively through two-step reaction scheme mentioned above. Labeled peptides were mixed together and desalted using Oasis HLB cartridges before performing LC-MS/MS analysis.

Matrix-Assisted Laser Desorption/Ionization (MALDI)-MS analysis.

Samples were prepared by premixing 1 µL of DiLeuEN-labeled glycopepties with 1 µL DHB matrix (100 mg/mL in 5% DMA, 47.5% MeOH and 47.5% water), and 1 µL of each matrix/sample mixture was spotted onto the MALDI target plate. A MALDI-LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, Bremen, Germany) was used for MALDI-MS analysis. Ionization was performed using a laser energy of 18 µJ. Spectra were acquired in the Orbitrap mass analyzer within a mass range of *m/z* 1,000-4,000 at a mass resolution of 60 K (at *m/z* 400).

LC-MS/MS analysis.

A self-fabricated nano-C18 column (15 cm, 75 µm i.d., 1.7 µm Ethylene Bridged Hybrid C18 packing material) was used for sample separation. A Dionex Ultimate 3000 nanoLC system was coupled to an Orbitrap Fusion Lumos Mass Spectrometer (Thermo Scientific, Bremen, Germany) for all LC-MS/MS analyses. Mobile phase A was deionized water containing 0.1% FA, and mobile phase B was ACN containing 0.1% FA. The flow rate was set at 0.3 µL/min, and the injection volume was 2 µL. The following gradient was used

(time, % mobile phase B) unless otherwise specified: (0 min, 3%), (18 min, 3%), (120 min, 30%), (130 min, 75%), (131 min, 95%), (140 min, 95%), (141 min, 3%), (155 min, 3%).

The following mass spectrometer parameters were used for data acquisition. Samples were ionized in positive ion mode with a spray voltage of 3 kV. S-lens radio frequency (RF) level was set to be 30, and capillary temperature was set to be 300 °C. Full MS scans were acquired at m/z 300-2000 with resolving power of 120 K. Maximum injection time of 150 ms, automatic gain control (AGC) target value of 2e5, and 1 microscan were used for full MS scans. Top 15 precursors with highest charge states were selected for MS/MS in an order of intensity. EThcD was performed with calibrated charge-dependent reaction time supplemented by 25% HCD activation. The isolation window was set as 2.5 *m/z* and the dynamic exclusion of acquired precursors was set to 15 sec with a ± 20 ppm tolerance. The first fixed mass is 100 m/z in order to obtain reporter ions.

Data analysis.

XCalibur Qual Browser was used for data processing. Proteome Discoverer and Byonic were used for data analysis. The mass tolerance for precursor ions set to 10 ppm and 20 ppm for fragment ions respectively. Fragmentation method set as EThcD. The enzyme was full trypsin and maximal missed cleavage was 2. Cysteine carbamidomethylation (+57.021 Da), lysine and N-Terminal dimethylation (+28.031 Da), Aspartic acid, Glutamic acid and C-Terminal DiLeuEN-amidation (+183.174 Da for non-isotopic version while +187.181 Da for 4-plex DiLeuEN) were set as fixed modification. Variable modifications contained oxidation on Met (M +15.995 Da, rare 2) and N-glycosylation (common 1). The protein database was SwissProt-bovine or human. In addition, the sialylated glycans were adjusted with corresponding additional mass (+183.174 or +187.181 per sialic glycan) in glycan database. The N-glycopeptides identified were filtered to 1%

FDR and Byonic score ≥ 150. Quantification was performed with a reporter ion integration tolerance of 20 ppm for the most confident centroid. Reporter ion ratio values for proteins groups were exported and analyzed in Excel workbook.

Results and Discussion

Tandem mass spectrometry has been widely adopted to elucidate the sequences of biomolecules such as glycopeptides with great success[38, 39]. Traditional approaches opt to analyze enzymatically released glycans and their attached peptides separately, due to heterogeneous nature of the two parts, leading to loss of site-specific information and a gap between glycomics and proteomics[40-42]. Collisional-based fragmentation and electron-based fragmentation method produce complementary results. Though stepped collision energy method produces much evenly distributed glycan and peptide fragments in single MS/MS spectra, glycan fragments resulting from glycosidic bond cleavage dominate MS/MS spectra generated by traditional CID or HCD with little knowledge of glycosylation sites and amino acid sequences. Whereas ETD promotes cleavages of N-C_α bonds that results in *c/z* type ions with retention of labile PTMs on the amino acid side chains, providing a useful way to characterize intact glycopeptides[11-13]. However, due to the free-radical-driven mechanism, ETD fragmentation suffers from poor efficiency on low charge precursor ions. In order to alleviate this handicap, two major strategies are commonly employed[19]. A straightforward approach is to increase charge states of targeted precursor ions. This could be achieved by altering electrospray conditions or introducing covalent attachment with either a fixed charge or basic functionalities that are able to enhance the protonation[19, 43, 44]. An alternative method is supplemental activation, in which extra energy is introduced into the fragmentation process and noncovalent interactions contributing to ETnoD are thus disrupted[24, 45, 46]. Several

activation approaches have been explored, with the most widely adopted being EThcD, which involves activation of all ETD products with high-energy collisional dissociation.

In light of both approaches, we designed a set of novel isobaric tags named DiLeuEN for high-throughput relative quantification of intact glycopeptides. Previously, derivatization of the C-termini of peptides has been reported to improve ETD efficiency for peptide fragmentation. Several reagents, such as (2-Aminoethyl)-trimethylammonium (AETMA), *N*,*N*-dimethylethylenediamine (DMEDA), and 3-(dimethylamino)-1-propylamine (C3-methylTert) were further developed[19, 23, 47]. Although the detailed structures of these derivatization reagents are slightly different, they all share the same core structure and can be regarded as the derivatives of *ethylenediamine*. Through chemical derivatization with C-terminus of a peptide, this charge-reversal strategy replaces the most acidic groups as well as negative-charge-carrying carboxylate acid with more basic groups and commonly positive-charge-carrying amide, thus increasing the charge states of the peptide of interest (**Figure S3**). However, these reagents inherently lack multiplexing capacity, preventing their use for high-throughput quantitative analysis. Our group developed DiLeu isobaric tags for multiplex protein and peptide relative quantification a decade ago. DiLeu features comparable performance and reduced cost compared with commercially available reagent iTRAQ or TMT[48]. The synthesis of DiLeu is straightforward, which offers an opportunity to develop an array of natural amino acid-based tandem mass tagging reagents. We exploited DiLeu multiplexing capacity from 4-plex to 12 plex, and more recently expanded to 21-plex, demonstrating its application to high-throughput quantitative analysis[49-51]. Isobaric multiplex reagents for carbonyl-containing compound (SUGAR) tags were developed subsequently, aiming to label released N-glycans, while DiLeuPMP tags were designed for quantitative O-glycomics.[52-54]. In the current study, by incorporating DiLeu onto ethylenediamine, a set of 4-plex DiLeuEN

isobaric tags developed to bridge the gap between glycomics and glycoproteomics. These tags are designed to provide benefits of both ethylenediamine as a good derivatization functionality for ETD fragmentation enhancement and DiLeu as an isobaric label for MS^2 relative quantification. DiLeuEN takes three steps to synthesize with high yield in regular analytical lab settings, and 4 out of 12-plex DiLeu isotope configuration have been chosen for 4-plex DiLeuEN architectures (**Figure 1**). Isotopic version of DiLeuEN can be produced following the same route accordingly using commercially available isotope reagents. Deuterium atom is not utilized in isotopic form to avoid possible chromatography retention time shift[55]. Thus, for channels 115 and 116, an extra ^{18}O exchange step is required after DiLeu is made.

The feasibility of the DiLeuEN isobaric tags for glycosylation analysis was first evaluated using standard glycopeptides and glycoproteins. Several amide coupling reactions were investigated for derivatization of carboxylic acids in peptides previously[19, 23, 47]. Although the popular water soluble coupling reagent EDC are widely used, the labeling efficiency has been reported to be undesirable, especially for aspartic acid and glutamic acid side chain carboxyls[23]. The completeness of the labeling reaction is increasingly important with cumulatively sample complexity. Incomplete labeling dramatically increases spectral complexity and requires dynamic modifications to consider each moiety, making the database searching speed for glycopeptide identification considerably slower compared to the situations with complete derivatization, where only static modification needs to be considered. The unmodified or partially modified peptides exist as separate chemical species with different retention times and masses than the desired fully modified peptides, leading to inaccurate quantification.

We evaluated the labeling efficiency of DiLeuEN tag using several standard peptides and adopted a two-step reaction scheme. The native amine groups were protected from interfering the

sequential amidation reaction by dimethylation firstly. Then the amidation reaction between peptides and DiLeuEN was promoted by strong coupling reagents, PyAOP. This two-step reaction exhibits complete reaction yield. Sialylglycopeptide (SGP) was used as a standard as it has two sialic acids at the terminal of glycan moiety which carry carboxylic acid groups. After dimethylation, the N-terminus and free amine side chain of SGP were completely dimethylated. Both C-terminus of the peptide and carboxylic acid group on sialic acid were completely modified after amidation (**Figure 2 A-C**). The labile sialic acids are commonly missing due to in-source fragmentation. Noticeably, the stability of sialic acids was greatly enhanced after amidation, which is in accordance with a similar study previously[56]. A peptide mixture containing three peptides which have aspartic acid (D) and glutamic acid (E) sidechain carboxyls within their sequences were also tested. The results demonstrated complete labeling efficiency similar to SGP (**Figure S4**). Moreover, the peptide charge states were observed to be increased substantially upon derivatization in positive-mode electrospray ionization (ESI). As shown in **Figure 2 D-F**, only charge states of +2, +3, and +4 were present in the ESI-MS spectra of unmodified SGP, among which the +2 and +3 precursor ions were the dominant species. Identical charge distribution was observed after dimethylation, but with lower fraction of +2 precursor ion. Notably, by installing an easily protonated tertiary amine moiety, DiLeuEN labeling dramatically elevated the charge states of SGP precursor ions to +4, +5 and +6, among which the +5 charge state is the most abundant species. The average charge state of precursor ions was boosted from +2.7 to +4.8. It is also worth noting that the simplification of the spectrum may be another potential benefit of DiLeuEN derivatization. Before amidation, lots of Na^+/K^+ adducts were observed for each species whereas only protonated precursor ions were preserved after derivatization. This is due to the block of carboxylic acid group on sialic acids, which usually carries metal ions[57].

Bovine fetuin was used next as a standard glycoprotein to further validate the feasibility of our labeling method. Fetuin was digested by trypsin before performing the two-step derivatization reaction. We successfully characterized a total of 361 glycoforms in fetuin, covering all three reported N-glycosylation sites (**Table S1**). As shown in **Figure 3 and S5**, only a few fragments were observed for the native sialoglycopeptides in fetuin, whereas extensive fragmentation along the peptide backbone with consecutive *c* and *z* ions for DiLeuEN-labeled fetuin was observed. As expected, the charge states of peptides were increased after DiLeuEN derivatization (**Figure S6**). This observation confirms that an increase in the charges of intact glycopeptides can greatly improve their ETD fragmentation efficiency as well as sequence coverage. Furthermore, the chemical derivatization did not significantly affect the chromatography behaviors of the glycopeptides, as indicated by the similar pattern of extracted ion chromatography (XIC) of HexNAc (**Figure S7**). To evaluate quantitative performance, we prepared 4-plex DiLeuEN labeled digested fetuin individually and mixed them at 1:2:5:10 ratio in triplicates and analyzed at a resolving power of 60k (at *m/z* 200). **Figure 4A** shows a close correlation between experimental and theoretical values. Across all channels, the median ratios were measured within 5% of the expected values with average coefficients of variation (CV) of 7%, 13% and 19%, respectively.

Moreover, given that this strategy is simple and easily adopted, we further applied this method to study glycoproteomes in pancreatic cancer cells (Panc1) to verify its applicability to complex biological samples. The proteins from Panc1 cells were extracted and digested following common proteomics procedure prior to HILIC enrichment[37]. The enriched glycopeptides were divided and processed with two-step labeling procedure with or without DiLeuEN. For the samples without DiLeuEN labeling, same amount of DMSO were used to minimize the differences of sample loss potentially caused by chemical labeling and sample cleanup. A total of 997 intact N-

glycopeptides were identified, representing 320 glycosylation sites from 220 glycoproteins as shown in **Figure 4B**. In comparison, only 738 glycopeptides were identified without labeling. Because of DiLeuEN derivatization, there were about 35% increase in the number of intact N-glycopeptides, 53% increase in glycosylation sites, and 20% increase in glycoproteins.

The glycopeptide samples are normally dissolved and stored in 0.1% formic acid (FA) solution prior to LC-MS/MS analyses, where 0.1% FA provide an acidic condition (pH=2.7) to facilitate protonation during ESI. Although sialic acids are noted to be labile under acidic environment, there are surprisingly few reports quantitatively assessing the changes of sialic acids on glycopeptides in acidic condition[58]. The DiLeuEN tags provide an excellent tool for quantitative investigation of sialoglycopeptides. We incubated equal amounts of Panc1 tryptic digests either in 0.1% FA or keeping lyophilized for two weeks. Duplicates from both digests were used to constitute 4 samples in total. The lyophilized samples were labeled with DiLeuEN channels 115 and 116, whereas samples incubated in 0.1% FA were labeled with channels 117 and 118. Then they were combined at 1:1:1:1 ratio (based on peptide assay, Pierce) prior to sample clean-up and LC-MS/MS analysis. For quantified peptides, the ratio of average signal intensity between 117+118 and 115+116 channels represents relative concentration changes of certain peptides over time. The results demonstrated that among overall 534 quantified glycopeptides, most of which remained stable and did not change statistically. However, the signal of 37 of sialoglycopeptides exhibited significant decrease after incubation in acidic condition (**Figure 4C and Table S2**). The incubation of SGP under the same condition resulted in the signal decrease of native precursors and emerging new peaks corresponding to one or two sialic acids loss (**Figure S8**), which suggested the signal decrease was due to the decay of sialic acids on sialoglycopeptides under acidic environment over time. Nevertheless, it is noted that among all quantified 139

sialoglycopeptides, most of which remained stable, at least statistically. We reasoned that this may be due to the differential stability of sialic acid linkage isomers in acidic conditions[58, 59]. Unfortunately, performed under strong coupling condition, DiLeuEN does not discriminate linkage isomers and thus cannot provide detailed structural information. Overall, these results suggest that for glycoproteomic analysis, the prepared samples shall be analyzed as early as possible before significant sample degradation occurs.

Conclusions

In the present study, we have successfully developed a novel class of isobaric tandem mass tagging reagent, DiLeuEN. With optimized two-step labeling procedure, complete labeling was achieved and the challenge in ETD fragmentation of low-charge precursor ions was addressed. Through chemical derivatization, the charges and ionization efficiency as well as fragmentation performance of intact glycopeptides were significantly improved. Isobaric quantification features of DiLeu tags were maintained in 4-plex DiLeuEN which offers precise and high-throughput relative quantitative analysis with glycosylation mapping in the meantime. Benefiting from EThcD, site-specific intact glycopeptide structure characterization and quantification can be achieved in a single spectrum. Furthermore, we applied this strategy to the analysis of complex biological samples to demonstrate its versatility and found potential degradation of sialoglycopeptides under acidic environment. Overall, DiLeuEN proves to be a useful tool for processing multiple samples at one time in glycoproteomics analysis. Future work will be conducted to expand the throughput even further and extend this methodology to investigate disease related biomarker discoveries.

Acknowledgements

This work was funded in part by the National Institutes of Health (NIH) grants RF1 AG052324, U01CA231081, R01 DK071801, and P41GM108538. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531) and Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. LL acknowledges NIH grant support R21AG065728 as well as 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. Ruhaak, L. R.; Xu, G. G.; Li, Q. Y.; Goonatilleke, E.; Lebrilla, C. B., Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chem. Rev.* **2018,** *118* (17), 7886-7930.

2. Ohtsubo, K.; Marth, J. D., Glycosylation in cellular mechanisms of health and disease. *Cell* **2006,** *126* (5), 855-867.

3. Ungar, D., Golgi linked protein glycosylation and associated diseases. *Semin. Cell Dev. Biol.* **2009,** *20* (7), 762-769.

4. Gudelj, I.; Lauc, G., Protein N-Glycosylation in Cardiovascular Diseases and Related Risk Factors. *Curr. Cardiovasc. Risk Rep.* **2018,** *12* (6).

5. Kristic, J.; Lauc, G., Ubiquitous Importance of Protein Glycosylation. In *High-Throughput Glycomics and Glycoproteomics: Methods and Protocols*, Lauc, G.; Wuhrer, M., Eds. 2017; Vol. 1503, pp 1-12.

6. Kanninen, K.; Goldsteins, G.; Auriola, S.; Alafuzoff, I.; Koistinaho, J., Glycosylation changes in Alzheimer's disease as revealed by a proteomic approach. *Neurosci. Lett.* **2004,** *367* (2), 235-240.

7. Kim, U.; Oh, M. J.; Lee, J.; Hwang, H. Y.; An, H. J., MS-Based Technologies for the Study of Site-Specific Glycosylation. *Mass Spectrom. Lett.* **2017,** *8* (4).

8. Illiano, A.; Pinto, G.; Melchiorre, C.; Carpentieri, A.; Faraco, V.; Amoresano, A., Protein Glycosylation Investigated by Mass Spectrometry: An Overview. *Cells* **2020,** *9* (9).

9. Fang, P.; Ji, Y. L.; Silbern, I.; Doebele, C.; Ninov, M.; Lenz, C.; Oellerich, T.; Pan, K. T.; Urlaub, H., A streamlined pipeline for multiplexed quantitative site-specific N-glycoproteomics. *Nat. Commun.* **2020,** *11* (1).

10. Zhu, Z. K.; Desaire, H., Carbohydrates on Proteins: Site-Specific Glycosylation Analysis by Mass Spectrometry. *Annu. Rev. Anal. Chem., Vol 8* **2015,** *8*, 463-483.

11. Hu, H.; Khatri, K.; Klein, J.; Leymarie, N.; Zaia, J., A review of methods for interpretation of glycopeptide tandem mass spectral data. *Glycoconjugate J.* **2016,** *33* (3), 285-296.

12. Lakbub, J. C.; Su, X. M.; Zhu, Z. K.; Patabandige, M. W.; Hua, D.; Go, E. P.; Desaire, H., Two New Tools for Glycopeptide Analysis Researchers: A Glycopeptide Decoy Generator and a Large Data Set of Assigned CID Spectra of Glycopeptides. *J. Proteome Res.* **2017,** *16* (8), 3002-3008.

13. Chandler, K. B.; Pompach, P.; Goldman, R.; Edwards, N., Exploring Site-Specific N-Glycosylation Microheterogeneity of Haptoglobin Using Glycopeptide CID Tandem Mass Spectra and Glycan Database Search. *J. Proteome Res.* **2013,** *12* (8), 3652-3666.

14. Yang, H.; Yang, C. X.; Sun, T. L., Characterization of glycopeptides using a stepped higher-energy C-trap dissociation approach on a hybrid quadrupole orbitrap. *Rapid Commun. Mass Spectrom.* **2018,** *32* (16), 1353-1362.

15. Wang, Y.; Tian, Z. X., New Energy Setup Strategy for Intact N-Glycopeptides Characterization Using Higher-Energy Collisional Dissociation. *J. Am. Soc. Mass Spectrom.* **2020,** *31* (3), 651-657.

16. Riley, N. M.; Coon, J. J., The Role of Electron Transfer Dissociation in Modern Proteomics. *Anal. Chem.* **2018,** *90* (1), 40-64.

17. Madsen, J. A.; Brodbelt, J. S., Simplifying Fragmentation Patterns of Multiply Charged Peptides by N-Terminal Derivatization and Electron Transfer Collision Activated Dissociation. *Anal. Chem.* **2009,** *81* (9), 3645-3653.

18. Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A., Electron-transfer ion/ion reactions of doubly protonated peptides: Effect of elevated bath gas temperature. *Anal. Chem.* **2005,** *77* (17), 5662-5669.

19. Ko, B. J.; Brodbelt, J. S., Enhanced Electron Transfer Dissociation of Peptides Modified at C-terminus with Fixed Charges. *J. Am. Soc. Mass Spectrom.* **2012,** *23* (11), 1991-2000.

20. Frey, B. L.; Krusemark, C. J.; Ledvina, A. R.; Coon, J. J.; Belshaw, P. J.; Smith, L. M., Ion-ion reactions with fixed-charge modified proteins to produce ions in a single, very high charge state. *Int. J. Mass Spectrom.* **2008,** *276* (2-3), 136-143.

21. Krusemark, C. J.; Frey, B. L.; Belshaw, P. J.; Smith, L. M., Modifying the Charge State Distribution of Proteins in Electrospray Ionization Mass Spectrometry by Chemical Derivatization. *J. Am. Soc. Mass Spectrom.* **2009,** *20* (9), 1617-1625.

22. Li, X. J.; Cournoyer, J. J.; Lin, C.; O'Connor, P. B., The Effect of Fixed Charge Modifications on Electron Capture Dissociation. *J. Am. Soc. Mass Spectrom.* **2008,** *19* (10), 1514-1526.

23. Frey, B. L.; Ladror, D. T.; Sondalle, S. B.; Krusemark, C. J.; Jue, A. L.; Coon, J. J.; Smith, L. M., Chemical Derivatization of Peptide Carboxyl Groups for Highly Efficient Electron Transfer Dissociation. *J. Am. Soc. Mass Spectrom.* **2013,** *24* (11), 1710-1721.

24. Frese, C. K.; Altelaar, A. F. M.; van den Toorn, H.; Nolting, D.; Griep-Raming, J.; Heck, A. J. R.; Mohammed, S., Toward Full Peptide Sequence Coverage by Dual Fragmentation Combining Electron-Transfer and Higher-Energy Collision Dissociation Tandem Mass Spectrometry. *Anal. Chem.* **2012,** *84* (22), 9668-9673.

25. 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* (9), 1751-1764.

26. Morton, T. H., Isotopic Labelling in Mass Spectrometry. *Encyclopedia of Spectroscopy and Spectrometry, 3rd Edition, Vol 2: G-M* **2017**, 501-511.

27. Lill, J., Proteomic tools for quantitation by mass spectrometry. *Mass Spectrom. Rev.* **2003,** *22* (3), 182-194.

28. Ashman, K.; Canas, M. I. R.; Luque-Garcia, J. L.; Martinez, F. G., *Stable Isotopic Labeling for Proteomics*. 2011; p 549-573.

29. Iliuk, A.; Galan, J.; Tao, W. A., Playing tag with quantitative proteomics. *Anal. Bioanal. Chem.* **2009,** *393* (2), 503-513.

30. Cheng, L.; Pisitkun, T.; Knepper, M. A.; Hoffert, J. D., Peptide Labeling Using Isobaric Tagging Reagents for Quantitative Phosphoproteomics. *Phospho-Proteomics: Methods and Protocols, 2nd Edition* **2016,** *1355*, 53-70.

31. High, A. A.; Tan, H. Y.; Pagala, V. R.; Niu, M. M.; Cho, J. H.; Wang, X. S.; Bai, B.; Peng, J. M., Deep Proteome Profiling by Isobaric Labeling, Extensive Liquid Chromatography, Mass Spectrometry, and Software-assisted Quantification. *JOVE-J VIS EXP* **2017,** (129).

32. Erdjument-Bromage, H.; Huang, F. K.; Neubert, T. A., Sample Preparation for Relative Quantitation of Proteins Using Tandem Mass Tags (TMT) and Mass Spectrometry (MS). In *Glioblastoma: Methods and Protocols*, Placantonakis, D. G., Ed. 2018; Vol. 1741, pp 135-149.

33. Nunez, E. V.; Domont, G. B.; Nogueira, F. C. S., iTRAQ-Based Shotgun Proteomics Approach for Relative Protein Quantification. In *Multiplex Biomarker Techniques: Methods and Applications*, Guest, P. C., Ed. 2017; Vol. 1546, pp 267-274.

34. Rauniyar, N.; Yates, J. R., Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics. *J. Proteome Res.* **2014,** *13* (12), 5293-5309.

35. Zhu, H. B.; Qiu, C.; Gryniewicz-Ruzicka, C. M.; Keire, D. A.; Ye, H. P., Multiplexed Comparative Analysis of Intact Glycopeptides Using Electron-Transfer Dissociation and Synchronous Precursor Selection Based Triple-Stage Mass Spectrometry. *Anal. Chem.* **2020,** *92* (11), 7547-7555.

36. Yang, L. J.; Sun, Z. Y.; Zhang, L.; Cai, Y.; Peng, Y.; Cao, T.; Zhang, Y.; Lu, H. J., Chemical labeling for fine mapping of IgG N-glycosylation by ETD-MS. *Chem. Sci.* **2019,** *10* (40), 9302-9307.

37. Modern Proteomics - Sample Preparation, Analysis and Practical Applications. *Modern Proteomics - Sample Preparation, Analysis and Practical Applications* **2016,** *919*, 1-533.

38. Nilsson, J., Liquid chromatography-tandem mass spectrometry-based fragmentation analysis of glycopeptides. *Glycoconjugate J.* **2016,** *33* (3), 261-272.

39. Zhang, Y. L.; Yu, C. Y.; Song, E. W.; Li, S. C.; Mechref, Y.; Tang, H. X.; Liu, X. W., Identification of Glycopeptides with Multiple Hydroxylysine O-Glycosylation Sites by Tandem Mass Spectrometry. *J. Proteome Res.* **2015,** *14* (12), 5099-5108.

40. Thaysen-Andersen, M.; Packer, N. H., Advances in LC-MS/MS-based glycoproteomics: Getting closer to system-wide site-specific mapping of the N- and O-glycoproteome. *BBA-PROTEINS PROTEOM* **2014,** *1844* (9), 1437-1452.

41. Chen, W. X.; Smeekens, J. M.; Wu, R. H., A Universal Chemical Enrichment Method for Mapping the Yeast N-glycoproteome by Mass Spectrometry (MS). *Mol. Cell. Proteomics* **2014,** *13* (6), 1563-1572.

42. Tian, Y. A.; Zhou, Y.; Elliott, S.; Aebersold, R.; Zhang, H., Solid-phase extraction of N-linked glycopeptides. *Nat. Protoc.* **2007,** *2* (2), 334-339.

43. Bonvin, G.; Rudaz, S.; Schappler, J., In-spray supercharging of intact proteins by capillary electrophoresis-electrospray ionization-mass spectrometry using sheath liquid interface. *Anal. Chim. Acta* **2014,** *813*, 97-105.

44. Gong, X. Y.; Li, C.; Zhai, R.; Xie, J.; Jiang, Y.; Fang, X., Supercharging of Proteins by Salts during Polarity Reversed Nano-Electrospray Ionization. *Anal. Chem.* **2019,** *91* (3), 1826-1837.

45. McAlister, G. C.; Phanstiel, D.; Good, D. M.; Berggren, W. T.; Coon, J. J., Implementation of electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer. *Anal. Chem.* **2007,** *79* (10), 3525-3534.

46. Tsybin, Y. O.; He, H.; Emmett, M. R.; Hendrickson, C. L.; Marshall, A. G., Ion activation in electron capture dissociation to distinguish between N-terminal and C-terminal productions. *Anal. Chem.* **2007,** *79* (20), 7596-7602.

47. Somasundaram, P.; Koudelka, T.; Linke, D.; Tholey, A., C-Terminal Charge-Reversal Derivatization and Parallel Use of Multiple Proteases Facilitates Identification of Protein C-Termini by C-Terminomics. *J. Proteome Res.* **2016,** *15* (4), 1369-1378.

48. Xiang, F.; Ye, H.; Chen, R. B.; Fu, Q.; Li, L. J., N,N-Dimethyl Leucines as Novel Isobaric Tandem Mass Tags for Quantitative Proteomics and Peptidomics. *Anal. Chem.* **2010,** *82* (7), 2817-2825.

49. Greer, T.; Hao, L.; Nechyporenko, A.; Lee, S.; Vezina, C. M.; Ricke, W. A.; Marker, P. C.; Bjorling, D. E.; Bushman, W.; Li, L. J., Custom 4-Plex DiLeu Isobaric Labels Enable Relative Quantification of Urinary Proteins in Men with Lower Urinary Tract Symptoms (LUTS). *Plos One* **2015,** *10* (8).

50. Frost, D. C.; Li, L. J., High-Throughput Quantitative Proteomics Enabled by Mass Defect-Based 12-Plex DiLeu Isobaric Tags. *Quantitative Proteomics by Mass Spectrometry, 2nd Edition* **2016,** *1410*, 169-194.

51. Frost, D. C.; Feng, Y.; Li, L. J., 21-plex DiLeu Isobaric Tags for High-Throughput Quantitative Proteomics. *Anal. Chem.* **2020,** *92* (12), 8228-8234.

52. Feng, Y.; Chen, B. M.; Yu, Q. Y.; Zhong, X. F.; Frost, D. C.; Ikonomidou, C.; Li, L. J., Isobaric Multiplex Labeling Reagents for Carbonyl-Containing Compound (SUGAR) Tags: A Probe for Quantitative Glycomic Analysis. *Anal. Chem.* **2019,** *91* (4), 3141-3146.

53. Feng, Y.; Li, M. Y.; Lin, Y. Y.; Chen, B. M.; Li, L. J., Multiplex Quantitative Glycomics Enabled by Periodate Oxidation and Triplex Mass Defect Isobaric Multiplex Reagents for Carbonyl-Containing Compound Tags. *Anal. Chem.* **2019,** *91* (18), 11932-11937.

54. Li, M.; Gu, T. J.; Lin, X.; Li, L., DiLeuPMP: A Multiplexed Isobaric Labeling Method for Quantitative Analysis of. *Anal. Chem.* **2021**.

55. Yan, X. J.; Sun, L. L.; Dovichi, N. J.; Champion, M. M., Minimal deuterium isotope effects in quantitation of dimethyl-labeled complex proteomes analyzed with capillary zone electrophoresis/mass spectrometry. *Electrophoresis* **2020,** *41* (15), 1374-1378.

56. Nishikaze, T.; Kawabata, S.; Tanaka, K., In-Depth Structural Characterization of N-Linked Glycopeptides Using Complete Derivatization for Carboxyl Groups Followed by Positive- and Negative-Ion Tandem Mass Spectrometry. *Anal. Chem.* **2014,** *86* (11), 5360-5369.

57. Saladini, M.; Menabue, L.; Ferrari, E., Binding ability of sialic acid towards biological and toxic metal ions. NMR, potentiometric and spectroscopic study. *J. Inorg. Biochem.* **2002,** *88* (1), 61-68.

58. Sonnenburg, J. L.; van Halbeek, H.; Varki, A., Characterization of the acid stability of glycosidically linked neuraminic acid - Use in detecting de-N-acetyl-gangliosides in human melanoma. *J. Biol. Chem.* **2002,** *277* (20), 17502-17510.

59. de Haan, N.; Yang, S.; Cipollo, J.; Wuhrer, M., Glycomics studies using sialic acid derivatization and mass spectrometry. *Nat. Rev. Chem.* **2020,** *4* (5), 229-242.

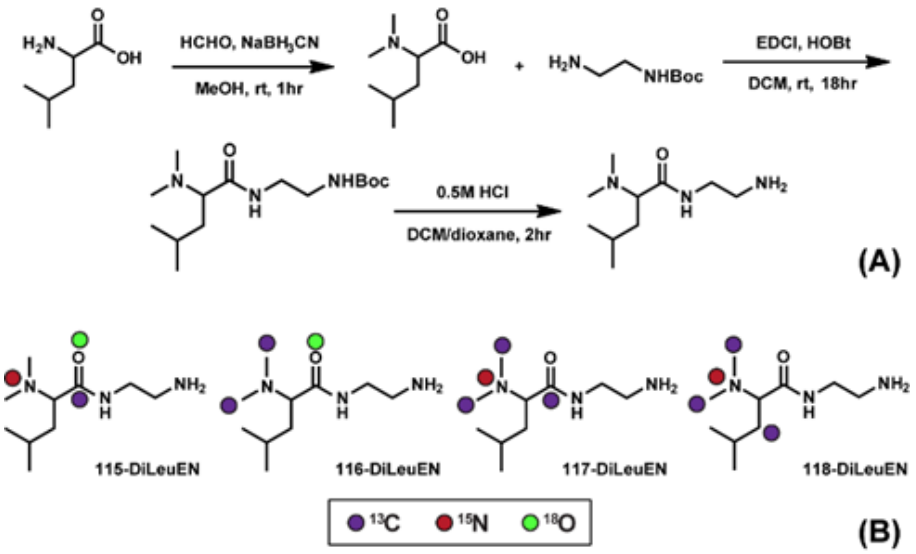


Figure 1. (A) Structure and synthetic routes of DiLeuEN tag; (B) Isotope configurations of 4-plex DiLeuEN tags (purple, red, and green dots represent ^{13}C, ^{15}N, and ^{18}O isotope position respectively).

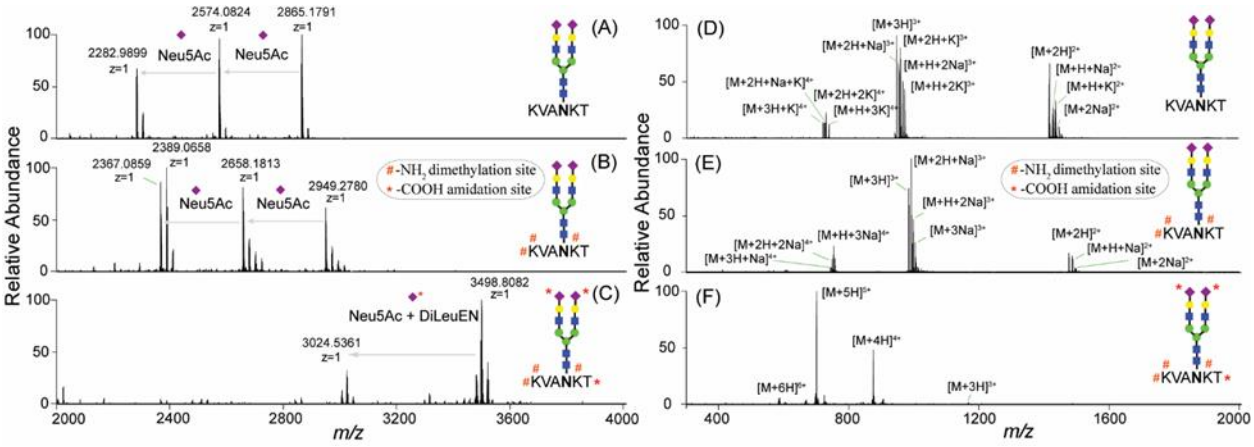


Figure 2. (A-C) MALDI-MS profiling of native, dimethylated and DiLeuEN labeled SGP; (D-F) ESI-MS profiling of native, dimethylated and DiLeuEN labeled SGP.

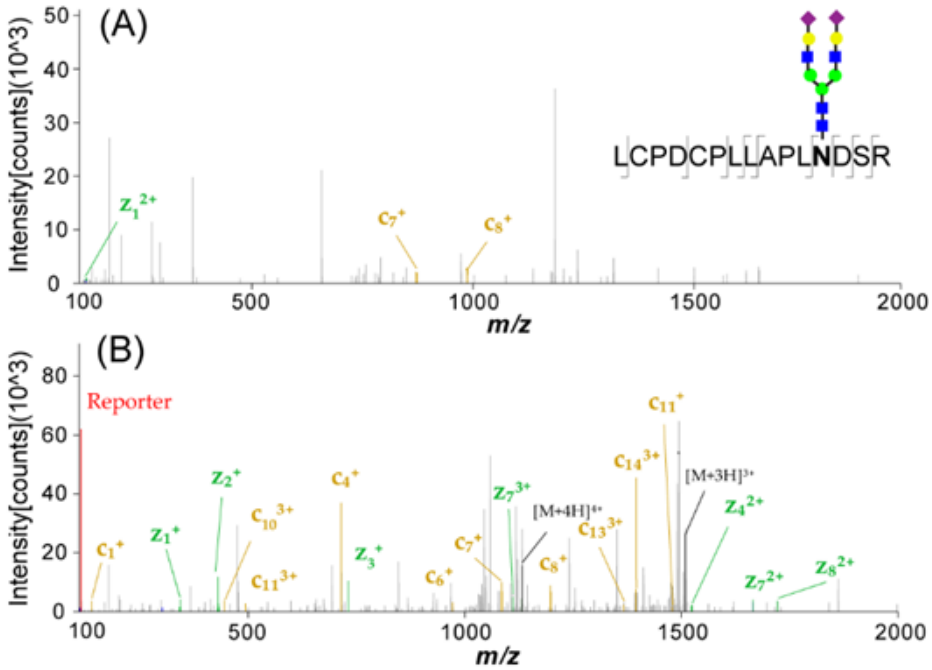


Figure 3. Representative EThcD-MS/MS spectra for the sialoglycopeptides from bovine fetuin. (A) underivatized, (B) DiLeuEN derivatized. (To emphasize ETD fragmentation improvement, b/y ions and glycan fragments are not annotated)

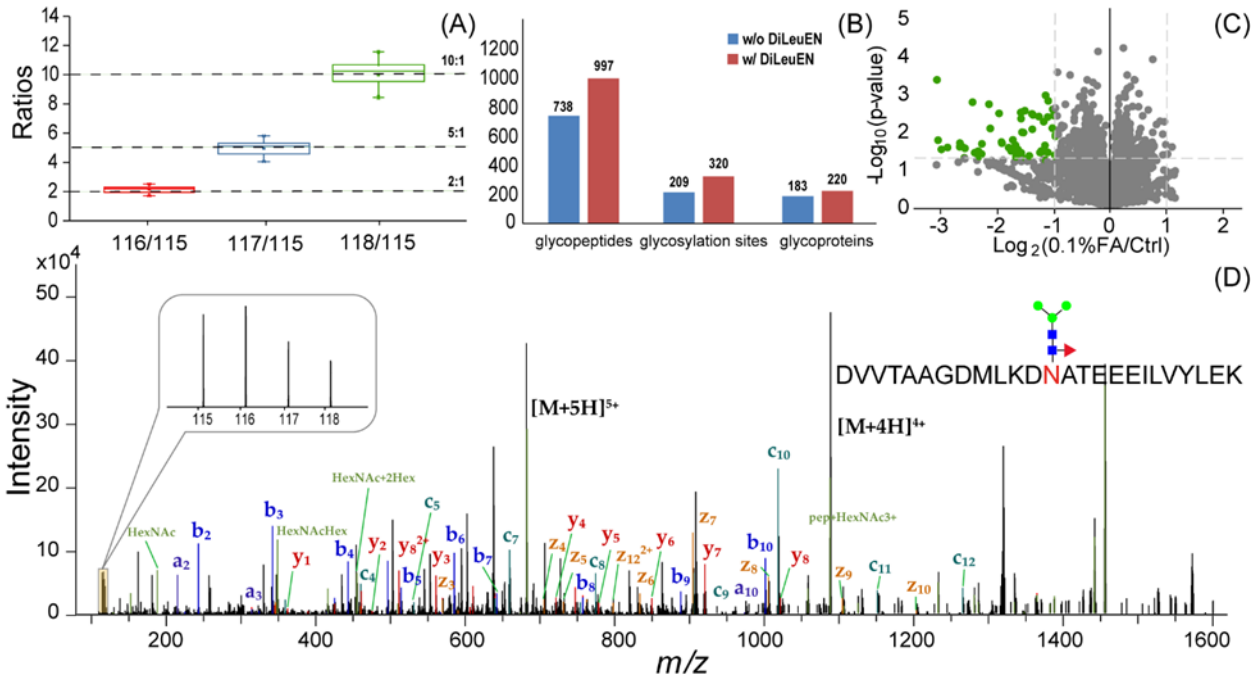


Figure 4. (A) Box plot of the relative quantification performance of 4-plex DiLeuEN labeled tryptic digest of bovine fetuin; (B) The comparison of the number of intact glycopeptides, glycosylation sites and glycoproteins between native and DiLeuEN labeled tryptic digest samples of Panc1 cell; (C) Volcano plot showing the quantification result of tryptic digest peptides of Panc1 cell incubated in 0.1% FA versus lyophilized control (37 sialylglcopeptides decayed, green dots); (D) Representative EThcD-MS/MS spectra for the 4-plex DiLeuEN labeled glycopeptide from Panc1 cell.

Supplemental Information

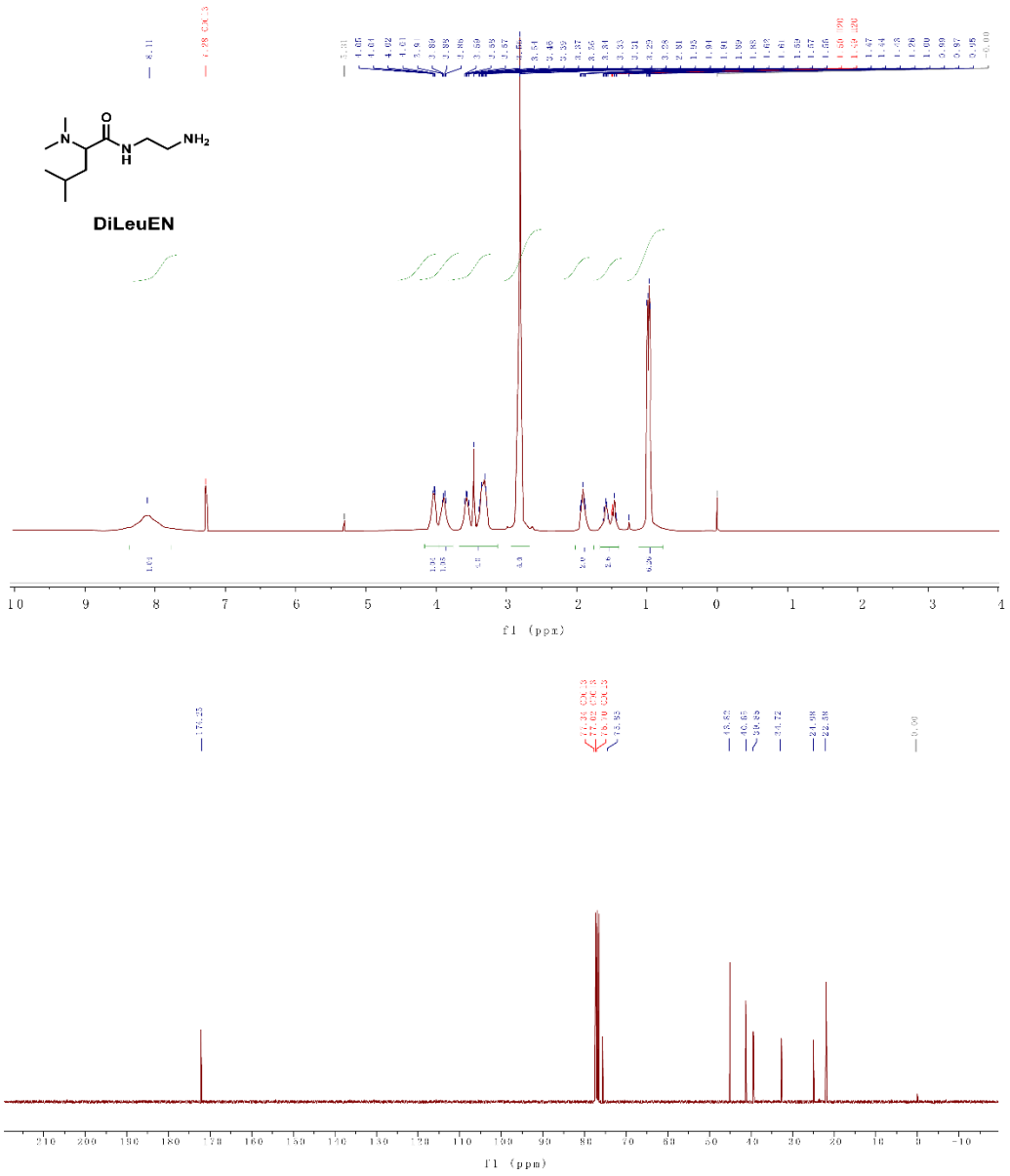


Figure S1. ^{1}H-NMR and ^{13}C-NMR of DiLeuEN

^{1}H NMR (400 MHz, $CDCl_3$): δ 0.95 (dd, J = 9.2, 6.5 Hz, 6H), 1.43-1.62 (m, 2H), 1.94 (dd, J = 8.5, 5.3Hz, 2H), 2.81 (s, 6H), 3.28-3.61 (m, 4H), 3.88 (t, J = 6.0, 1H), 4.02 (t, J = 6.0, 1H), 8.11(s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 22.6, 25.0, 34.7, 39.9, 40.7,43.8, 75.9, 174.2

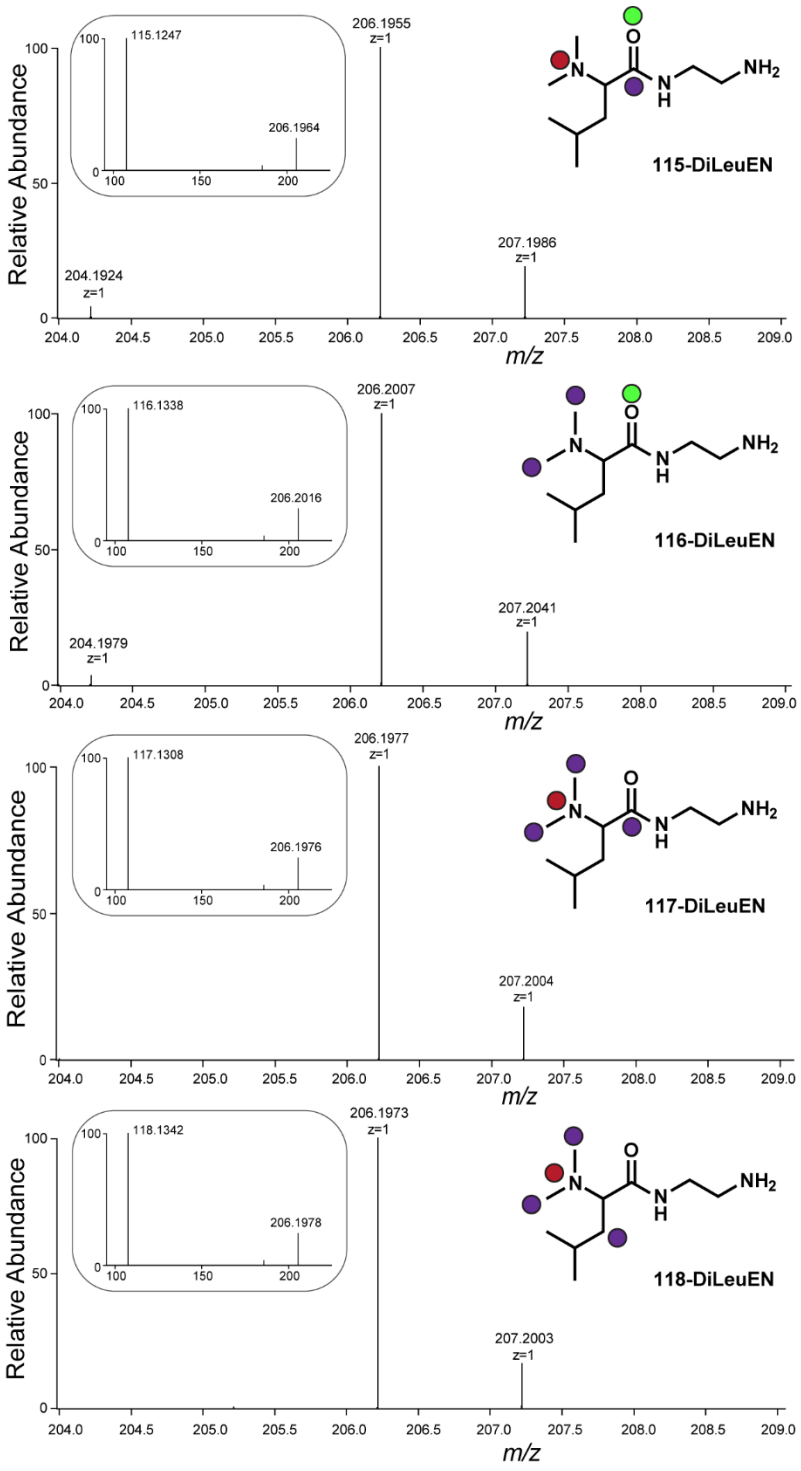


Figure S2. Isotope purities of 4-plex DiLeuEN. Inlets are MS^2 of precursor ions (HCD @25).

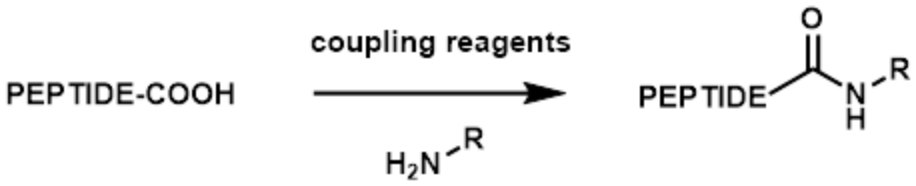


Figure S3. Scheme of charge-reversal strategy (common coupling reagents include EDCI, PyAOP, CDI).

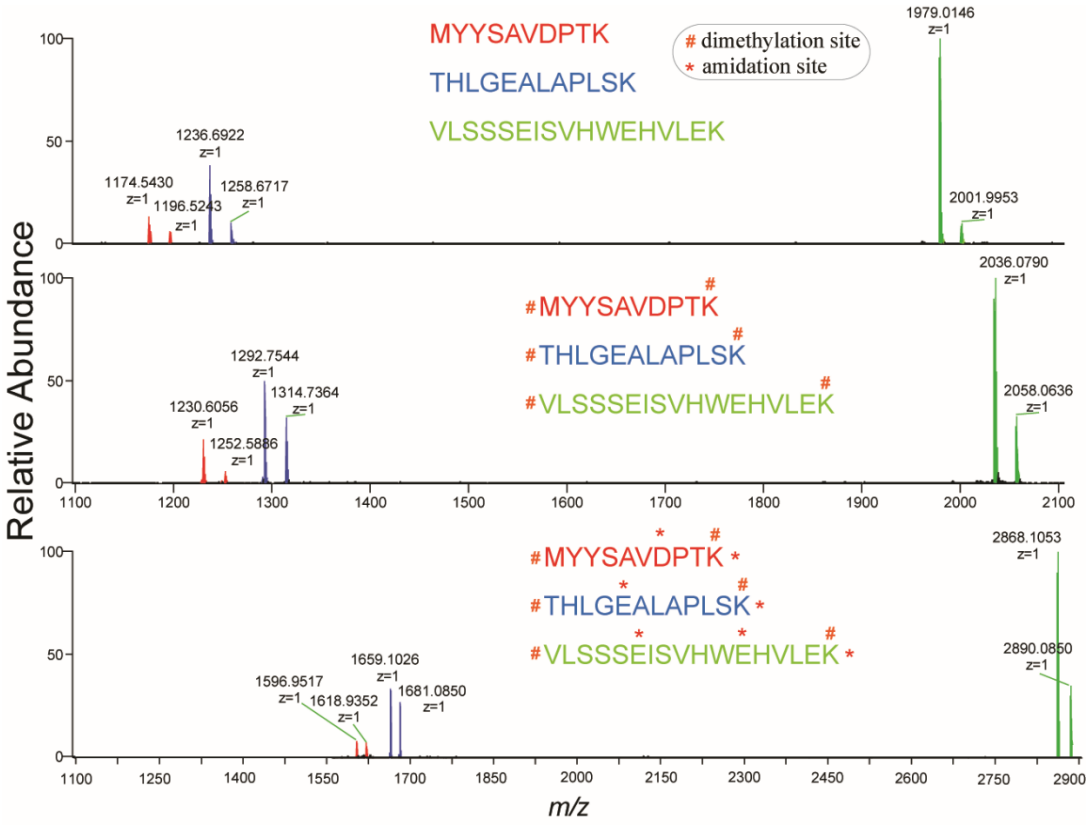


Figure S4. MALDI-MS profiling of a peptide-mix through two-step derivatization.

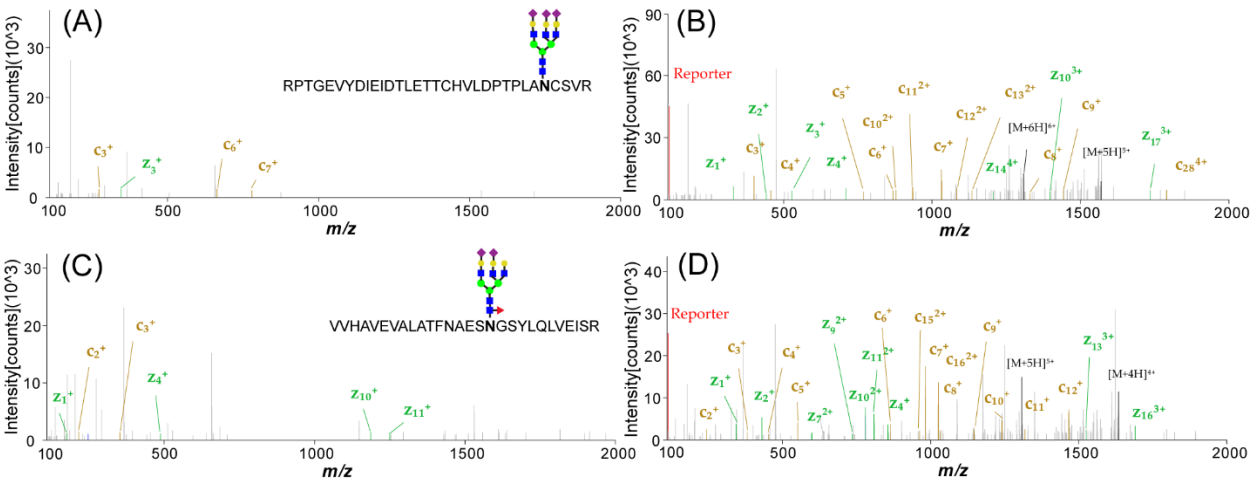


Figure S5. Representative EThcD-MS/MS spectra for the sialoglycopeptides from bovine fetuin. (A, C) underivatized, (B, D) DiLeuEN derivatized. (To emphasize ETD fragmentation improvement, b/y ions and glycan fragments are not annotated)

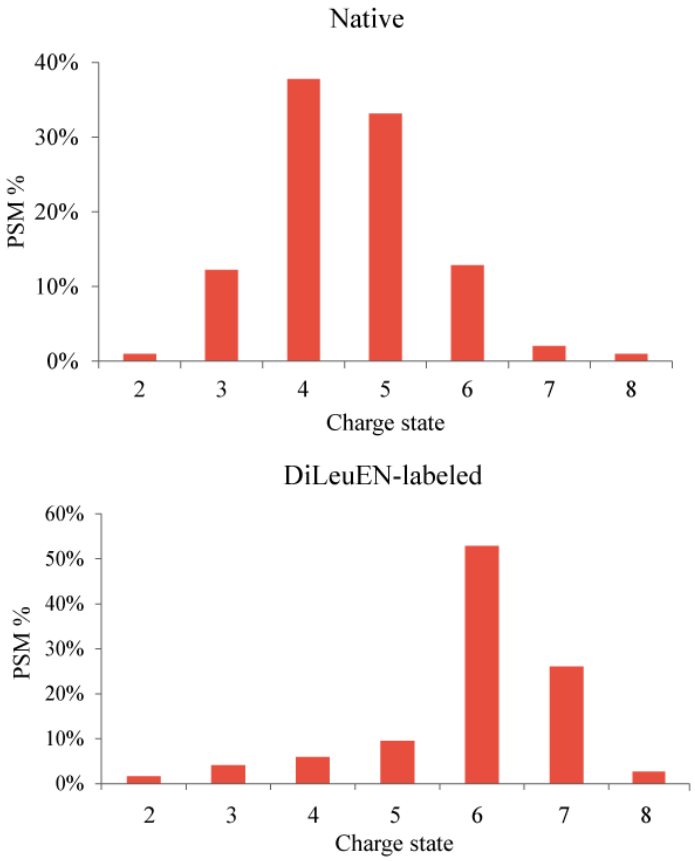


Figure S6. Charge distribution between native and DiLeuEN labeled bovine fetuin digest.

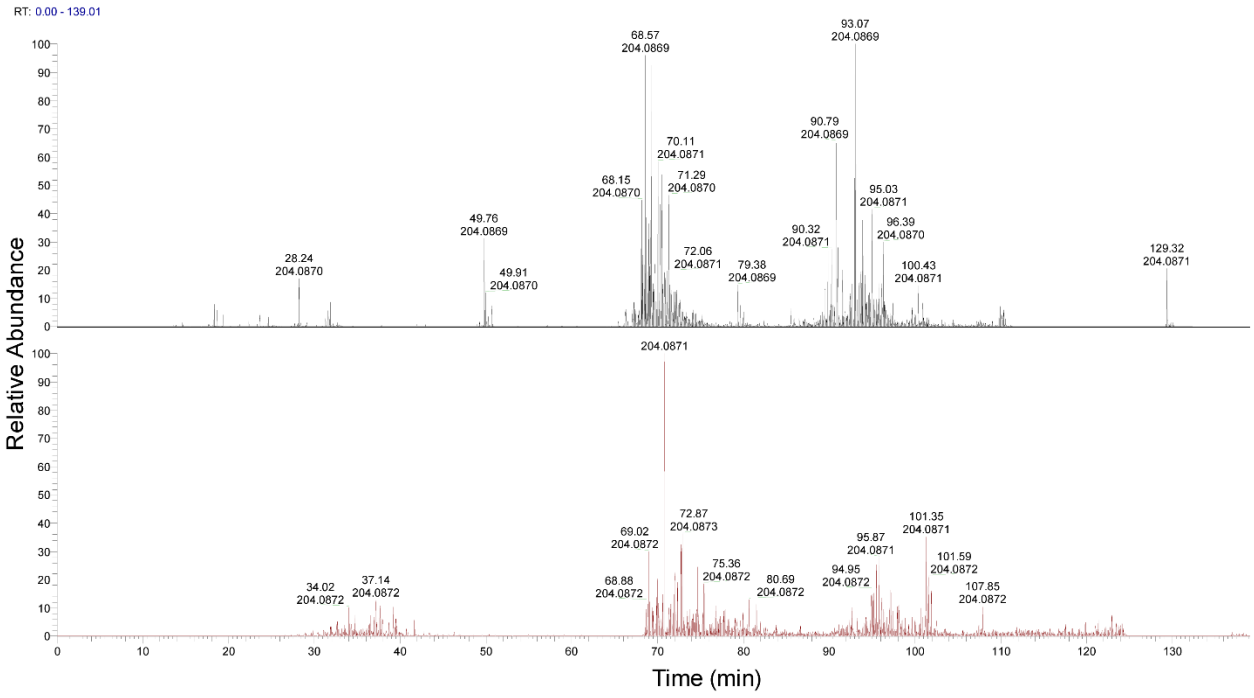


Figure S7. Extracted Ion Chromatogram **(**XIC) of glycan HexNAc (204.0872) for unlabeled (upper) and DiLeuEN labeled (lower) bovine fetuin digest.

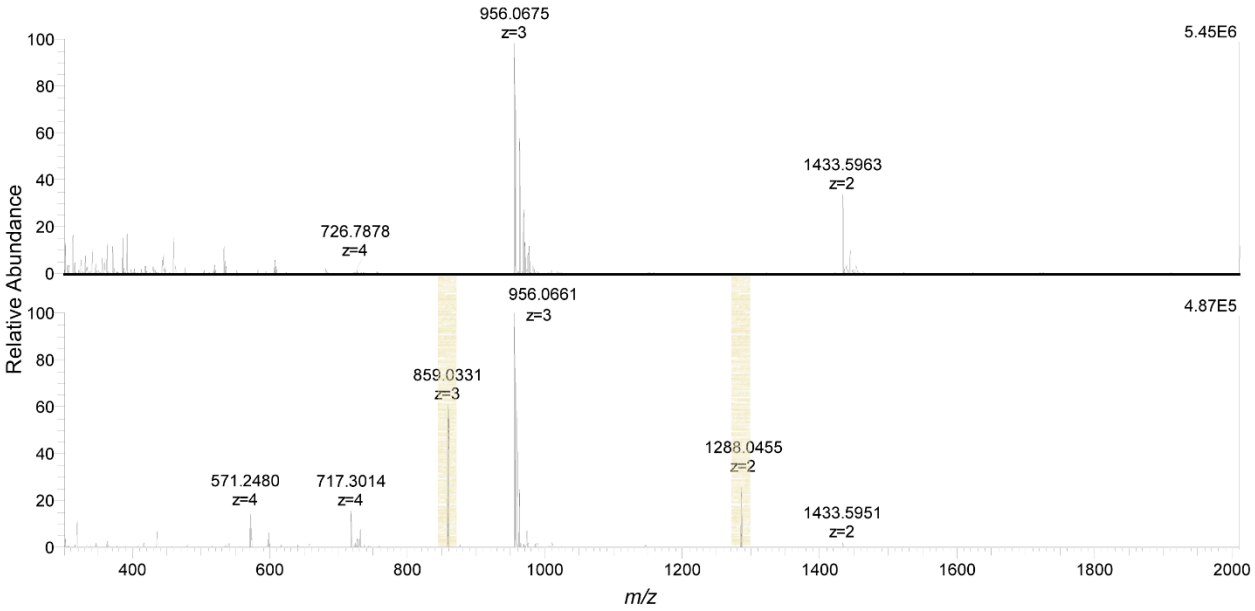


Figure S8. ESI-MS of SGP native (upper) and 0.1% FA incubated (lower). Emerging peaks corresponding to sialic acids loss.

Chapter 6

Development of Cleavable DiLeu-Biotin-Azide (cDBA) Isobaric Tags and Isobaric Tandem Orthogonal Proteolysis Activity-based Protein Profiling (isoBOP-ABPP) Platform for High-throughput Quantitative Chemical Proteomics

Abstract

The human genome project revealed less than 25,000 protein-coding genes (1). However, the total number of proteins in human proteome is estimated at over one million. Part of this complexity and diversity arises from post-translational modifications (PTMs), a process where the proteins are covalently modified during or after their assembly on the ribosome. PTM plays an important role in protein functions because they regulate activity, localization, and interaction with other cellular molecules such as glycans, nucleic acids, lipids, and cofactors (2).

Aberrant states of PTMs are frequently implicated in many human diseases, such as cancers, diabetes, and neurodegenerative diseases (3-5). In addition, it is recently recognized that PTMs act in combination on proteins for modulation and regulation purposes, and this PTMs crosstalk further expand the landscape of proteomes in eukaryotes, providing a fine-tuning mechanism in regulating protein function, localization, and interaction with other molecules (6-8). Therefore, identifying and quantifying PTMs is critical for gaining a comprehensive understanding of molecular mechanisms underlying physiological activities and human diseases. Among the protein-coding amino acids, cysteine is the most intrinsically nucleophilic amino acid and sensitive to oxidation. The pKa of the free cysteine thiol is between 8 and 9, meaning that only slight perturbations in the local protein microenvironment can result in ionized thiolate groups with enhanced reactivity at physiological pH (9). In addition to its role in catalysis, cysteine is subject to several forms of oxidative post-translational modifications, including sulphenation (SOH), sulphination (SO_2H), nitrosylation (SNO), disulphide formation and glutathionylation, which endow it with the ability to serve as a regulatory switch on proteins that is responsive to the cellular redox state (10). Elevated oxidative/nitrosative stress has been recognized as the key features of many neurodegenerative disorders such as Parkinson's diseases (PD), Alzheimer's disease (AD),

and amyotrophic lateral sclerosis (ALS) (11, 12). Mass spectrometry (MS)-based proteomics has gained great popularity in detecting and structurally defining covalent changes in a protein. However, the lack of versatile and high-throughput tools limited studies of PTM dynamics. We aim to address the following questions:

(1) How to achieve identification and quantification of PTM globally in a high-throughput manner with economic cost and simple procedures?

(2) Accumulating evidence suggest abnormal S-nitrosylation of several proteins are involved in AD pathogenesis. How do global S-nitrosylation profiles change in cerebrospinal fluids (CSF) during different stages of AD progression?

(3) Will a novel chemical tool offer versatile analytical capacity for various PTMs? Can we use it to tackle PTM crosstalk and elucidate how combination of multiple PTMs can be generated and affect protein functions and interactions?

In order to address these challenges and fill in existing knowledge gaps, we divided this project into three stages:

Stage 1: To design and synthesize multiplexed DiLeu-Biotin-Azide (DBA) tag and establish isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP) method to facilitate quantitative PTM analysis. We will first design and synthesize several candidates of DiLeu-Biotin-Azide (DBA) tag, followed by systematic evaluation of the performance of these candidates including labeling/enrichment efficiency and fragmentation pattern in tandem mass spectrometry (MS/MS). The structure of the best candidate will then be selected as the final choice, and isotopic version of DiLeu will be incorporated to construct a set of isobaric multiplexed DBA tags.

Stage 2: To discover and investigate global S-nitrosylation in Alzheimer's diseases using multiplexed DBA tag based on quantitative proteomics approach for a better understanding of molecular mechanisms underlying neurodegenerative diseases. S-nitrosylation plays a pivotal role in many neurodegenerative diseases and the DBA tag offers a high-throughput quantitative tool for global analysis of S-nitrosylation. Quantitative proteomics will be conducted using DBA tag on cells samples from AD mouse model and CSF samples from human patients. Bioinformatics analysis will help to identify differentially expressed proteins or modification sites as putative biomarkers or targets followed by subsequent biological validation.

Stage 3: To expand the scope of application of DBA tag and establish a pan-PTM high-throughput quantitative proteomics platform for various PTM study and crosstalk investigation. The utility of DBA tag can be further explored due to the exquisite design of its chemical structure and unique features of biorthogonal reactions. Besides various cysteine oxidation, protein carbonylation and citrullination are also amenable to chemical tagging. Thus, by altering alkyne probe to chemical selectively label different PTM sites, various PTM could be derivatized, captured, enriched, and analyzed. The versatility of the application of DBA tag makes it an ideal pan-PTM high throughput quantitative analysis platform. Furthermore, the combination of different PTM constitutes the next level of complexity for proteomics research and this crosstalk research field could also be benefited from our pan-PTM proteomics platform.

Introduction

Nitrosative stress and the role of protein S-nitrosylation in AD.

Numerous studies suggest that the dysregulated redox signaling is crucial in the pathophysiology, inflammatory responses and neuroprogressive nature of major depression (13-15). Reactive oxygen and nitrogen species (ROS and RNS), including peroxynitrite, superoxides, peroxides and nitric oxide (NO), are produced during normal physiological processes and through interacting with proteins, fatty acids and DNA, perform numerous roles in regulation of cellular function. However, excess ROS/RNS can lead to structural and functional changes resulting in cellular injury. These potentially toxic effects are offset under normal physiological conditions by intrinsic antioxidant mechanisms but increased oxidative and nitrosative stress (O&NS) may cause damage to cellular components, induce harmful autoimmune responses, and ultimately lead to failure of normal cellular processes.

NO, as an important RNS, is a small, highly diffusible signaling molecule generated by three different isoforms of NO synthase (NOS) from L-arginine and NADPH in mammalian cells (16). NO induces a PTM of proteins, called S-nitrosylation. S-nitrosylation is formed via the reaction of NO with thiyl radical or between NO+, which is one electron oxidized form of NO, and the thiol group of cysteine residues in the target protein (17). This nitrosylation reaction forms an S-nitrosothiol (-SNO) and is commonly reversible. It is important to note that S-nitrosylation has unequivocally been shown to function not only in redox homeostasis and toxicity, but also in conveying or regulating physiological cellular signals (18).

Elevated levels of nitrosative stress were found in human AD brains (19). Genetic mutations associated with neurodegenerative diseases, environmental toxins such as certain pesticides, and misfolded proteins including Aβ oligomers can all lead to excessive nitrosative stress. The resulting aberrant S-nitrosylation of numerous proteins has been implicated in the pathogenesis of several neurodegenerative disorders. For example, S-nitrosylation of XIAP and

PDI can cause ‘loss-of-function’ by impairing E3 ligase ubiquitin ligase activity and molecular chaperone activity, respectively (20, 21). These processes can contribute to accumulation of neurotoxic proteins and activation of apoptotic pathways. Additionally, oligomeric Aβ peptide can result in increased generation of NO and S-nitrosylation of Drp1, resulting in excessive mitochondrial fragmentation, bioenergetic compromise, and consequent synaptic damage. Additionally, Cdk5 is activated when Aβ increases calpain activity to cleave the Cdk5 regulatory subunit p35 to p25. The resulting neurotoxic kinase activity of Cdk5 is further enhanced by S-nitrosylation. Formation of SNO-Cdk5 may also contribute to spine loss by transnitrosylating Drp1, with the resultant SNO-Drp1 participating in mitochondrial fragmentation (22-24). Overall, aberrant S-nitrosylation can trigger many neurodegenerative signaling pathways. Here, we will utilize novel chemical proteomics methods to investigate S-nitrosylation and study its role in Alzheimer’s diseases.

Isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP).

While human genome sequencing and genetics enable making direct connections between mutations and human disorders at an unprecedented rate, there has been a widening gap between the number of genes and their encoded proteins that have been linked to health and diseases. As a matter of fact, only around 2% of all predicted human gene products are currently targeted with small-molecule drugs and only 10–15% of all human proteins are ‘druggable,’ with only a 25% overlap between druggable protein targets and known disease-modifying targets (25, 26). This discrepancy is in part because many proteins may not have any obvious binding pocket for pharmacological interrogation or because of a lack of high throughput screening technologies for functional assays to identify small-molecule modulators against these targets. The advent of advanced strategies of chemical proteomics and the activity-based protein profiling (ABPP)

approach has empowered significant expansion of the scope of proteins that can be pharmacologically evaluated in living systems, thus promoting identification and prioritization of new therapeutic targets (27-30).

In a typical ABPP experiment for target identification, ABPP probes are first incubated with living cells or cell lysates (**Fig. 1**), along with appropriate negative controls (such as DMSO, inactivated drug, or the original compound in competitive assays). Following target protein binding (based on specific activity between probe and substrate proteins), the probe–protein complexes are enriched through affinity purification against the probe. Next, nonspecific binding proteins are removed through multiple washes before target proteins are eluted. The protein targets are then identified with methods including gel electrophoresis and MS. The use of negative controls allows for relative quantification and mitigates the interference of background and nonspecific binding proteins to some extent. ABPP probes generally have three fundamental components: (i) a reactive group for binding or modifying the target proteins at certain sites, such as the active sites of a given enzyme class or the drug-binding pockets; (ii) a reporter moiety for identification and purification of binding proteins; (iii) a linker, occasionally cleavable, extending the distance between the reporter and the reactive group to avoid steric hindrance (31, 32).

Uniquely, Cravatt and colleagues designed an isotopic tandem orthogonal proteolysis–activity-based protein profiling (isoTOP-ABPP) strategy to profile the intrinsic reactivity of cysteine residues in native proteomes (33, 34). The isoTOP-ABPP approach employed biorthogonal reaction—the copper(I)-catalyzed azide/alkyne cycloaddition (CuAAC) reaction—to bridge the probes and proteins which have been modified with precursor alkyne handle molecules in advance. In addition, the linker containing a tobacco etch virus (TEV)-protease recognition peptide could be efficiently cleaved after enrichment and the isotopically labelled

valine could be used for relative comparison in MS measurement (35). However, the bulky nature and synthetic complexity of its chemical structure limited its widespread application and the quantification strategy using isotopic labeling based on MS1 limited its multiplex capacity.

By incorporating isobaric labeling strategy based on fragment ions in MS/MS scan, high-throughput quantitative analysis can be achieved. Commercially available tandem mass tags (TMT/TMTpro) and isobaric tags for relative and absolute quantitation (iTRAQ) have become a mainstream technique in relative protein quantification by examining up to 16 biological samples simultaneously in a single liquid chromatography mass spectrometry (LC-MS) run (36-38). The in-house-constructed *N*,*N*-dimethyl leucine (DiLeu) tags developed in our group are a cost-effective alternative to TMT and iTRAQ tagging approaches (39). Here, we will combine ABPP approach and DiLeu isobaric labeling strategy to establish a novel method coined isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP) to achieve targeted protein screening as well as high-throughput quantitative analysis.

Pan-PTM high-throughput quantitative proteomics platform.

As PTMs play important roles in many physiological and pathological processes, accurate and convenient analysis of each type of PTM is of high significance and relevance. However, due to the various types of PTMs and their diverse functions and localizations on proteins, there is no universal method capable of analyzing every/each type of these PTMs. Most PTMs can be detected by MS-based proteomics, either as a mass increment or a mass deficit relative to the nascent unmodified protein, and MS/MS enables characterization of modified proteins via amino acid sequencing and site-specific localization of post-translationally modified amino acid residues. But the general low abundance of PTMs usually require upstream enrichment to aid LC-MS analysis, where different strategies including affinity purification, antibody recognition and chemical

modification are employed (40-45). Moreover, MS is not inherently quantitative, which demands the introduction of stable isotopes into the analytes of interest based on stable isotope dilution principle (46, 47). With the development of biorthogonal reaction, artificial chemical probes could be incorporated into biomolecules with greater flexibility and hence more sophisticated architecture of the chemical structures could be designed and implemented. Here, we create a pan-PTM high-throughput quantitative proteomics platform based on the isoBOP-ABPP method integrating selective PTM probing, efficient PTM enrichment, and isobaric labeling for quantification that allows versatile PTM analysis and potential application to the study of PTM crosstalks in complex biological systems.

Methods and Approaches

Multiplexed quantification with DiLeu isobaric tags.

4-plex DiLeu isobaric tags. DiLeu tags were originally developed by our lab as a novel, cost-effective alternative to commercial isobaric tags (e.g., iTRAQ, TMT) (48). The structure of DiLeu resembles other isobaric tags in that it features a reporter group, a mass-balance group, and a reactive group for labeling of primary amines on peptide N-termini and lysine side chains. The first generation of the 4-plex DiLeu isobaric tags yield reporter ions at *m/z* 115, 116, 117, and 118 upon MS/MS analysis of labeled peptides. The multiplex set can be synthesized in-house using established chemistry in only a few steps with significant cost savings, while offering excellent labeling efficiency and enabling accurate relative quantification (49, 50).

Mass defect-based 12-plex DiLeu isobaric tags. We recently expanded the multiplexing capacity of the DiLeu reagent to enable simultaneous quantification of up to twelve samples in a

single MS/MS spectrum using high-resolution MS platforms (49). We borrowed the concept of neutron encoding (NeuCode) method originally developed by Coon and colleagues (51). Mass defect, the cause of a subtle mass change, arises from the fact that nuclear binding energy, the energy required to break down a nucleus into its component nucleons, is different for each isotope of every element (52). Through calculated incorporation of 13C, 15N, and 2H stable isotopes in the reporter structure, we were able to design eight additional reporter isotopologues that differ in mass from the original four reporters by ~6 mDa to yield an isobaric 12-plex set consisting of two 115 labels, three 116 labels, three 117 labels, and four 118 labels. In doing so, we tripled the multiplexing capacity without increasing synthetic complexity or requiring custom isotopic reagents. Resolving each of the 12-plex DiLeu reporter ions requires an MS/MS resolving power (RP) of at least 30,000 (@ 400 m/z), which is readily approachable for modern orbitrap mass spectrometer. Multiplexing offers the ability of parallel sample processing, thereby reducing the number of MS runs and increasing the overlap of proteins detected and quantified across many experiments. In addition, the concise chemical structure of DiLeu and simple synthetic scheme offer a great opportunity for further developing them into advanced architecture for sophisticated applications (53, 54).

Design and synthesis of the prototype of DBA tag and evaluation of its performance.

The DBA tag has three unique features: (i) It contains a biotin moiety which is widely used as an enrichment group. The biotin-streptavidin interaction is the strongest noncovalent biological binding known so far, having a dissociation constant, Kd, in the order of $4x10^{-14}$ M. The strength and specificity of the interaction render it to be one of the most widely used affinity pairs in molecular, immunological, and cellular assays (55); (ii) It contains an azide group which serves a conjugation site for alkyne modified molecules, where copper(I)-catalyzed azide/alkyne

cycloaddition (CuAAC) reaction, known as click chemistry, is readily performed with high efficiency (56, 57); (iii) It contains a DiLeu moiety as a reporter group, which would facilitate MS/MS based relative quantification. Combining all three features together, various chemical structures can be conceived; considering synthetic difficulty, one of the most concise and simplified structure was chosen and synthesized as a prototype (**Fig. 2**). This prototype DBA tag has a linear structure, in which one end is the biotin moiety for enrichment and the other end is DiLeu reporter ion, with an azide handle branched out in the middle. It can be made with four steps of simple organic reactions at moderate yield, which is relatively easily accessible for common analytical lab settings. In addition, the incorporation of DiLeu reporter ion group was set as the final step for the synthetic route, which is intended for modularized synthesis of multiplex isotopic channels. To evaluate its performance, MS/MS of the tag itself was performed under higher collisional dissociation (HCD) mode at normalized collision energy (NCE) 30, intense reporter ion signal was observed corresponding to the cleavage of DiLeu and biotin fragments. Both of these two peaks could be used as signature of modification. For chemical labeling condition, standard click chemistry protocol utilizing $CuSO_4$ and sodium ascorbate to general Cu(I) *in situ* which catalyzed alkyne and azide cycloaddition was employed (58). The alkyne modified biomolecules were prepared using standard peptide—atrial natriuretic peptide (ANP)—which contains a free cysteine residue in the middle of the sequence and thiol-reactive probe 2-indo-N-(prop-2-yn-yl)acetamide (IPM). The reaction parameters such as concentrations for the substrates, reaction temperature, and reaction time were further optimized.

Enrichment efficiency comparison between DBA tag and iodoTMT and total cysteine analysis in HeLa cell line.

As mentioned above, the DBA tag shall possess three unique features, among which the most useful and distinct one is enrichment enabled by biotin-streptavidin interaction. According to the design, alkyne-modified biomolecules are converted to DBA probe derivatized form, thus containing biotin moiety and could be captured by streptavidin modified solid phase extraction (SPE). Those non-alkyne-modified molecules are washed off and discarded. To investigate the enrichment efficiency, two model proteins—bovine serum albumin (BSA) and β-casein—were used. BSA has multiple cysteine residues within its sequence, while β-casein only has one cysteine residue in its sequence. When spiking tryptic digested BSA to abundant β-casein, cysteine residues only exist as a minor portion of total peptides. By using IPM (an IAA-like alkyne molecule) to label BSA / β-casein mixture, cysteine residues were installed an alkyne handle, which were then conjugated to the DBA tag through click chemistry. Enrichment efficiency can be calculated by taking ratio of abundance of cysteine containing peptides / total peptides. IodoTMT is a set of commercially available isobaric tags that enable multiplexed quantification of cysteine-containing proteins using tandem mass spectrometry (59, 60). Therefore, we compared our DBA tag against iodoTMT. As shown in **Fig. 3a**, DBA tag exhibited comparable or slightly better enrichment performance compared to iodoTMT.

In addition, among the many developed methods for the detection of SNO, the biotin switch technique (BST) reported by Jaffrey et al. is the gold standard (61). BST and its various modified forms represent indirect detection strategies, in which S-nitrosothiols are selectively reduced with a mild reducing reagent (e.g., ascorbate) and labeled with another tag forming a stable bond for further detection, enrichment, or quantification on the same site (62, 63). Therefore, by performing total cysteine analysis in proteome, the potential application of DBA tag to probe SNO could be assessed. HeLa cell lines were used as complex biological samples. Around 15% of

tryptic peptides are cysteine containing peptides in *homo sapiens* (64), this was validated by cysteine analysis without enrichment (**Fig. 3b**). After DBA tag labeling and enrichment, majority of peptides were removed, and the percentage of cysteine-containing peptides could reach to above 80%. This result suggested that our new approach provides the capability to identify protein S-nitrosylation sites and quantify changes in S-nitrosylation PTMs by mass spectrometry.

Optimization of instrumental parameters for DBA tag labeled peptide analysis.

Liquid chromatography−tandem mass spectrometry (LC−MS/MS) strategies are frequently used to identify specific PTMs. However, labile PTMs such as S-nitrosylation are unstable during collisional induced fragmentation, thus requiring certain chemical derivatization. Chemical derivatization can not only stabilize the primordial modification sites but also change chemical physical properties of peptides on LC and/or MS. Detailed optimizations of instrumental parameters are necessary for obtaining optimal analysis results (65). We performed PTM analysis on Orbitrap Fusion Lumos, a commercialized tribrid mass spectrometer, where precursor ions can be isolated in either quadrupole (Q) or ion trap (IT), and the fragment product ions can be analyzed in both low-resolution ion trap and high-resolution Orbitrap analyzers (66). This lays the foundation of isobaric quantification based on minute mass differentiation of fragment ions. Higher resolution can increase mass accuracy and the sensitivity of low abundance precursor ion detection but decrease the data acquisition rate. Thus, there is always a trade-off between mass spectral resolution and instrument duty cycle. Many other parameters including mass resolution, maximum injection time, isolation window, automatic gain control (AGC) target of MS2 and normalized collision energy (NCE) were screened and optimized further. All the optimizations were performed in top speed mode, which could maximize the number of high-quality MS spectrum acquisition within each duty cycle. Selected reaction monitoring (SRM) -- also called

multiple reaction monitoring—is emerging as a targeted proteomics approach that a predefined precursor ion and one of its fragments are selected by the two mass filters and monitored over time for precise quantification (67). DBA tag labeled peptides typically generated signature fragments of DiLeu during MS/MS, which commonly used as reporter ions for relative quantification, but could also be utilized for SRM method development. In addition to characteristic reporter ion, those unique precursor peptides representing targeted proteins are essential information need to know in advance.

Results and Discussion

Design and synthesize of multiplexed DiLeu-Biotin-Azide (DBA) tag and establish isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP) method to facilitate quantitative analysis of post-translational modifications (PTM).

Chemical proteomics, in particular activity-based protein profiling (ABPP), combining with bio-orthogonal click chemistry, is widely utilized both *in vitro* and *in vivo* (68-71). ABPP is an ideal strategy that can faithfully recapitulate protein-small molecule interactions *in situ* (i.e., in live cells), and tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP) employs on-bead trypsin and TEV digestions to simultaneously identify both probe-labelled proteins and their exact sites of probe modification at the same time (72, 73). However, the bulky structure of the tag in classical isoTOP-ABPP and the reliance on full-MS quantification has limited its usage in broader PTM proteomics studies and high-throughput analysis. Here, we incorporated our in-house constructed DiLeu isobaric tags into the structure of a novel chemical probe, DBA tag,

establishing an isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP) method to overcome the limitations of isoTOP-ABPP.

As shown in **Fig. 4**, several candidates are designed and synthesized as the prototype of DBA tags (protoDBA). DBA tag contains three parts: (i) DiLeu as reporter ion in MS/MS; (ii) biotin moiety for streptavidin enrichment and (iii) an azide group for probing alkyne-modified substrates through CuAAC click chemistry. The general architecture of the tag is a linear molecule with a short side chain. DiLeu and biotin groups are placed on two ends of the molecule with azide functionality left at the terminus on the side chain. Candidate A is the most compact one containing only three essential moieties bridged by an ethanediamine as a linker. While candidates B-E possess an extra feature, a cleavable linker between biotin and DiLeu/azide groups. This is for efficient and convenient release of captured biomolecules from the biotin-streptavidin complexes. Although there were reports suggested that the biotin-streptavidin interaction could be broken to recover intact biotin and streptavidin parts respectively (74, 75), it seems that the mainstream techniques tend to avoid putting cumbersome biotin parts on the conjugates before LC-MS/MS (76-79). Therefore, we designed two different structures for cleavable DBA tag. The candidate B has an acid sensitive linker mimicking amine protecting group Boc which could be hydrolyzed under acidic condition, whereas candidate C contains a diol group which can be cleaved upon mild peroxide oxidation. The candidate D possess a basic labile ester linker which can be facilely released under mild and volatile basic environment – 5% ammonium hydroxide, and the candidate E has a popular cleavable linker DADPS which is acidic labile and utilized in many chemical probes. Some other candidates such as the one with photocleavable linker can also be proposed (80). After candidates are synthesized, their performance was evaluated using peptide standard and the optimal structure shall then be determined.

Design and synthesize the prototype DBA tags in a compact form and concise manner.

The general architecture of DBA tag is designed as a linear molecule with two essential parts—DiLeu and biotin groups—on each end of it, while the azide moiety is at the terminus of the side chain. This design borrows from the first chemical *in vitro* labeling method that uses a biotin tag to label proteins containing cysteine residues, isotope-coded affinity tag (ICAT) (47, 81). The major novelty is modular synthetic design which makes incorporation of isotopes facile and straightforward. The construction of backbone structure starts from biotin acid, through amide coupling reaction, the structure is extended allowing functionalities to be installed. Tactfully, the fusion of DiLeu group is the last step, where isotopic version of DiLeu could be utilized for the synthesis of multiple channels of isobaric DBA tag. Considering the difficulty of recovering biotin from streptavidin beads, in addition to a compact structure Candidate A, we designed another four cleavable structures Candidate B - E, in which either acid/base hydrolysis or mild oxidation could be performed to efficiently remove biotin moiety from the conjugated biomolecules.

Synthesis of Candidate A:

Biotin acid was coupled with N-Boc-ethylenediamine firstly before the Boc protecting group was removed in acidic condition. The biotin-amine was then coupled with a 6-Azido-N-Fmoc-lysine to produce a biotin-azide module 1 after deprotecting Fmoc group. At last, dimethylated leucine was installed onto 1 to make the final product Candidate A. It takes six steps in total and by far is the shortest synthetic route to make a compact DBA tag. (**Fig. 5a Route A**)

Synthesis of Candidate B:

Tertiary alcohol was introduced into biotin acid first. The acid cleavable moiety, while requiring an electrophilic locus, also necessitated the presence of a terminally protected primary

amino group for the attachment of the remaining DiLeu-Azide part. To this end, we envisaged utilizing chlorocarbonate reagent for modifying tertiary alcohol which could be substituted subsequently by an azide amino acid (82). When azide module is installed, DiLeu group shall be incorporated lastly. In brief, N-Boc-ethylenediamine is used to convert carboxyl group of DiLeu into an amine functionality, after which the condensation between amine on DiLeu module and carboxyl group on Biotin/Azide module could be implemented to afford the final product. The overall route takes seven steps. (**Fig. 5a Route B**)

Synthesis of Candidate C:

Photocleavable linker (83) and diol linker are both widely adopted cleavable linker in chemical proteomics. To synthesize a DBA tag with a diol linker, 1,4-diaminobutane-2,3-diol was used as the core. By using amide coupling, the core can be extended with biotin and azide lysine to produce the precursor intermediate. Then the intermediate product was deprotected before labeling with DiLeu to get the final product. The route takes four steps in total. (**Fig. 5b Route C**)

Synthesis of Candidate D and E:

The candidate D and E shared same synthetic route in part. The DiLeu was conjugated to azide lysine first following two-step amide coupling and two-step deprotection reaction. The intermediate product with a primary alcohol group was then linked with either a carboxylic acid to afford an ester or labeled with DADPS to afford a base labile cDBA tag. The two routes take five or six steps, respectively. (**Fig. 5b Route D and E**)

Evaluation of the performance of prototype DBA tags and determination of the optimal structure.

As mentioned above, the designed DBA tag shall process three characteristics: efficient labeling against alkyne group by click chemistry; competent enrichment performance provided by biotin-streptavidin interaction and unique DiLeu reporter ion generated in MS/MS for relative quantification. Therefore, the performance of synthesized candidates will be evaluated from the following aspects: (i) labeling condition and efficiency; (ii) enrichment efficiency and (iii) fragmentation behavior in MS/MS. Standard peptides or proteins could be used for evaluation and the one with the best performance will be chosen as the optimal structure of DBA tag.

i. Labeling condition and efficiency.

Atrial natriuretic peptide (ANP) was used as a standard peptide for method development and labeling condition optimization. ANP is a natriuretic peptide hormone secreted from the cardiac atria and the main function of ANP is causing a reduction in expanded extracellular fluid (ECF) volume by increasing renal sodium excretion (84). The first 28 amino-acid sequence contains a disulfide bridge (Cys:7-23), in which a shorter peptide with a free cysteine residue in the middle of its sequence could be prepared through reduction and digestion. This segment of ANP was labeled by IPM first to convert thiol group on cysteine residue into an alkyne handle, and protoDBA was then tagged with alkyne modified ANP segment through CuAAC click chemistry. The labeling conditions including substrate concentrations, choice of metal ligands, temperature, reaction time etc. were screened and carefully optimized as in previous studies (85, 86). The reaction was monitored by matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS). As shown in **Fig. 6**, both alkylation and click chemistry exhibit excellent labeling efficiency. Mass spectral peaks corresponding to the modified mass were clearly observed with minimal signal of reactant barely detected. This result suggests the high

efficiency of click chemistry and thiol alkylation, making it readily applied to cysteine modification analysis.

ii. Enrichment efficiency.

Enrichment performance was evaluated at the peptide level first. By spiking protoDBA labeled ANP peptide (**Fig. 7a**) into a peptide mix (**Fig. 7b**), the new peptide mixture constitutes an ideal environment for enrichment performance test (**Fig. 7c**). Because only modified ANP contains a biotin group, the other peptides cannot be bonded onto streptavidin beads and were washed away. Only protoDBA labeled ANP was preserved in elution buffer (**Fig. 7d**). To test enrichment performance on a more general level, standard proteins BSA and β-casein were used. Tryptic digestion was performed first on both proteins and then BSA was spiked into β-casein at 1:5 and 1:10 ratio. Since BSA is rich of cysteine residues within its sequence while β-casein barely contain any cysteine residue. When performing enrichment protocol similar to that applied to the peptide mixture, the signal intensities from cysteine-containing peptides over total peptides can be used to calculate the enrichment efficiency. As illustrated in **Fig. 3**, comparable results were acquired compared to commercially available reagent, iodoTMT. These results demonstrated that the biotin enrichment moiety worked well as designed.

iii. Fragmentation behavior of labeled peptides in MS/MS

Fragmentation of protoDBA tag itself generates two signature ions at *m/z* 114 and *m/z* 287, corresponding to the fragments of DiLeu and biotin backbone (**Fig. 2**). These two signature ions can be also observed at the MS/MS of protoDBA tag labeled peptides from BSA. By altering normalized collisional energy (NCE) to an optimal value, abundant b/y ions of peptide backbone fragmentation accompanied with decent signals of two reporter ions were acquire in a single

MS/MS spectrum. As illustrated in **Fig. 8**, the site of which derivatized by protoDBA could be unambiguously assigned, for both singly tagged and doubly tagged species, in particular, for the two adjacent cysteine containing peptides. This result again confirmed the efficient labeling process and demonstrated robust MS/MS fragmentation behavior of protoDBA tag, which will benefit site specific identification and fragment-ion-based relative quantification.

Multiplexed DBA tag synthesis for high-throughput analysis of PTM.

Since the modular synthesis allows DiLeu module to be installed at the last step, the intermediate product containing Biotin/Azide backbone could be prepared in a large quantity. The isotopic version of DiLeu can then be incorporated into the synthesis to make a set of multiplexed isobaric DBA tags. The unique features of DiLeu as a tandem mass tag include strong reporter ion signals and economic synthesis, as well as the introduction of isobaric moiety in a single step. This feature can minimize the cost of required isotopic reagents and make multiplexed DBA tag affordable. The first generation of DiLeu has four channels and by combining the Neucode concept, it was developed into the second generation 12-plex (**Fig. 9**) and the third generation 21-plex (87). Higher mass spectral resolution is needed to fully resolve reporter ions from different channels as the throughput increase. However, ultra-high resolution commonly accompanied with increased duty cycle with slower data acquisition rate and reduced target identification. Here, 12-plex DiLeu tags are chosen for multiplexed DBA synthesis and a set of 12-plex isobaric DBA tags was designed and constructed based on the structure of candidate D.

Establishment of isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP) workflow based on DBA tag.

To identify the specific reactive amino acid sites of the target protein by using small molecules, Cravatt and co-workers developed a strategy called isotopic tandem orthogonal proteolysis–ABPP (isoTOP-ABPP) based on TEV tag (33). It was then quickly applied to many areas including identifying disease-relevant targets, mapping hyper-reactive and functional hotspots and developing inhibitors against ligandable hotspots (88-91). An isotopically light or heavy valine in TEV tag was employed for quantitative ratiometric MS-based proteomics analysis. Obviously, TEV tag relies on isotopic labeling and can only compare two samples at a time. The DBA tag we proposed here employs isobaric labeling and thus can expand the quantitation channels for high-throughput analysis. We coined our method "isoBOP-ABPP" which stands for isobaric tandem orthogonal proteolysis activity-based protein profiling based on DBA tag. The general workflow of isoBOP-ABPP can be established and it is resembling isoTOP-ABPP. Briefly, proteomes are labeled *in vivo* or *in vitro* with an alkyne probe toward a particular enzyme class or amino acid modification sites. Then enzymatic digests of the proteomes are labeled with different channels of DBA tags through CuAAC click chemistry. Since the DBA tags are in excess and click chemistry is highly efficient, alkyne probe modified proteome in each sample will be thoroughly converted to isobaric DBA tagged forms. Subsequently, different samples were combined followed by biotin-streptavidin enrichment prior to running LC-MS/MS analysis. Premix ratio test starts with same amounts of proteins or peptides and identical procedure is implemented for all channels. The results of premix ratio test could be used for correction of the ratio obtained from the real sample analysis.

We have already synthesized Candidate A as the protoDBA tag and evaluated several crucial performances. We anticipate successful development of multiplexed DBA tag eventually and do not expect technical difficulties in establishing the isoBOP-ABPP platform. These multiplexed

isobaric DBA tags can be used as a competing product of TEV tag and offer the ability to quantify up to 12 samples in a single experiment via high-resolution MS. However, the Candidate A has the potential challenge of coeluting with peptide analytes in enrichment as it relies on the release of biotin from streptavidin beads. This issue can be resolved with the design of Candidate B or C due to the addition of cleavable linker.

Future Directions

To discover and investigate global S-nitrosylation in Alzheimer's diseases using multiplexed DBA tag based on quantitative proteomics approach for deepening understanding of molecular mechanisms leading to neurodegenerative diseases.

Emerging evidence have suggested that the occurrence of aberrant S-nitrosylation of proteins could lead to protein misfolding, mitochondrial fission, synaptic damage, or apoptosis, thus contributing to the pathogenesis of Alzheimer's disease (AD) (92-94). For instance, S-nitrosylated NMDAR and caspases induced by low levels of NO can exert neuroprotective effects via suppression of neuronal cell death, whereas S-nitrosylation of Drp1 and Cdk5 lead to excessive mitochondrial fission and thus synaptic damage in AD (95, 96). Additionally, since HSP-90 and PDI are molecular chaperones and XIAP is a ubiquitin E3 ligase, S-nitrosylation of these proteins may contribute to protein misfolding and aggregation in aging neurons (97, 98). Moreover, many S-nitrosylated proteins detected in AD brains are involved in energy metabolism pathway, such as ALDOC and GAPDH (99-101). Hence, this redox reaction could lead to impaired metabolism, contributing to AD pathology. Future studies include the use of shotgun proteomics approach assisted by isoBOP-ABPP to identify and quantify global S-nitrosylation and eventually find a

panel of potential candidate biomarkers in AD. We hypothesize that additional biomarkers may exist in CSF of AD patients and can be used for diagnosis and multiplex quantification via isobaric tagging will facilitate and accelerate our search for such novel biomarkers. In addition, patients from multiple AD stages may have different signature biomarkers for monitoring disease progression.

We will establish SNO analysis platform based on isoBOP-ABPP using cell samples first to validate its feasibility. Further investigation of global S-nitrosylation will be conducted on Alzheimer's diseases mouse model and CSF samples collected from various groups (healthy control, preclinical individuals, patients with mild cognitive impairment (MCI) and AD), and then apply bioinformatics tools to identify differentially expressed putative biomarkers and for AD stages classification.

SNO analysis in GSNO-treated and non-treated HeLa extracts using isoBOP-ABPP.

Among methods for studying protein S-nitrosylation, the biotin switch technique (BST) or its variants have gained great popularity because of the ease with which it can detect individual S-nitrosylated (SNO) proteins in biological samples (102). BST was brought up by Jaffrey et al. in 2001 and consists of three principal steps: (i) blocking of free cysteine thiols by S-methylthiolation with methylmethane thiosulfonate (MMTS; a reactive thiosulfonate); (ii) conversion of SNOs to thiols with a mild reductant ascorbate; and (iii) in situ labeling by S-biotinylation of the nascent thiols with biotin–HPDP, a reactive mixed disulfide of biotin. The degree of biotinylation (and thus S-nitrosylation) is determined by either anti-biotin immunoblotting or streptavidin pulldown followed by immunoblotting for the protein(s) of interest (61). Our DBA tags constitute another variant of BST for studying S-nitrosylation. Instead of using biotin-HPDP, IPM and DBA tag will

be used for chemical labeling. IsoBOP-ABPP enabled MS proteomics will thus achieve high-throughput SNO analysis.

The feasibility of DBA tag-based BST will be tested first using HeLa cell lines. S-Nitrosoglutathione (GSNO) is an endogenous SNO and is a source of bioavailable NO. Cell lysate treated with GSNO will undergo transnitrosylation (**Fig. 10**), exposing the SNO hyper-reactive cysteine residue and potential substrates of SNO in vivo. Therefore, GSNO stimulated cell lysates can provide a positive control of SNO samples (103). Two channels of DBA tags will label GSNO-treated and un-treated HeLa cell lysates respectively and the following steps will be performed according to the isoBOP-ABPP framework. DBA tag-based BST allows us to discover proteome-wide SNO targets and consensus motifs. Mapping the identified SNO proteins to the HeLa cell proteome based on normalized spectral abundance factors (NSAF) will reveal the SNO targets dynamic range relative to the whole proteome. In addition, differential groups defined by RGSNO:control ratios correspond to distinct reactivity of Cys sites to GSNO. These results can verify feasibility and validity of DBA tag-based BST enabled by isoBOP-ABPP, while will provide a blueprint and reference for subsequent global SNO analysis with rodent model and human specimen.

SNO analysis in Alzheimer's diseases mouse model using isoBOP-ABPP.

Two key molecules tau and amyloid-β (Aβ) have been extensively implicated in AD, forming fibrillar aggregates via oligomeric intermediates (104,105). Aβ is a small peptide that is derived from the larger amyloid precursor protein (APP). Aggregates of Aβ form histological lesions known as amyloid plaques, which, when surrounded by neurites filled with fibrillary tau, give rise to what are referred to as neuritic plaques (106, 107). Various transgenic mouse models, based on the overexpression of mutant forms of human APP and recapitulating the AD phenotype

are currently used to investigate mechanisms underlying disease pathology (108). Here, the APP/PS-1 transgenic mouse model, is selected for demonstration of the isoBOP-ABPP method as several SNO-modified proteins have previously been reported in the brain and synaptosomes from AD subjects (109, 110).

APP/PS-1 (hereafter referred to as AD) and wild type (WT) mice (N = 3 for each genotype) will be euthanized using CO_2 before brain tissues is harvested and homogenized in an ice-cold phosphate buffer saline (PBS) solution containing 8 M urea. For SNO analysis, 1 mg brain proteins from individual WT or AD mice are treated with IAA to block free sulfhydryl groups. After tryptic digestion, ascorbate is added to the peptide mixture to selectively reduce SNO. Reduced peptides containing nascent sulfhydryl groups are labeled by IPM. Six channels of DBA tags are used to conjugate IPM-modified peptides in WT and AD samples, respectively. The six samples are combined to a single mixture and subject to streptavidin resin for enrichment. Enriched and tagged peptide samples are eluted from the resin and fractionated by SCX and analyzed using nanoLC-MS/MS. Protein level ratios, SNO site occupancy in WT or AD, and SNO ratios between WT and AD can be calculated by using corresponding reporter ion intensities. The results can be examined by comparison with previous literatures and instruct ensuing application on human samples.

Global SNO analysis in CSF samples from human patients using isoBOP-ABPP.

At present, a definitive diagnosis of AD can only be obtained at autopsy via the histological quantification of two AD hallmarks: brain amyloid plaques and intraneuronal neurofibrillary tangles (111). Thus, there is an urgent need for the discovery of new biomarkers that will aid in identifying the disease at its most early clinical stages as well as preclinically. Cerebrospinal fluid (CSF), given its contiguity with the brain interstitial space, represents the most direct means to study the biochemical changes occurring in the central nervous system (CNS). People have found

certain protein changes directly associated with the progression of AD, for instance, levels of CSF amyloid-β42 (Aβ42) are reduced in AD, even in early and preclinical stages of the disease, and levels of CSF total tau and phosphorylated tau (p-tau) increase in AD and accelerate during later disease stages, concomitant with neurofibrillary tangle formation and synapse and neuron loss (112, 113). However, there are limited research focusing on PTM analysis in various AD stages using CSF samples including S-nitrosylation. To investigate the global SNO changes along with the progression of the disease, and to confidently identify potential candidate biomarkers, high-throughput quantitative tools are essential. Therefore, we propose to use isoBOP-ABPP enabled by 12-plex DBA tags to perform quantitative proteomics in CSF samples from different stages of AD.

As shown in **Fig. 11**, four groups (control, preclinical, MCI and AD) of CSF samples are collected corresponding to different stages of disease progression. Each of which has three biological replicates making up twelve samples in total. Identical proteolysis and BST procedure is applied for all samples, and 12-plex DBA reagents are used for chemical labeling of 12 samples respectively, after which they are combined together to a single mixture and performed downstream SNO enrichment and LC-MS/MS analysis. Accurate quantification can be achieved by comparing relative intensities of the zoom-in low mass region for 12 reporter ions, and differentially regulated SNO proteins or sites will be identified for further investigation.

The development of 12-plex DBA reagents will provide unprecedented throughput and capacity for multiplexed PTM quantitative proteomics measurements that will facilitate AD biomarker discovery and deepen understanding of SNO in AD. We expect smooth development of DBA-based BST method and its application to AD animal model. Moreover, certain identified SNO proteins or sites that are differentially regulated in different CSF sample groups can be

regarded as potential biomarkers. For novel biomarkers we will validate them by the SRM targeted proteomics approach.

To expand the scope of application of DBA tag and establish a pan-PTM high-throughput quantitative proteomics platform for various PTM study and crosstalk investigation.

With the advent of isoTOP-ABPP technique, many novel discoveries were made in PTM research (114-116). The broad utility of isoTOP-ABPP stems from the exquisite design of chemical labeling strategy. It separates the target-site-oriented warhead from functionalities possessing more general purposes such as enrichment and isotope incorporation, thus avoiding redesigning specific chemical probes case-by-case. Click chemistry bridges alkyne and azide in an efficient and robust manner both *in vitro* and *in vivo*, making classical TEV tag widely used in various applications (88, 117-119). Similarly, our DBA tag and isoBOP-ABPP pipeline are expected to inherit merits of TEV tag as well as isoTOP-ABPP platform, with distinctive feature of larger sample processing volume. Moreover, more available channels of tags enable quantifying multiple PTMs simultaneously, allowing investigation of PTM crosstalk and interactions.

We will take a step forward on cysteine oxidative modification analysis in addition to S-nitrosylation first using DBA tag and isoBOP-ABPP platform. Furthermore, broader application can be explored with other PTMs including carbonylation and citrullination via chemical proteomics. To this end, bio-orthogonally probing and fishing multiple PTMs in a one-pot manner and quantitatively profiling their interactions and combinatorial effects could be envisioned.

Application of DBA tag to the analysis of oxidative post-translational modification (OxiPTM).

The prerequisite of using DBA tag and isoBOP-ABPP platform for PTM analysis is the prior selective tagging of modification sites with an alkyne handle. Since DBA tag only contains azide

group for conjugation enabled by click chemistry, the selectivity of PTM probing is achieved by precursor alkyne modified molecules. In addition to S-nitrosylation, the thiol group of the cysteine side chain is subject to a variety of oxiPTMs (**Fig. 12**). Each oxidation state of cysteine exhibits different reactivity which, in turn, may determine their functions. These differences in reactivity may be subtle, presenting an interesting challenge in developing chemoselective probes to detect these modifications (120-122). Currently, chemical labeling approaches have been reported for S-sulphenation (-SOH) and S-sulphination (-SO2H) utilizing special chemical reagents DYn-2 and DiaAlk with high chemoselectivity (123, 124). These probes can successfully convert oxidative thiol into a terminal alkyne, which provides an anchor for DBA tag and makes application of isoBOP-ABPP feasible.

Peptide ANP will be treated with H_2O_2 and $FeSO_4$ to prepare sulphenated and sulphinated standard peptides for method development (125). DYn-2 and DiaAlk will then be used to probe SOH and SO_2H, respectively. Labeling performance and chemoselectivity will be monitored by MALDI-TOF MS. After optimization of labeling condition, integrating scheme of modified BST and isoBOP-ABPP method, brain tissues from AD mouse model or CSF from different stages of AD patients can be used as complex biological samples, for exploration of broader application of DBA reagents and deepen understanding of OxiPTM in organism.

Application of DBA tag to the analysis of carbonylation and citrullination PTM.

Carbonylation and citrullination are two other PTMs that will be investigated using DBA tag enabled isoBOP-ABPP approach. In addition to OxiPTM on cysteine residue, ROS can also induce peptide cleavage which results in producing electrophilic carbonyl groups, also known as protein carbonylation (126). (**Fig.13a**) Recently, quantitative chemoproteomic approaches to globally and site-specifically profile electrophilic carbonyl products formed from peptide

backbone cleavages in human proteomes were reported (127, 128). In these studies, researchers employed a nucleophilic probe, alkynyl hydrazine (HZyne), to selectively label and stabilize carbonylated proteins in cell lysates. Then modified version of isoTOP-ABPP platforms were applied for quantitative analysis. Therefore, we propose to conduct similar studies, using DBA tag to label HZyne converted carbonylated proteins/peptides and to perform quantitative comparison based on isoBOP-ABPP.

Citrullination is an arginine modification that is catalyzed by a group of hydrolases called protein arginine deiminases (PADs), in which the positively charged guanidinium is hydrolyzed to the neutral urea (**Fig.13b**), which alters the charge and hydrogen bonding potential of this residue and impact important biological processes. Thompson et al. previously developed citrulline-specific probe based on the chemoselective reaction that occurs between glyoxals and either citrulline or arginine under acidic or basic conditions, respectively (129). The initial tag is a rhodamine-derivative and used for visualization of the citrullination in both purified proteins and proteins present in complex mixtures such as serum. Later, biotin-derivative was developed to support chemical proteomics analysis (130, 131). Enlightened by their work, we propose to design and synthesize an alkyne containing molecules which also possess phenylglyoxal moiety. This reagent is expected to introduce a terminal alkyne onto citrullinated sites through selective reaction, which offers opportunity for downstream incorporation of DBA tag and isoBOP-ABPP enabled quantitative proteomics.

Multiplexed DBA tag enabled PTM crosstalk investigation.

Typically, PTMs are studied in isolation to determine possible regulation imparted by an individual PTM (132). However, PTM profiling studies have shown that numerous proteins can be modified by multiple PTMs, a finding that suggests the possibility of regulatory crosstalk

between unique modification events. For example, the phosphorylation of specific cell-cycle regulators to create phosphodegrons that are recognized by ubiquitin ligases to direct ubiquitylation and subsequent degradation is a well-characterized example of PTM cross-talk (133). Despite classic examples, the breadth of PTM crosstalk, and whether crosstalk is a widely used mechanism of regulation, remains to be determined. Previously, serial PTM enrichment approaches were reported to catalog multiple modification events in a single sample using MS-based approaches (6, 134). More recently, simultaneous affinity enrichment of two PTMs for quantification and site localization provided a quick and cost-effective alternative (135). In light of the effectiveness of biorthogonal reaction and compatibility of isoBOP-ABPP platform, we propose to use DBA tag enabled isoBOP-ABPP to simultaneously enrich multiple PTMs and facilitate deciphering PTM crosstalk.

Nitric oxide (NO) is generated, in most cell types, from arginine and O_2 by NO synthases (NOSs) and could possibly modify cysteine residue to afford S-nitrosylation. In the meantime, the other product in this pathway is the citrulline originated from arginine (136). Despite close connection, the correlation and interaction between S-nitrosylation and citrullination has rarely been studied. We will have specific chemical probes targeting SNO and citrullination sites in hand, both of which contain the alkyne group. When treating proteome with both probes, respective labeling of SNO and citrullination is envisioned. Sequential click chemistry will introduce DBA tag onto the derivatized PTMs sites which makes it feasible to simultaneously enrich both SNO and citrullination via the isoBOP-ABPP platform (**Fig. 14**). We will use CSF samples from AD patients to perform the proof-of-principle experiments, as both S-nitrosylation and citrullination have been reported to play important roles in neurodegenerative diseases but few studies investigate their relationship and connection. Once the co-enrichment of SNO and citrullination

proves to be success, this design and framework can be widely used for investigation of multiple PTMs at a time and enhance our understanding of the interaction and crosstalk between different PTMs in various biological systems.

We expect to establish a pan-PTM high-throughput quantitative proteomics platform based on the DBA tag. Application scope can be extended but not limited to S-sulphenation, S-sulphination, carbonylation and citrullination. One potential concern with the investigation of crosstalk between S-nitrosylation and citrullination is the difficulty of bioinformatic analysis. The lack of specific search engine limited bioinformatic analysis to be conducted using existing software such as MaxQuant and PEAKS. The default configuration of these search engine may not be appropriate or perform the best analysis in terms of PTM crosstalk. Therefore, custom-designed bioinformatic tool for specifically analyzing multiple DBA tag labeled peptides need to be developed to generate useful biological insights.

Conclusions:

This project uniquely integrates advances of biorthogonal reaction empowered chemical proteomics and DiLeu tag-based multiplex quantitative proteomics to enable a novel chemical tool, the cDBA tag, to be developed. This invention will build a new analytical platform, isoBOP-ABPP, allowing for high-throughput quantitative analysis of PTMs and facilitating more comprehensive discovery and validation of SNO biomarkers in AD CSF. In addition, the versatility of the isoBOP-ABPP platform could also support Pan-PTM and PTM crosstalk studies. These new tools will enable us to gain deeper understanding of PTMs and the advances in technology and new insights will have broad impact on discovery of biomarkers in various diseases.

References:

1. Collins, F. S., Lander, E. S., Rogers, J., Waterston, R. H., and Int Human Genome Sequencing, C. (2004) Finishing the euchromatic sequence of the human genome. Nature 431, 931-945

2. Jensen, O. N. (2004) Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. Current Opinion in Chemical Biology 8, 33-41

3. Chaiyawat, P., Netsirisawan, P., Svasti, J., and Champattanachai, V. (2014) Aberrant O-GIcNAcylated proteins: new perspectives in breast and colorectal cancer. Frontiers in Endocrinology 5

4. Morino, K., Petersen, K. F., Dufour, S., Befroy, D., Frattini, J., Shatzkes, N., Neschen, S., White, M. F., Bilz, S., Sono, S., Pypaert, M., and Shulman, G. I. (2005) Reduced mitochondrial density and increased IRS-1 serine phosphorylation in muscle of insulin-resistant offspring of type 2 diabetic parents. Journal of Clinical Investigation 115, 3587-3593

5. Martin, L., Latypova, X., and Terro, F. (2011) Post-translational modifications of tau protein: Implications for Alzheimer's disease. Neurochemistry International 58, 458-471

6. Swaney, D. L., Beltrao, P., Starita, L., Guo, A., Rush, J., Fields, S., Krogan, N. J., and Villen, J. (2013) Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. Nature Methods 10, 676-+

7. Yao, Q., Li, H., Liu, B. Q., Huang, X. Y., and Guo, L. (2011) SUMOylation-regulated Protein Phosphorylation, Evidence from Quantitative Phosphoproteomics Analyses. Journal of Biological Chemistry 286, 27342-27349

8. Lu, Z. K., Cheng, Z. Y., Zhao, Y. M., and Volchenboum, S. L. (2011) Bioinformatic Analysis and Post-Translational Modification Crosstalk Prediction of Lysine Acetylation. Plos One 6

9. Bulaj, G., Kortemme, T., and Goldenberg, D. P. (1998) Ionization-reactivity relationships for cysteine thiols in polypeptides. Biochemistry 37, 8965-8972

10. Reddie, K. G., and Carroll, K. S. (2008) Expanding the functional diversity of proteins through cysteine oxidation. Current Opinion in Chemical Biology 12, 746-754

11. Nakamura, T., Tu, S. C., Akhtar, M. W., Sunico, C. R., Okamoto, S. I., and Lipton, S. A. (2013) Aberrant Protein S-Nitrosylation in Neurodegenerative Diseases. Neuron 78, 596-614

12. Nakamura, T., Prikhodko, O. A., Pirie, E., Nagar, S., Akhtar, M. W., Oh, C. K., McKercher, S. R., Ambasudhan, R., Okamoto, S., and Lipton, S. A. (2015) Aberrant protein S-nitrosylation contributes to the pathophysiology of neurodegenerative diseases. Neurobiology of Disease 84, 99-108

13. Castro, R., Pinzón, H. S., and Alvis-Guzman, N. (2015) A systematic review of observational studies on oxidative/nitrosative stress involvement in dengue pathogenesis. Colomb Med (Cali) 46, 135-143

14. Filipović, D., Todorović, N., Bernardi, R. E., and Gass, P. (2017) Oxidative and nitrosative stress pathways in the brain of socially isolated adult male rats demonstrating depressive- and anxiety-like symptoms. Brain Struct Funct 222, 1-20

15. Maes, M., Stevens, W. J., Declerck, L. S., Bridts, C. H., Peeters, D., Schotte, C., and Cosyns, P. (1993) Significantly increased expression of T-cell activation markers (interleukin-2 and HLA-DR) in depression: further evidence for an inflammatory process during that illness. Prog Neuropsychopharmacol Biol Psychiatry 17, 241-255

16. Marletta, M. A., Hurshman, A. R., and Rusche, K. M. (1998) Catalysis by nitric oxide synthase. Current Opinion in Chemical Biology 2, 656-663

17. Heinrich, T. A., da Silva, R. S., Miranda, K. M., Switzer, C. H., Wink, D. A., and Fukuto, J. M. (2013) Biological nitric oxide signalling: chemistry and terminology. British Journal of Pharmacology 169, 1417-1429

18. Hess, D. T., Matsumoto, A., Kim, S. O., Marshall, H. E., and Stamler, J. S. (2005) Protein S-nitrosylation: Purview and parameters. Nature Reviews Molecular Cell Biology 6, 150-166

19. Sayre, L. M., Perry, G., and Smith, M. A. (2008) Oxidative stress and neurotoxicity. Chemical Research in Toxicology 21, 172-188

20. Chung, K. K. K., Thomas, B., Li, X. J., Pletnikova, O., Troncoso, J. C., Marsh, L., Dawson, V. L., and Dawson, T. M. (2004) S-nitrosylation of Parkin regulates ubiquitination and compromises Parkin's protective function. Science 304, 1328-1331

21. Uehara, T., Nakamura, T., Yao, D. D., Shi, Z. Q., Gu, Z. Z., Ma, Y. L., Masliah, E., Nomura, Y., and Lipton, S. A. (2006) S-Nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration. Nature 441, 513-517

22. Shankar, G. M., Bloodgood, B. L., Townsend, M., Walsh, D. M., Selkoe, D. J., and Sabatini, B. L. (2007) Natural oligomers of the Alzheimer amyloid-beta protein induce reversible synapse loss by modulating an NMDA-type glutamate receptor-dependent signaling pathway. Journal of Neuroscience 27, 2866-2875

23. Qu, J., Nakamura, T., Cao, G., Holland, E. A., McKercher, S. R., and Lipton, S. A. (2011) S-Nitrosylation activates Cdk5 and contributes to synaptic spine loss induced by beta-amyloid peptide. Proceedings of the National Academy of Sciences of the United States of America 108, 14330-14335

24. Zhang, P., Yu, P. C., Tsang, A. H. K., Chen, Y., Fu, A. K. Y., Fu, W. Y., Chung, K. K., and Ip, N. Y. (2010) S-Nitrosylation of Cyclin-Dependent Kinase 5 (Cdk5) Regulates Its Kinase Activity and Dendrite Growth During Neuronal Development. Journal of Neuroscience 30, 14366-14370

25. Dixon, S. J., and Stockwell, B. R. (2009) Identifying druggable disease-modifying gene products. Current Opinion in Chemical Biology 13, 549-555

26. Parker, C. G., Galmozzi, A., Wang, Y. J., Correia, B. E., Sasaki, K., Joslyn, C. M., Kim, A. S., Cavallaro, C. L., Lawrence, R. M., Johnson, S. R., Narvaiza, I., Saez, E., and Cravatt, B. F. (2017) Ligand and Target Discovery by Fragment-Based Screening in Human Cells. Cell 168, 527-+

27. Deng, H., Lei, Q., Wu, Y. P., He, Y., and Li, W. M. (2020) Activity-based protein profiling: Recent advances in medicinal chemistry. European Journal of Medicinal Chemistry 191

28. Chen, X., Wong, Y. K., Wang, J. G., Zhang, J. B., Lee, Y. M., Shen, H. M., Lin, Q. S., and Hua, Z. C. (2017) Target identification with quantitative activity based protein profiling (ABPP). Proteomics 17

29. Roberts, A. M., Ward, C. C., and Nomura, D. K. (2017) Activity-based protein profiling for mapping and pharmacologically interrogating proteome-wide ligandable hotspots. Current Opinion in Biotechnology 43, 25-33

30. Moellering, R. E., and Cravatt, B. F. (2012) How Chemoproteomics Can Enable Drug Discovery and Development. Chemistry & Biology 19, 11-22

31. Chakrabarty, S., Kahler, J. P., van de Plassche, M. A. T., Vanhoutte, R., and Verhelst, S. H. L. (2019) Recent Advances in Activity-Based Protein Profiling of Proteases. Activity-Based Protein Profiling 420, 253-281

32. Yang, P. Y., and Liu, K. (2015) Activity-Based Protein Profiling: Recent Advances in Probe Development and Applications. Chembiochem 16, 712-724

33. Weerapana, E., Wang, C., Simon, G. M., Richter, F., Khare, S., Dillon, M. B. D., Bachovchin, D. A., Mowen, K., Baker, D., and Cravatt, B. F. (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. Nature 468, 790-U779

34. Weerapana, E., Wang, C., Simon, G. M., Richter, F., Khare, S., Dillon, M. B. D., Bachovchin, D. A., Mowen, K., Baker, D., and Cravatt, B. F. (2011) Quantitative reactivity profiling predicts functional cysteines in proteomes. Abstracts of Papers of the American Chemical Society 241

35. Sieber, S. A., and Cravatt, B. F. (2006) Analytical platforms for activity-based protein profiling - exploiting the versatility of chemistry for functional proteomics. Chemical Communications, 2311-2319

36. Rauniyar, N., and Yates, J. R. (2014) Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics. Journal of Proteome Research 13, 5293-5309

37. Nunez, E. V., Domont, G. B., and Nogueira, F. C. S. (2017) iTRAQ-Based Shotgun Proteomics Approach for Relative Protein Quantification. In Multiplex Biomarker Techniques: Methods and Applications (Guest, P. C., ed) Vol. 1546 pp. 267-274

38. Li, J. M., Van Vranken, J. G., Vaites, L. P., Schweppe, D. K., Huttlin, E. L., Etienne, C., Nandhikonda, P., Viner, R., Robitaille, A. M., Thompson, A. H., Kuhn, K., Pike, I., Bomgarden, R. D., Rogers, J. C., Gygi, S. P., and Paulo, J. A. (2020) TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. Nature Methods 17, 399-+

39. Xiang, F., Ye, H., Chen, R. B., Fu, Q., and Li, L. J. (2010) N,N-Dimethyl Leucines as Novel Isobaric Tandem Mass Tags for Quantitative Proteomics and Peptidomics. Analytical Chemistry 82, 2817-2825

40. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nature Biotechnology 20, 301-305

41. Stensballe, A., Andersen, S., and Jensen, O. N. (2001) Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. Proteomics 1, 207-222

42. Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jorgensen, T. J. D. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Molecular & Cellular Proteomics 4, 873-886

43. Yang, Z. P., and Hancock, W. S. (2004) Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. Journal of Chromatography A 1053, 79-88

44. Tian, Y. A., Zhou, Y., Elliott, S., Aebersold, R., and Zhang, H. (2007) Solid-phase extraction of N-linked glycopeptides. Nature Protocols 2, 334-339

45. Zhao, Y. M., and Jensen, O. N. (2009) Modification-specific proteomics: Strategies for characterization of post-translational modifications using enrichment techniques. Proteomics 9, 4632-4641

46. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. Analytical and Bioanalytical Chemistry 389, 1017-1031

47. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotechnology 17, 994-999

48. Xiang, F., Ye, H., Chen, R., Fu, Q., and Li, L. (2010) N, N-dimethyl leucines as novel isobaric tandem mass tags for quantitative proteomics and peptidomics. Analytical chemistry 82, 2817-2825

49. Frost, D. C., Greer, T., and Li, L. (2014) High-resolution enabled 12-plex DiLeu isobaric tags for quantitative proteomics. Analytical chemistry 87, 1646-1654

50. Hao, L., Zhong, X., Greer, T., Ye, H., and Li, L. (2015) Relative quantification of amine-containing metabolites using isobaric N, N-dimethyl leucine (DiLeu) reagents via LC-ESI-MS/MS and CE-ESI-MS/MS. Analyst 140, 467-475

51. Hebert, A. S., Merrill, A. E., Bailey, D. J., Still, A. J., Westphall, M. S., Strieter, E. R., Pagliarini, D. J., and Coon, J. J. (2013) Neutron-encoded mass signatures for multiplexed proteome quantification. Nature Methods 10, 332-+

52. Sleno, L. (2012) The use of mass defect in modern mass spectrometry. Journal of Mass Spectrometry 47, 226-236

53. Feng, Y., Chen, B. M., Yu, Q. Y., Zhong, X. F., Frost, D. C., Ikonomidou, C., and Li, L. J. (2019) Isobaric Multiplex Labeling Reagents for Carbonyl-Containing Compound (SUGAR) Tags: A Probe for Quantitative Glycomic Analysis. Analytical Chemistry 91, 3141-3146

54. Feng, Y., Li, M. Y., Lin, Y. Y., Chen, B. M., and Li, L. J. (2019) Multiplex Quantitative Glycomics Enabled by Periodate Oxidation and Triplex Mass Defect Isobaric Multiplex Reagents for Carbonyl-Containing Compound Tags. Analytical Chemistry 91, 11932-11937

55. Wong, J., Chilkoti, A., and Moy, V. T. (1999) Direct force measurements of the streptavidin-biotin interaction. Biomolecular Engineering 16, 45-55

56. Hems, E. S., Wagstaff, B. A., Saalbach, G., and Field, R. A. (2018) CuAAC click chemistry for the enhanced detection of novel alkyne-based natural product toxins. Chemical Communications 54, 12234-12237

57. Meldal, M., and Diness, F. (2020) Recent Fascinating Aspects of the CuAAC Click Reaction. Trends in Chemistry 2, 569-584

58. Chio, T. I., and Bane, S. L. (2020) Click Chemistry Conjugations. Antibody-Drug Conjugates: Methods and Protocols 2078, 83-97

59. Araki, K., Kusano, H., Sasaki, N., Tanaka, R., Hatta, T., Fukui, K., and Natsume, T. (2016) Redox Sensitivities of Global Cellular Cysteine Residues under Reductive and Oxidative Stress. Journal of Proteome Research 15, 2548-2559

60. Qu, Z., Meng, F. J., Bomgarden, R. D., Viner, R. I., Li, J. L., Rogers, J. C., Cheng, J. L., Greenlief, C. M., Cui, J. K., Lubahn, D. B., Sun, G. Y., and Gu, Z. Z. (2014) Proteomic Quantification and Site-Mapping of S-Nitrosylated Proteins Using Isobaric iodoTMT Reagents. Journal of Proteome Research 13, 3200-3211

61. Jaffrey, S. R., Erdjument-Bromage, H., Ferris, C. D., Tempst, P., and Snyder, S. H. (2001) Protein S-nitrosylation: a physiological signal for neuronal nitric oxide. Nature Cell Biology 3, 193-197

62. Chen, Y. J., Ching, W. C., and Lin, Y. P. (2013) Methods for detection and characterization of protein S-nitrosylation. Methods 62, 138-150

63. Lopez-Sanchez, L. M., Muntane, J., de la Mata, M., and Rodriguez-Ariza, A. (2009) Unraveling the S-nitrosoproteome: Tools and strategies. Proteomics 9, 808-818

64. Giron, P., Dayon, L., and Sanchez, J. C. (2011) CYSTEINE TAGGING FOR MS-BASED PROTEOMICS. Mass Spectrometry Reviews 30, 366-395

65. Petritis, K., Chaimbault, P., Elfakir, C., and Dreux, M. (2000) Parameter optimization for the analysis of underivatized protein amino acids by liquid chromatography and ionspray tandem mass spectrometry. Journal of Chromatography A 896, 253-263

66. Nie, L. T., Zhu, M. R., Sun, S. N., Zhai, L. H., Wu, Z. X., Qian, L. L., and Tan, M. J. (2016) An optimization of the LC-MS/MS workflow for deep proteome profiling on an Orbitrap Fusion. Analytical Methods 8, 425-434

67. Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Molecular Systems Biology 4

68. Speers, A. E., Adam, G. C., and Cravatt, B. F. (2003) Activity-based protein profiling in vivo using a copper(I)-catalyzed azide-alkyne 3+2 cycloaddition. Journal of the American Chemical Society 125, 4686-4687

69. Ovaa, H., van Swieten, P. F., Kessler, B. M., Leeuwenburgh, M. A., Fiebiger, E., van den Nieuwendijk, A., Galardy, P. J., van der Marel, G. A., Ploegh, H. L., and Overkleeft, H. S. (2003) Chemistry in living cells: Detection of active proteasomes by a two-step labeling strategy. Angewandte Chemie-International Edition 42, 3626-3629

70. Evans, M. J., and Cravatt, B. F. (2006) Mechanism-based profiling of enzyme families. Chemical Reviews 106, 3279-3301

71. Paulick, M. G., and Bogyo, M. (2008) Application of activity-based probes to the study of enzymes involved in cancer progression. Current Opinion in Genetics & Development 18, 97-106

72. Speers, A. E., and Cravatt, B. F. (2005) A tandem orthogonal proteolysis strategy for high-content chemical proteomics. Journal of the American Chemical Society 127, 10018-10019

73. Su, Y., Ge, J. Y., Zhu, B. W., Zheng, Y. G., Zhu, Q., and Yao, S. Q. (2013) Target identification of biologically active small molecules via in situ methods. Current Opinion in Chemical Biology 17, 768-775

74. Holmberg, A., Blomstergren, A., Nord, O., Lukacs, M., Lundeberg, J., and Uhlen, M. (2005) The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. Electrophoresis 26, 501-510

75. Hyre, D. E., Amon, L. M., Penzotti, J. E., Le Trong, I., Stenkamp, R. E., Lybrand, T. P., and Stayton, P. S. (2002) Early mechanistic events in biotin dissociation from streptavidin. Nature Structural Biology 9, 582-585

76. Rudolf, G. C., Heydenreuter, W., and Sieber, S. A. (2013) Chemical proteomics: ligation and cleavage of protein modifications. Current Opinion in Chemical Biology 17, 110-117

77. Szychowski, J., Mahdavi, A., Hodas, J. J. L., Bagert, J. D., Ngo, J. T., Landgraf, P., Dieterich, D. C., Schuman, E. M., and Tirrell, D. A. (2010) Cleavable Biotin Probes for Labeling of Biomolecules via Azide-Alkyne Cycloaddition. Journal of the American Chemical Society 132, 18351-18360

78. Fonovic, M., Verhelst, S. H. L., Sorum, M. T., and Bogyo, M. (2007) Proteomics evaluation of chemically cleavable activity-based probes. Molecular & Cellular Proteomics 6, 1761-1770

79. Geurink, P. P., Florea, B. I., Li, N., Witte, M. D., Verasdonck, J., Kuo, C. L., van der Marel, G. A., and Overkleeft, H. S. (2010) A Cleavable Linker Based on the Levulinoyl Ester for Activity-Based Protein Profiling. Angewandte Chemie-International Edition 49, 6802-6805

80. Yonezawa, H., Nishiyama, Y., Takeo, K., Iwatsubo, T., Tomita, T., Yokoshima, S., and Fukuyama, T. (2014) New photocleavable linker: alpha-Thioacetophenone-type linker. Bioorganic & Medicinal Chemistry Letters 24, 2831-2833

81. Craft, G. E., Chen, A., and Nairn, A. C. (2013) Recent advances in quantitative neuroproteomics. Methods 61, 186-218

82. Fauq, A. H., Kache, R., Khan, M. A., and Vega, I. E. (2006) Synthesis of acid-cleavable light isotope-coded affinity tags (ICAT-L) for potential use in proteomic expression profiling analysis. Bioconjugate Chemistry 17, 248-254

83. Neckers, D. C., Kellogg, R. M., Prins, W. L., and Schoustr.B. (1971) DEVELOPMENTAL PHOTOCHEMISTRY - NORRISH TYPE-II REACTION. Journal of Organic Chemistry 36, 1838-&

84. Black, L. S., Monroe, W. E., Lee, J. C., and Robertson, J. L. (1994) ATRIAL-NATRIURETIC-PEPTIDE. Compendium on Continuing Education for the Practicing Veterinarian 16, 717-&

85. Sun, N. N., Wang, Y., Wang, J., Sun, W. C., Yang, J. B., and Liu, N. (2020) Highly Efficient Peptide-Based Click Chemistry for Proteomic Profiling of Nascent Proteins. Analytical Chemistry 92, 8292-8297

86. Rudolf, G. C., and Sieber, S. A. (2013) Copper-Assisted Click Reactions for Activity-Based Proteomics: Fine-Tuned Ligands and Refined Conditions Extend the Scope of Application. Chembiochem 14, 2447-2455

87. Frost, D. C., Feng, Y., and Li, L. J. (2020) 21-plex DiLeu Isobaric Tags for High-Throughput Quantitative Proteomics. Analytical Chemistry 92, 8228-8234

88. Wang, C., and Cravatt, B. F. (2014) A Chemoproteomic Platform To Quantitatively Map Targets Of Lipid-Derived Electrophiles. Protein Science 23, 191-191

89. Zhou, Y. N., Wynia-Smith, S. L., Couvertier, S. M., Kalous, K. S., Marletta, M. A., Smith, B. C., and Weerapana, E. (2016) Chemoproteomic Strategy to Quantitatively Monitor Transnitrosation Uncovers Functionally Relevant S-Nitrosation Sites on Cathepsin D and HADH2. Cell Chemical Biology 23, 727-737

90. Bateman, L. A., Zaro, B. W., Miller, S. M., and Pratt, M. R. (2018) An Alkyne-Aspirin Chemical Reporter for the Detection of Aspirin-Dependent Protein Modification in Living Cells (vol 135, pg 14568, 2013). Journal of the American Chemical Society 140, 4954-4954

91. Bak, D. W., Gao, J. J., Wang, C., and Weerapana, E. (2018) A Quantitative Chemoproteomic Platform to Monitor Selenocysteine Reactivity within a Complex Proteome. Cell Chemical Biology 25, 1157-+

92. Wijasa, T. S., Sylvester, M., Brocke-Ahmadinejad, N., Schwartz, S., Santarelli, F., Gieselmann, V., Klockgether, T., Brosseron, F., and Heneka, M. T. (2020) Quantitative proteomics of synaptosome S-nitrosylation in Alzheimer's disease. Journal of Neurochemistry 152, 710-726

93. Zhao, Q. F., Yu, J. T., and Tan, L. (2015) S-Nitrosylation in Alzheimer's disease. Molecular Neurobiology 51, 268-280

94. Abrams, A. J., Farooq, A., and Wang, G. F. (2011) S-Nitrosylation of ApoE in Alzheimer's Disease. Biochemistry 50, 3405-3407

95. Nakamura, T., and Lipton, S. A. (2013) Emerging Role of Protein-Protein Transnitrosylation in Cell Signaling Pathways. Antioxidants & Redox Signaling 18, 239-249

96. Qu, J., Nakamura, T., Holland, E. A., McKercher, S. R., and Lipton, S. A. (2012) S-nitrosylation of Cdk5 Potential implications in amyloid-beta-related neurotoxicity in Alzheimer disease. Prion 6, 364-370

97. Retzlaff, M., Stahl, M., Eberl, H. C., Lagleder, S., Beck, J., Kessler, H., and Buchner, J. (2009) Hsp90 is regulated by a switch point in the C-terminal domain. Embo Reports 10, 1147-1153

98. Walker, A. K., Farg, M. A., Bye, C. R., McLean, C. A., Horne, M. K., and Atkin, J. D. (2010) Protein disulphide isomerase protects against protein aggregation and is S-nitrosylated in amyotrophic lateral sclerosis. Brain 133, 105-116

99. Chakravarti, R., Aulak, K. S., Fox, P. L., and Stuehr, D. J. (2010) GAPDH regulates cellular heme insertion into inducible nitric oxide synthase. Proceedings of the National Academy of Sciences of the United States of America 107, 18004-18009

100. Moskovitz, J. (2014) Detection and Localization of Methionine Sulfoxide Residues of Specific Proteins in Brain Tissue. Protein and Peptide Letters 21, 52-55

101. Sun, M. K., and Alkon, D. L. (2002) Carbonic anhydrase gating of attention: memory therapy and enhancement. Trends in Pharmacological Sciences 23, 83-89

102. Forrester, M. T., Foster, M. W., and Stamler, J. S. (2007) Assessment and application of the biotin switch technique for examining protein S-nitrosylation under conditions of pharmacologically induced oxidative stress. Journal of Biological Chemistry 282, 13977-13983

103. Mnatsakanyan, R., Markoutsa, S., Walbrunn, K., Roos, A., Verhelst, S. H. L., and Zahedi, R. P. (2019) Proteome-wide detection of S-nitrosylation targets and motifs using bioorthogonal cleavable-linker-based enrichment and switch technique. Nature Communications 10

104. Polanco, J. C., Li, C. Z., Bodea, L. G., Martinez-Marmol, R., Meunier, F. A., and Gotz, J. (2018) Amyloid-beta and tau complexity - towards improved biomarkers and targeted therapies. Nature Reviews Neurology 14, 22-39

105. Goedert, M. (2015) Alzheimer's and Parkinson's diseases: The prion concept in relation to assembled A beta, tau, and alpha-synuclein. Science 349

106. Glenner, G. G., and Wong, C. W. (2012) ALZHEIMER'S DISEASE: INITIAL REPORT OF THE PURIFICATION AND CHARACTERIZATION OF A NOVEL CEREBROVASCULAR AMYLOID PROTEIN (Reprinted from BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol 120, pg 885-890, 1984). Biochemical and Biophysical Research Communications 425, 534-539

107. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L., and Beyreuther, K. (1985) AMYLOID PLAQUE CORE PROTEIN IN ALZHEIMER-DISEASE AND DOWN SYNDROME. Proceedings of the National Academy of Sciences of the United States of America 82, 4245-4249

108. Zareba-Koziol, M., Szwajda, A., Dadlez, M., Wyslouch-Cieszynska, A., and Lalowski, M. (2014) Global Analysis of S-nitrosylation Sites in the Wild Type (APP) Transgenic Mouse Brain-Clues for Synaptic Pathology. Molecular & Cellular Proteomics 13, 2288-2305

109. Gu, L. Q., and Robinson, R. A. S. (2016) High-throughput endogenous measurement of S-nitrosylation in Alzheimer's disease using oxidized cysteine-selective cPILOT. Analyst 141, 3904-3915

110. Gu, L. Q., and Robinson, R. A. S. (2016) Proteomic approaches to quantify cysteine reversible modifications in aging and neurodegenerative diseases. Proteomics Clinical Applications 10, 1159-1177

111. Hyman, B. T., and Trojanowski, J. Q. (1997) Editorial on consensus recommendations for the postmortem diagnosis of Alzheimer disease from the National Institute on Aging and the Reagan Institute working group on diagnostic criteria for the neuropathological assessment of Alzheimer disease. Journal of Neuropathology and Experimental Neurology 56, 1095-1097

112. Fagan, A. M., Mintun, M. A., Shah, A. R., Aldea, P., Roe, C. M., Mach, R. H., Marcus, D., Morris, J. C., and Holtzman, D. M. (2009) Cerebrospinal fluid tau and ptau(181) increase with cortical amyloid deposition in cognitively normal individuals: Implications for future clinical trials of Alzheimer's disease. Embo Molecular Medicine 1, 371-380

113. Fagan, A. M., Head, D., Shah, A. R., Marcus, D., Mintun, M., Morris, J. C., and Holtzman, D. M. (2009) Decreased Cerebrospinal Fluid A beta(42) Correlates with Brain Atrophy in Cognitively Normal Elderly. Annals of Neurology 65, 176-183

114. Backus, K. M., Correia, B. E., Lum, K. M., Forli, S., Horning, B. D., Gonzalez-Paez, G. E., Chatterjee, S., Lanning, B. R., Teijaro, J. R., Olson, A. J., Wolan, D. W., and Cravatt, B. F. (2016) Proteome-wide covalent ligand discovery in native biological systems. Nature 534, 570-+

115. Bar-Peled, L., Kemper, E. K., Suciu, R. M., Vinogradova, E. V., Backus, K. M., Horning, B. D., Paul, T. A., Ichu, T. A., Svensson, R. U., Olucha, J., Chang, M. W., Kok, B. P., Zhu, Z., Ihle, N. T., Dix, M. M., Jiang, P., Hayward, M. M., Saez, E., Shaw, R. J., and Cravatt, B. F. (2017) Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. Cell 171, 696-+

116. van Esbroeck, A. C. M., Janssen, A. P. A., Cognetta, A. B., Ogasawara, D., Shpak, G., van der Kroeg, M., Kantae, V., Baggelaar, M. P., de Vrij, F. M. S., Deng, H., Allara, M., Fezza, F., Lin, Z., van der Wel, T., Soethoudt, M., Mock, E. D., den Dulk, H., Baak, I. L., Florea, B. I., Hendriks, G., de Petrocellis, L., Overkleeft, H. S., Hankemeier, T., De Zeeuw, C. I., Di Marzo, V.,

Maccarrone, M., Cravatt, B. F., Kushner, S. A., and van der Stelt, M. (2017) Activity-based protein profiling reveals off-target proteins of the FAAH inhibitor BIA 10-2474. Science 356, 1084-1087

117. Deng, X., Weerapana, E., Ulanovskaya, O., Sun, F., Liang, H. H., Ji, Q. J., Ye, Y., Fu, Y., Zhou, L., Li, J. X., Zhang, H. Y., Wang, C., Alvarez, S., Hicks, L. M., Lan, L. F., Wu, M., Cravatt, B. F., and He, C. (2013) Proteome-wide Quantification and Characterization of Oxidation-Sensitive Cysteines in Pathogenic Bacteria. Cell Host & Microbe 13, 358-370

118. Medina-Cleghorn, D., Bateman, L. A., Ford, B., Heslin, A., Fisher, K. J., Dalvie, E. D., and Nomura, D. K. (2015) Mapping Proteome-Wide Targets of Environmental Chemicals Using Reactivity-Based Chemoproteomic Platforms. Chemistry & Biology 22, 1394-1405

119. Huang, Z., Ogasawara, D., Seneviratne, U. I., Cognetta, A. B., Ende, C. W. A., Nason, D. M., Lapham, K., Litchfield, J., Johnson, D. S., and Cravatt, B. F. (2019) Global Portrait of Protein Targets of Metabolites of the Neurotoxic Compound BIA 10-2474. Acs Chemical Biology 14, 192-197

120. Alcock, L. J., Perkins, M. V., and Chalker, J. M. (2018) Chemical methods for mapping cysteine oxidation. Chemical Society Reviews 47, 231-268

121. Lo Conte, M., and Carroll, K. S. (2013) The Redox Biochemistry of Protein Sulfenylation and Sulfinylation. Journal of Biological Chemistry 288, 26480-26488

122. Miseta, A., and Csutora, P. (2000) Relationship between the occurrence of cysteine in proteins and the complexity of organisms. Molecular Biology and Evolution 17, 1232-1239

123. Yang, J., Gupta, V., Tallman, K. A., Porter, N. A., Carroll, K. S., and Liebler, D. C. (2015) Global, in situ, site-specific analysis of protein S-sulfenylation. Nature Protocols 10, 1022-1037

124. Akter, S., Fu, L., Jung, Y., Lo Conte, M., Lawson, J. R., Lowther, W. T., Sun, R., Liu, K. K., Yang, J., and Carroll, K. S. (2018) Chemical proteomics reveals new targets of cysteine sulfinic acid reductase. Nature Chemical Biology 14, 995-+

125. Shetty, V., Spellman, D. S., and Neubert, T. A. (2007) Characterization by tandem mass spectrometry of stable cysteine sulfenic acid in a cysteine switch peptide of matrix metalloproteinases. Journal of the American Society for Mass Spectrometry 18, 1544-1551

126. Fedorova, M., Bollineni, R. C., and Hoffmann, R. (2014) PROTEIN CARBONYLATION AS A MAJOR HALLMARK OF OXIDATIVE DAMAGE: UPDATE OF ANALYTICAL STRATEGIES. Mass Spectrometry Reviews 33, 79-97

127. Tian, C. P., Liu, K. K., Sun, R., Fu, L., and Yang, J. (2018) Chemoproteomics Reveals Unexpected Lysine/Arginine-Specific Cleavage of Peptide Chains as a Potential Protein Degradation Machinery. Analytical Chemistry 90, 794-800

128. Chen, Y., Cong, Y., Quan, B. Y., Lan, T., Chu, X. Y., Ye, Z., Hou, X. M., and Wang, C. (2017) Chemoproteomic profiling of targets of lipid-derived electrophiles by bioorthogonal aminooxy probe. Redox Biology 12, 712-718

129. Bicker, K. L., Subramanian, V., Chumanevich, A. A., Hofseth, L. J., and Thompson, P. R. (2012) Seeing Citrulline: Development of a Phenylglyoxal-Based Probe To Visualize Protein Citrullination. Journal of the American Chemical Society 134, 17015-17018

130. Lewallen, D. M., Bicker, K. L., Subramanian, V., Clancy, K. W., Slade, D. J., Martell, J., Dreyton, C. J., Sokolove, J., Weerapana, E., and Thompson, P. R. (2015) Chemical Proteomic Platform To Identify Citrullinated Proteins. Acs Chemical Biology 10, 2520-2528

131. Nemmara, V. V., and Thompson, P. R. (2019) Development of Activity-Based Proteomic Probes for Protein Citrullination. Activity-Based Protein Profiling 420, 233-251

132. Altelaar, A. F. M., Munoz, J., and Heck, A. J. R. (2013) Next-generation proteomics: towards an integrative view of proteome dynamics. Nature Reviews Genetics 14, 35-48

133. Mocciaro, A., and Rape, M. (2012) Emerging regulatory mechanisms in ubiquitin-dependent cell cycle control. Journal of Cell Science 125, 255-263

134. Mertins, P., Qiao, J. W., Patel, J., Udeshi, N. D., Clauser, K. R., Mani, D. R., Burgess, M. W., Gillette, M. A., Jaffe, J. D., and Carr, S. A. (2013) Integrated proteomic analysis of post-translational modifications by serial enrichment. Nature Methods 10, 634-+

135. Xie, X. S., Shah, S. M., Holtz, A. A., Rose, A. O., Basisty, N. H., and Schilling, B. R. (2020) Simultaneous Affinity Enrichment of Two Post-Translational Modifications for Quantification and Site Localization. Jove-Journal of Visualized Experiments

136. Lima, B., Forrester, M. T., Hess, D. T., and Stamler, J. S. (2010) S-Nitrosylation in Cardiovascular Signaling. Circulation Research 106, 633-646

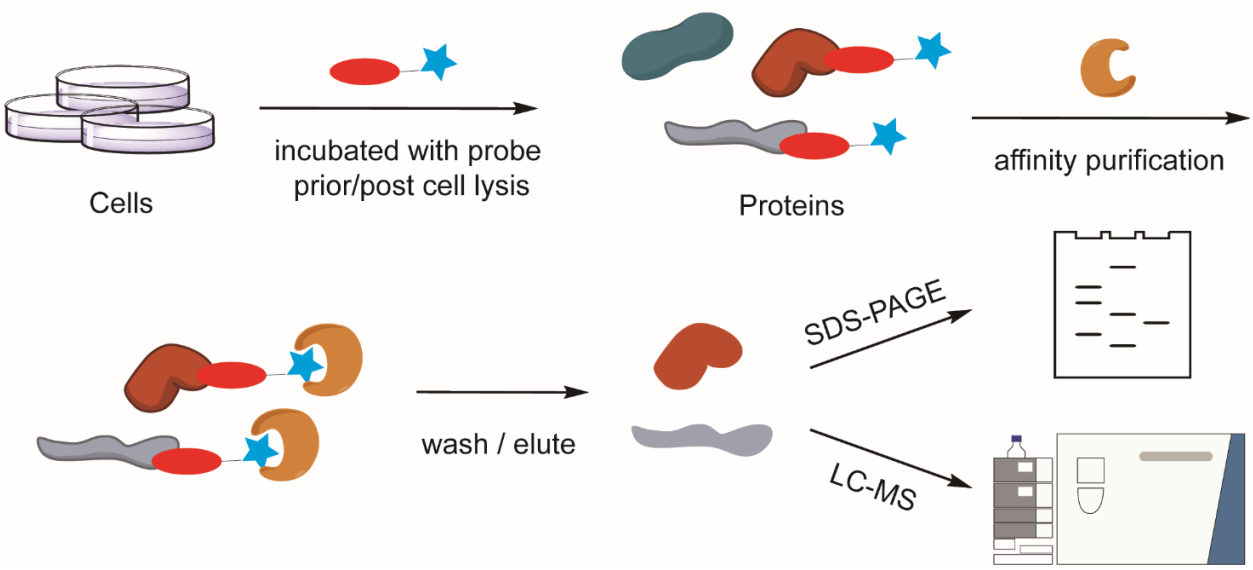


Figure 1. Workflow of ABPP for target identification.

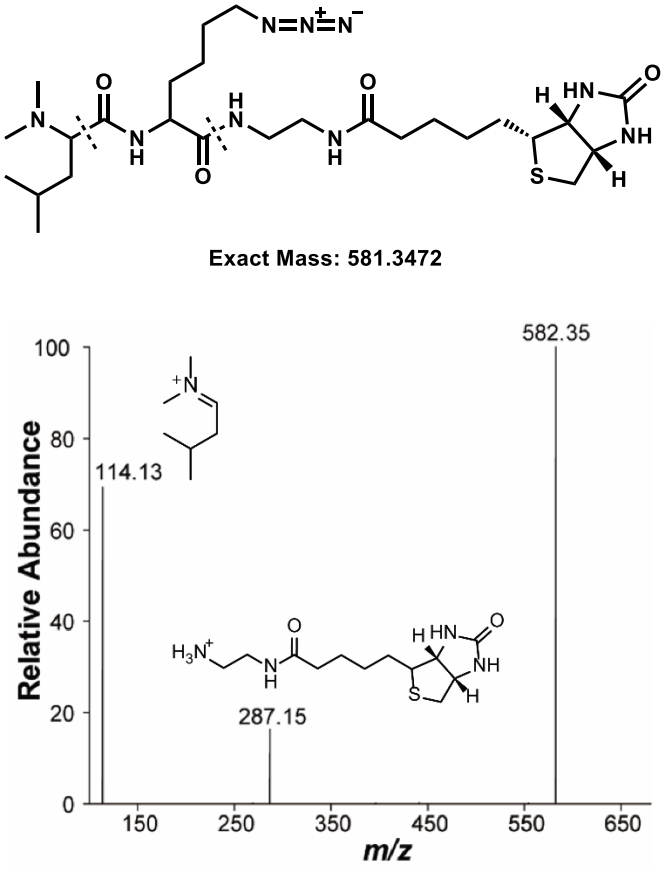


Figure 2. Structure of prototype DBA tag (top) and its MS/MS spectrum (bottom).

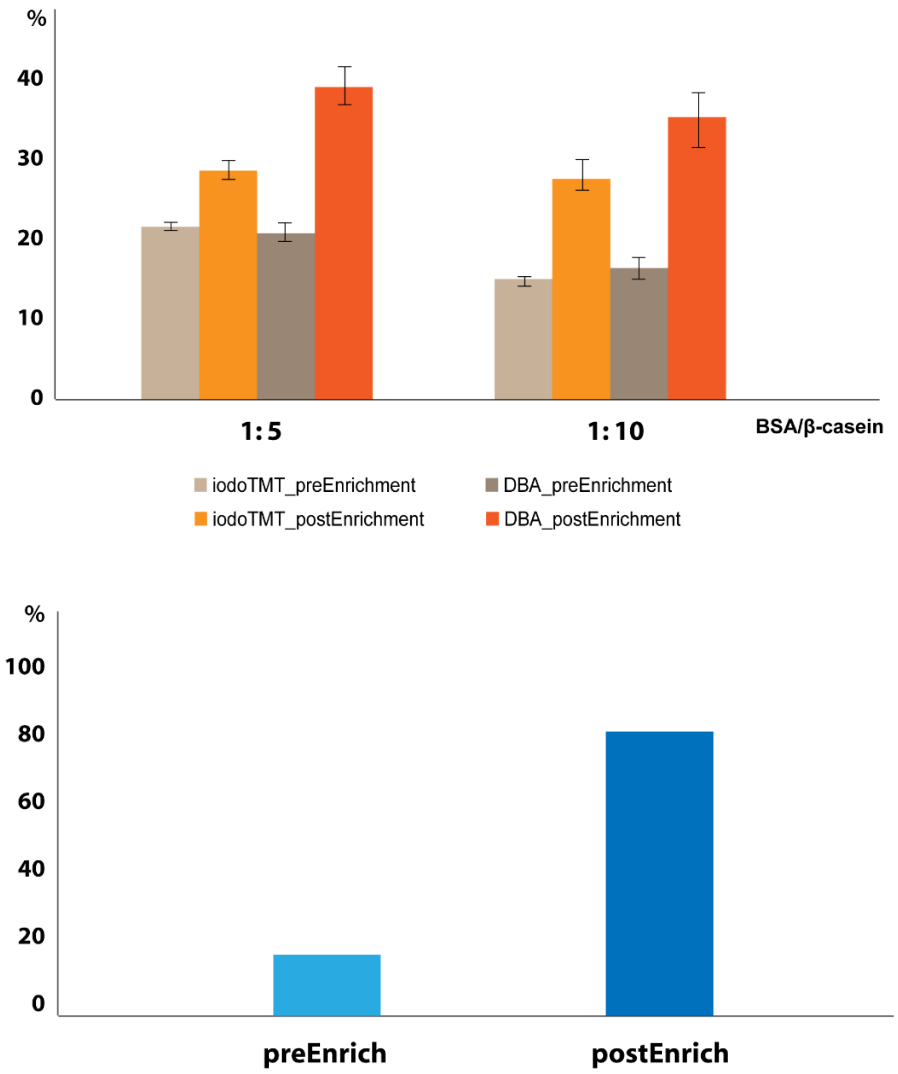


Figure 3. Enrichment efficiency comparison between DBA tag and iodoTMT (top) and total cysteine analysis in HeLa cell lines (bottom). Tryptic digested BSA was spiked into tryptic digested β-casein with a ratio at 1:5 or 1:10. The y-axis represents the percentage of cysteine-containing peptides / total peptides.

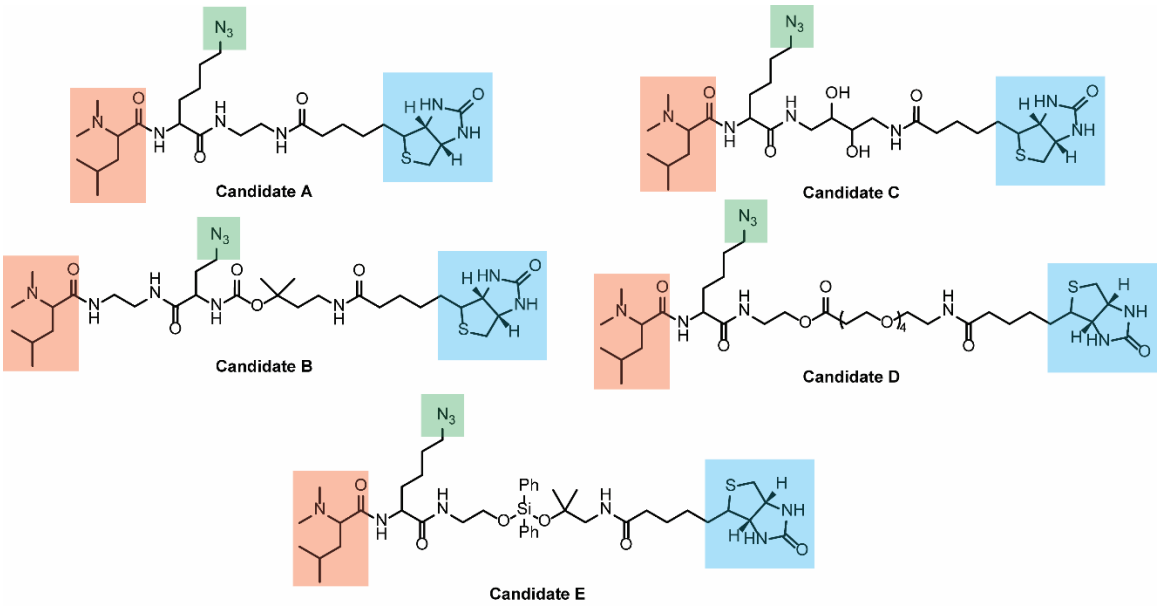


Figure 4. Chemical structures of DBA tag candidates. (red shade: DiLeu; blue shade: Biotin; green shade: Azide)

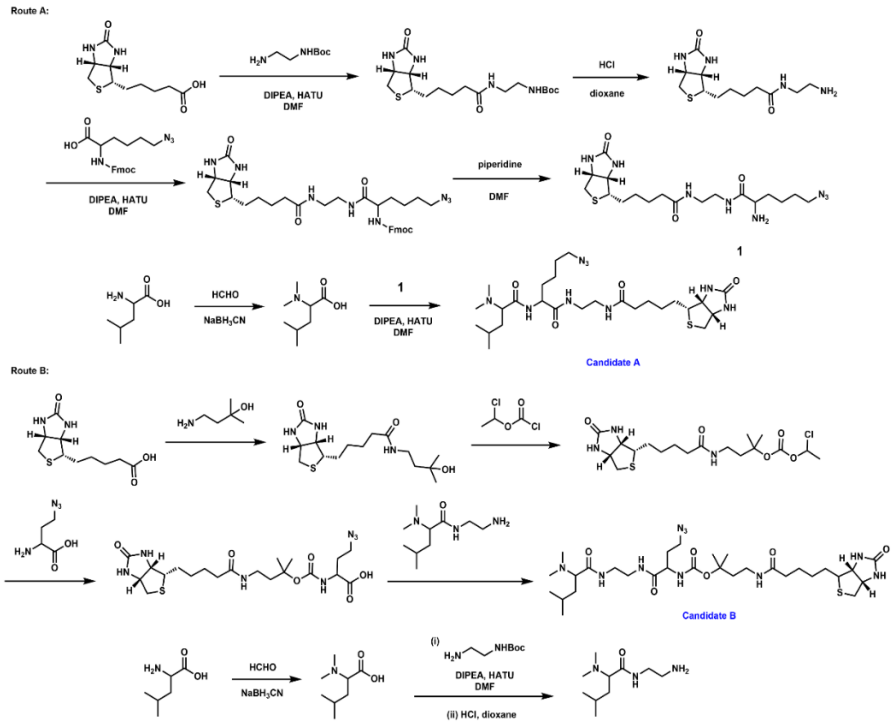


Figure 5a. Synthetic routes of DBA tag candidate A and B (candidate A: non-cleavable protoDBA tag; candidate B: DBA tag with an acid cleavable linker).

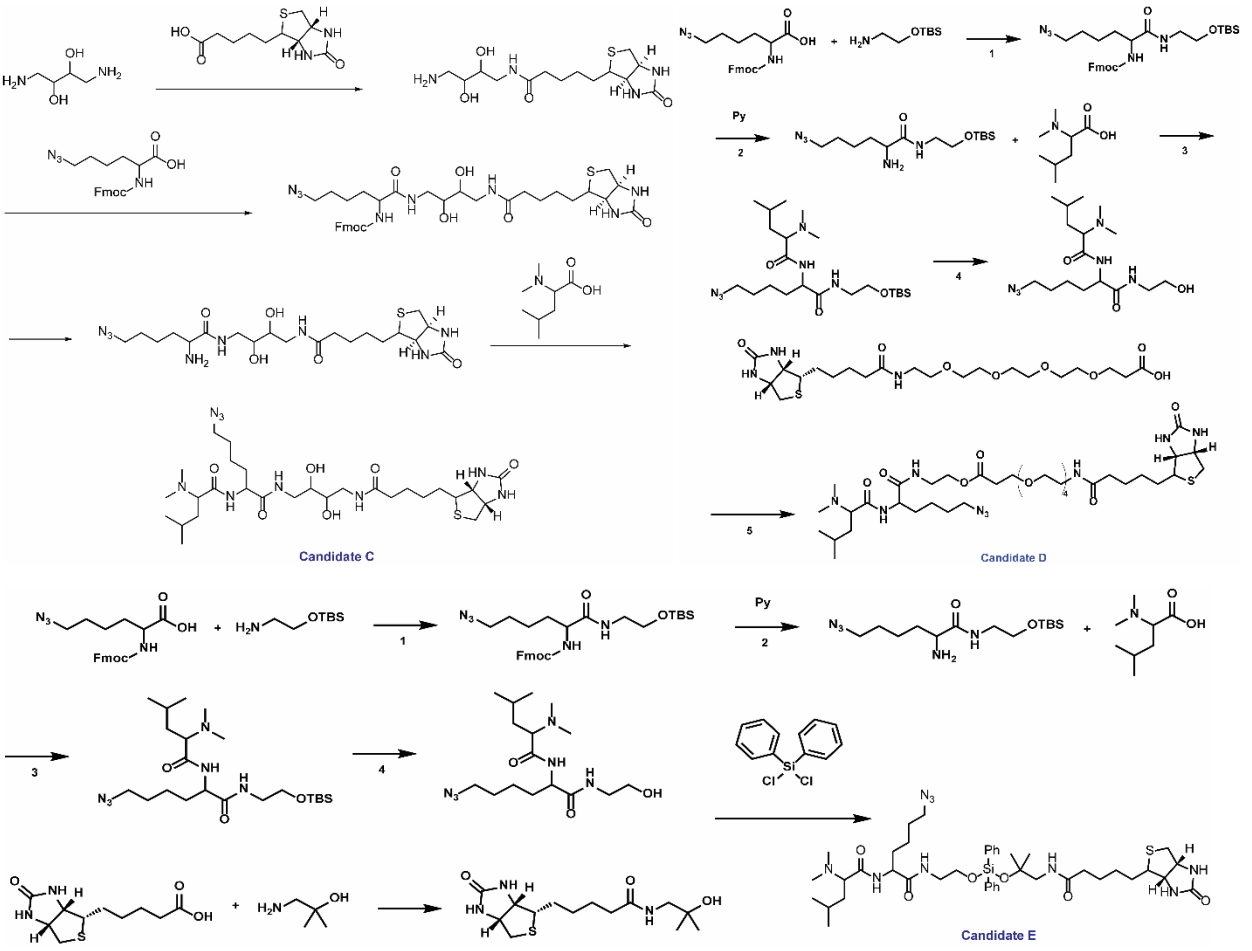


Figure 5b. Synthetic routes of DBA tag candidates C-E (candidate C: cleavable DBA tag with a diol linker; candidate D: cleavable DBA tag with an base cleavable linker, candidate E: cleavable DBA tag with DADPS linker).

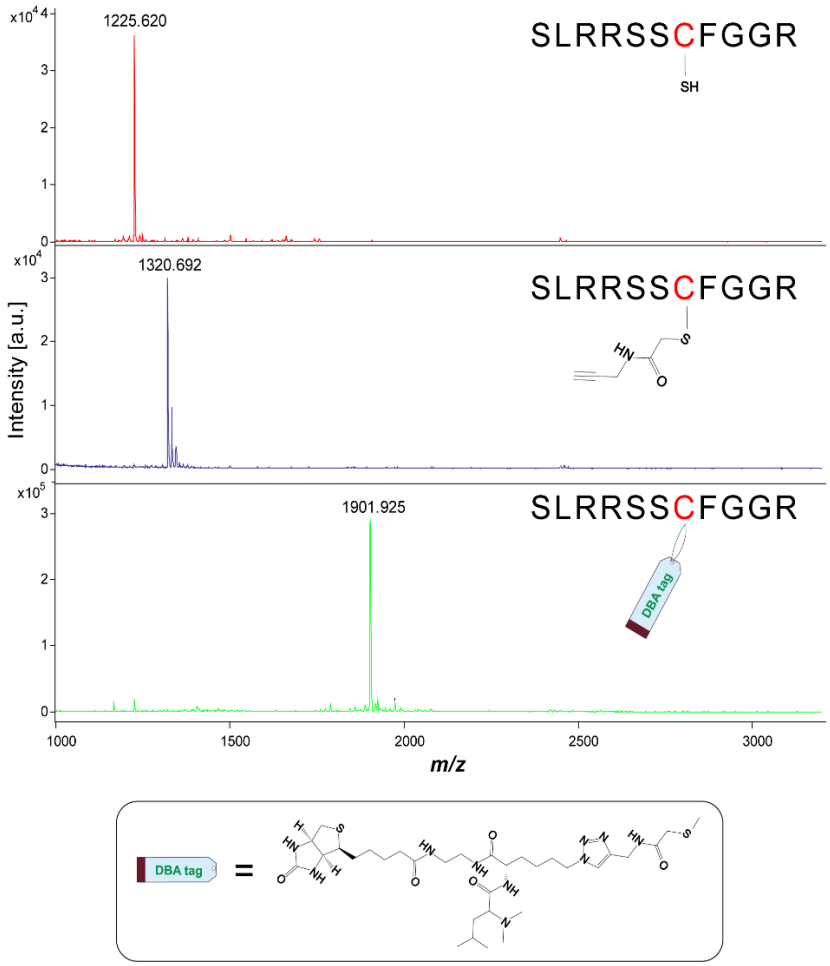


Figure 6. DBA tag labeling condition and performance evaluated using standard peptide ANP by MALDI-TOF MS. Results acquired using Candidate A (Top: native ANP peptide; middle: ANP labeled by IPM; bottom: ANP labeled by protoDBA candidate A).

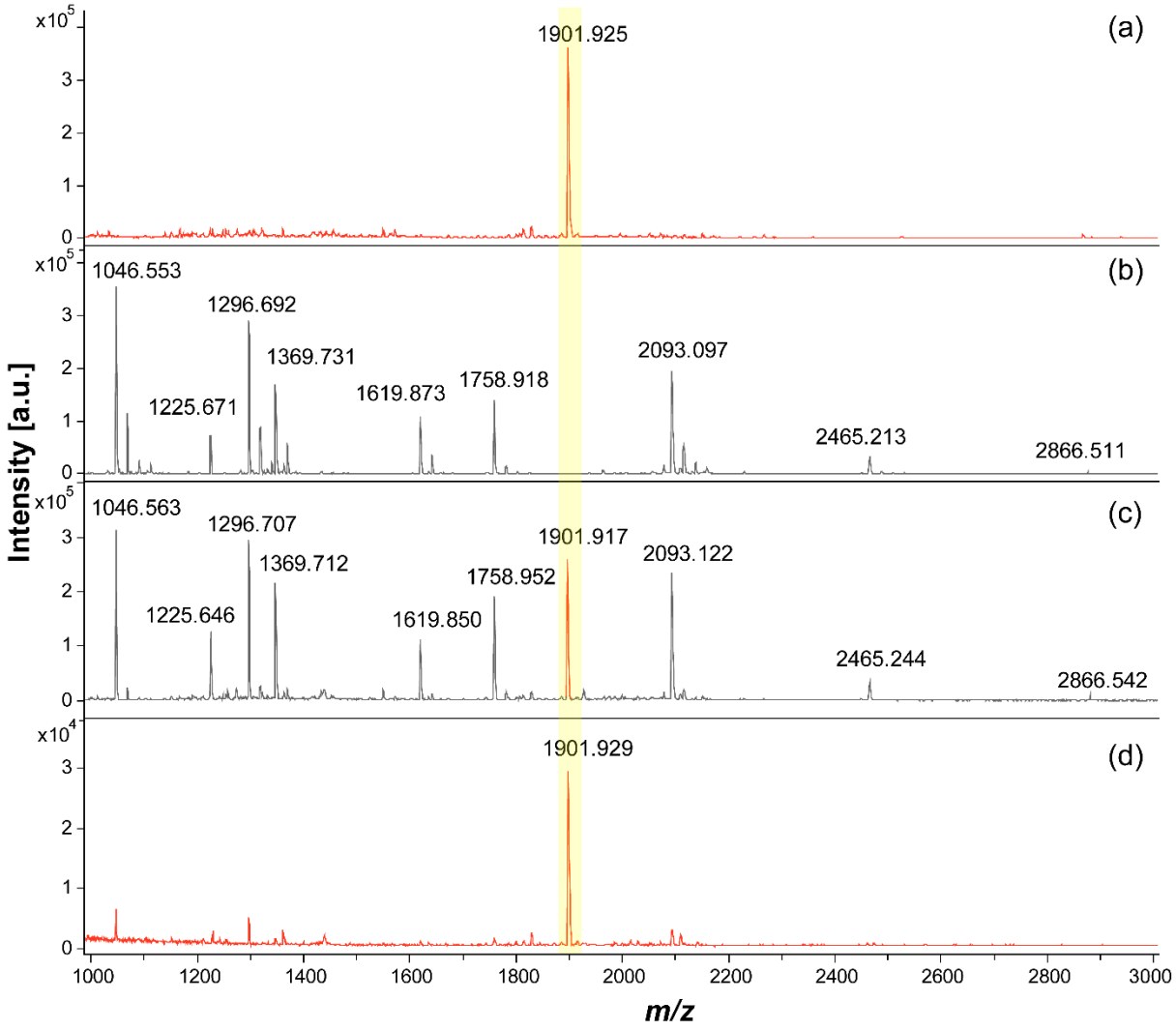


Figure 7. Enrichment efficiency evaluation on peptide mixture (a: ANP labeled by protoDBA candidate A; b: peptide mix; c: labeled ANP spiked into peptide mixture before enrichment; d: after enrichment).

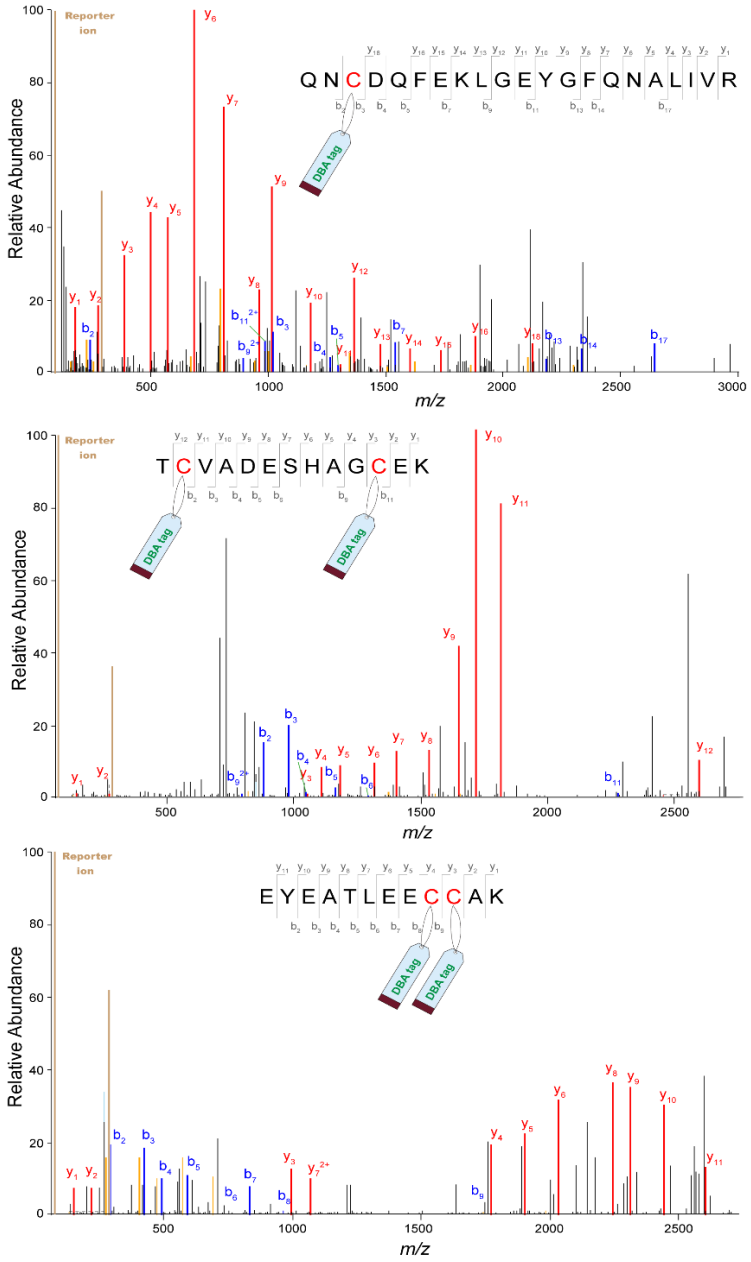


Figure 8. MS/MS of DBA labeled peptides from BSA (modification sites of DBA tags are unambiguously assigned by Maxquant; multiple sites can also be pinpointed).

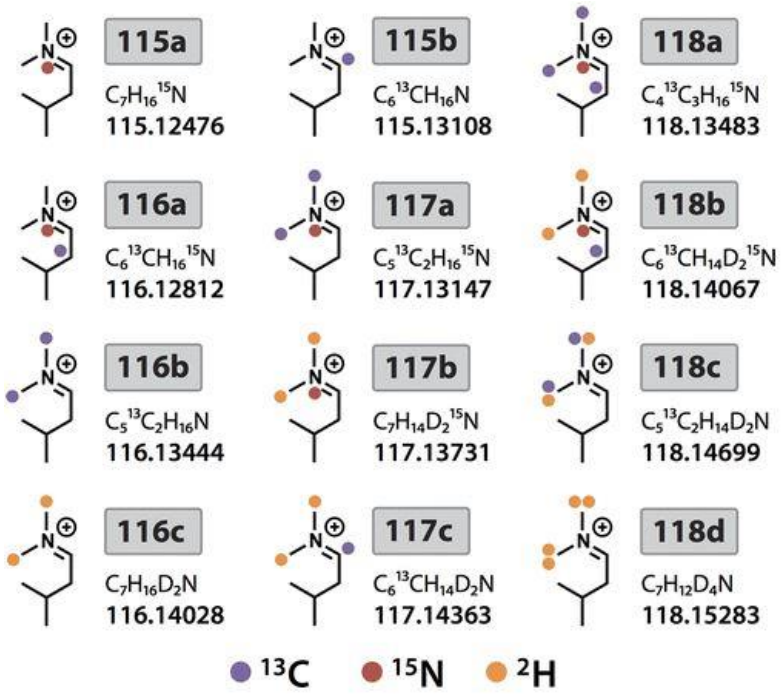


Figure 9. The 12-plex DiLeu reporter ion structures showing stable isotope positions.

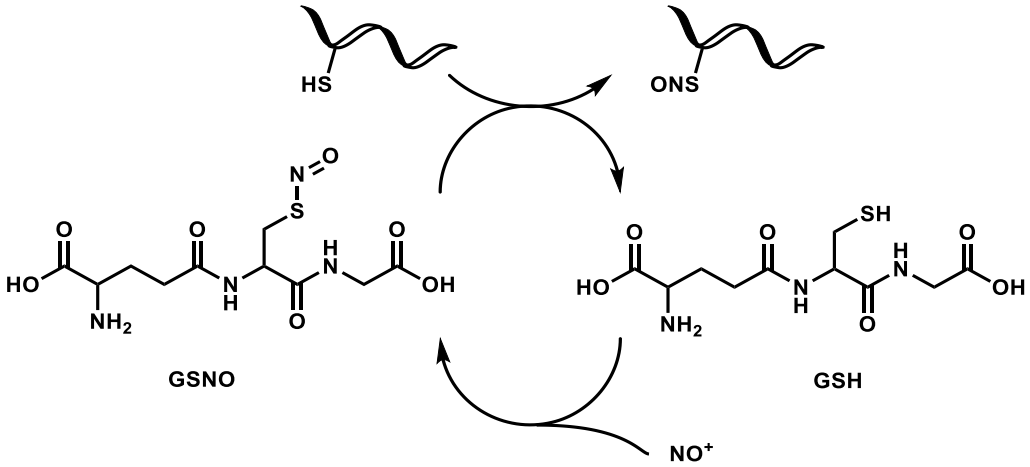


Figure 10. GSNO-mediated transnitrosylation.

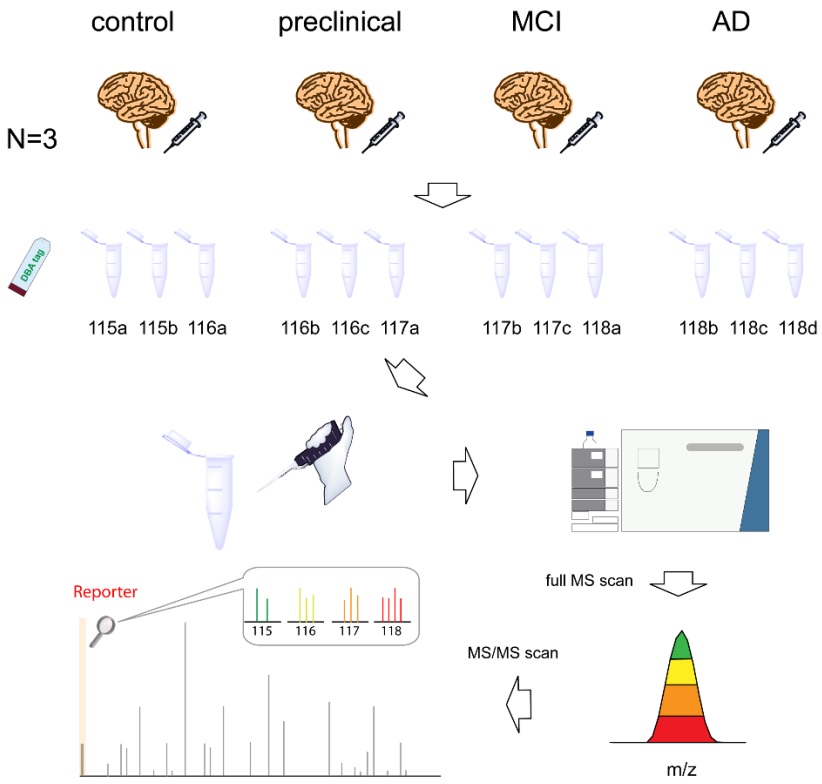


Figure 11. Workflow of 12-plex DBA tag enabled isoBOP-ABPP quantitative proteomics for SNO analysis of CSF samples collected from patients at different stages of AD.

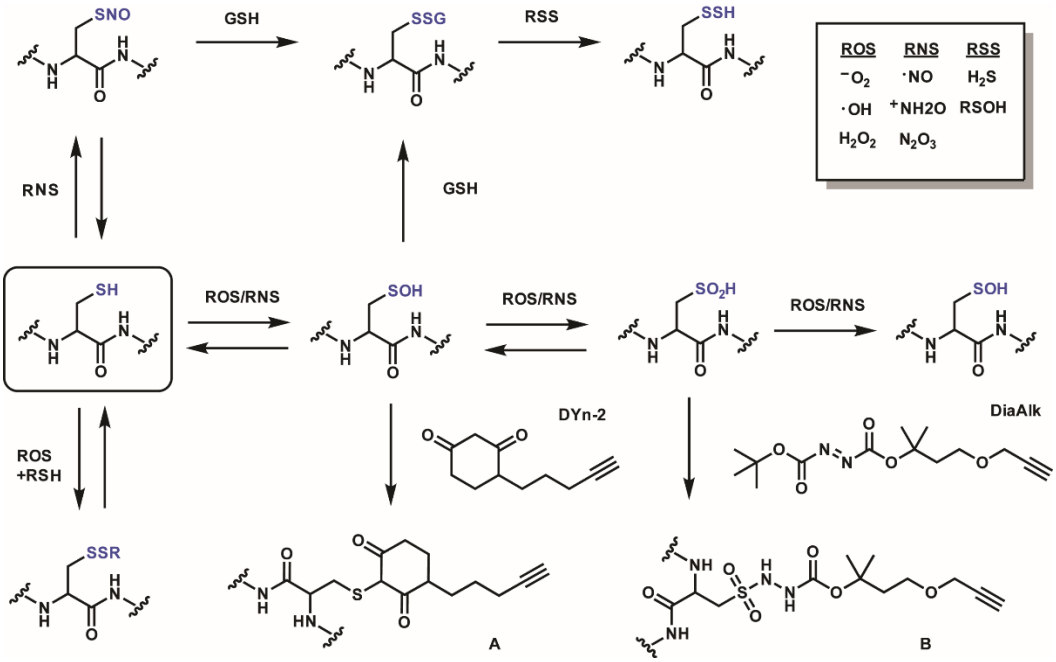


Figure 12. Biologically relevant oxidative PTMs of cysteine and chemical probes for -SOH and -SO_2H.

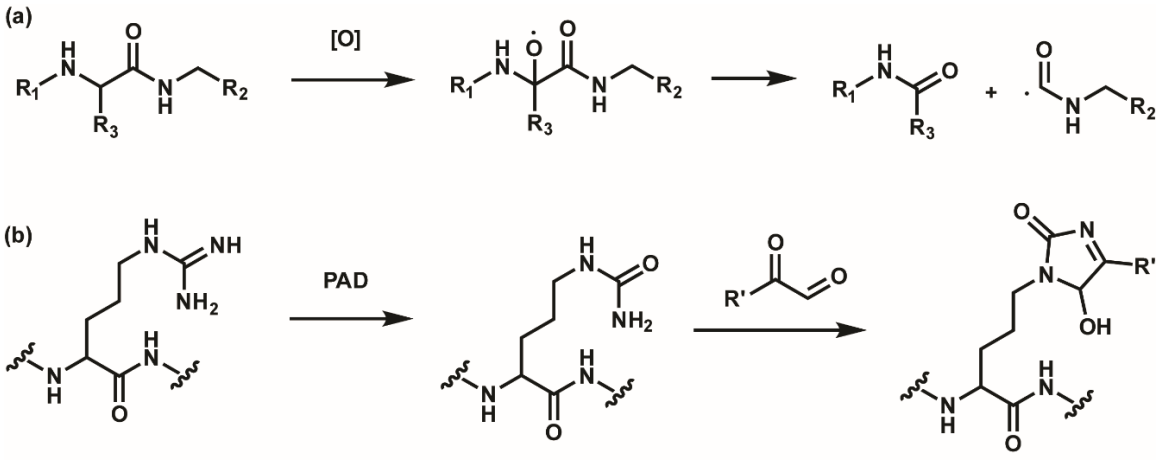


Figure 13. (a) Oxygen radical induced peptide backbone bond cleavage results in peptide carbonylation; (b) Arginine citrullination and phenylglyoxal chemical derivatization.

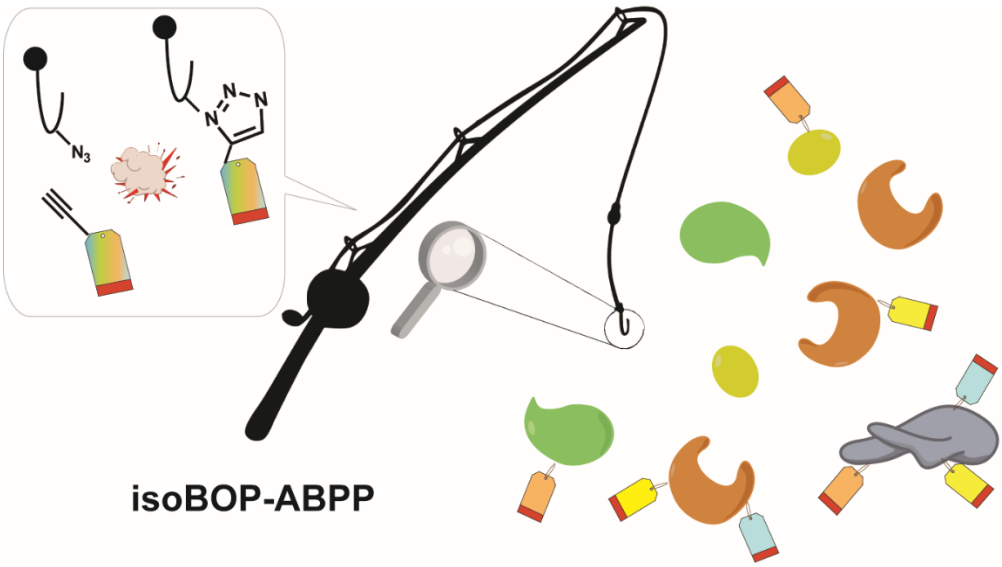


Figure 14. IsoBOP-ABPP provides an analytical platform for probing and quantifying multiple PTMs.

Chapter 7

Conclusions and Future Directions

Conclusions

This dissertation is devoted to the development and application of novel chemical probes to facilitate the analysis of biomolecules such as glycans, peptides and proteins.

The first two chapters in this thesis center around the topic of glycomics (Chapters 2-3). Glycosylation is one of the most important post-translational modifications (PTMs) with essential physiological functions, including protein folding, cell signaling and immune response. Thus, various qualitative and quantitative glycomics analysis strategies have been developed. Recently, isobaric multiplex reagents for carbonyl containing compound (SUGAR) tags were developed for quantitative glycomics with multiplexing capacity and increased reporter ion yield. To further improve quantification efficiency and enable quantifying low abundance species, mass defect based triplex SUGAR (mdSUGAR) tag has been designed and introduced in Chapter 2. In addition, we also introduced additional reaction sites for mdSUGAR at terminal sialic acid by periodate oxidation of polyhydroxy chain to extend the mass difference and lower the requirement for resolving power. As a result, mdSUGAR tags show complete labeling efficiency, improved fragmentation pattern and accurate quantification. Moreover, the quantitative performance of the mdSUGAR tags in complex system has been systematically evaluated and demonstrated reliable results. While extensive efforts have been devoted to the analysis of N-glycans, high throughput quantitative analysis of O-glycans is often overlooked and underexplored. This is partially due to the lack of universal enzyme for the release of O-glycans from protein backbone. Furthermore, the traditional chemical releasing method suffers from severe side reactions and involves tedious sample preparation procedures. Here, a multiplexed isobaric labeling method enabled by N, N-dimethyl leucine containing pyrazolone analog (DiLeuPMP) is introduced in Chapter 3. This method combines the release and labeling of O-glycans in a one-pot reaction and achieves accurate

MS2-based relative quantification with the ability to process four samples at a time. The method has been applied to core-1 O-glycan standard and three glycoproteins first and the results demonstrated its validity. Then complex biological specimen, the human serum sample was analyzed. Overall, this method provides an effective and reliable approach for the profiling and high-throughput quantitative analysis of O-glycans in complex samples.

Chapter 4 discusses our development of a streamlined workflow to selectively enrich sialylglycopeptides (SGPs). Despite the important roles of protein sialylation in biological processes such as cellular interaction and cancer progression, simple and effective methods for the analysis of intact SGPs are still limited. Analyses of low-abundance SGPs typically require efficient enrichment prior to comprehensive liquid chromatography-mass spectrometry (LC-MS)-based analysis. Here, a novel workflow combining mild periodate oxidation, hydrazide chemistry, copper-catalyzed azide/alkyne cycloaddition (CuAAC) click chemistry, and dynamic covalent exchange, has been developed for selective enrichment of SGPs. The intact SGPs could be separated easily from protein tryptic digests and the signature ions were produced during LC-MS/MS for unambiguous identification. The structure of the signature ions and corresponding dynamic covalent exchange was confirmed by using isotopic reagent. Under the optimized condition, over 70% enrichment efficiency of SGPs was achieved for bovine fetuin digests, and the method was successfully applied to complex biological samples, such as a mouse lung tissue extract. The high enrichment efficiency, good reproducibility, and easily adopted procedure without the need to generate specialized materials make this method a promising tool for broad applications in SGPs analysis.

Chapter 5 shifts focus from qualitative analysis of sialylglycopeptides to quantitative analysis. To characterize intact glycopeptide, mass spectrometry (MS)-based glycoproteomics

have employed versatile fragmentation methods, among which Electron-Transfer/Higher-Energy Collision Dissociation (EThcD) has gained great popularity. However, the inherent limitation of EThcD in fragmenting low-charge ions has prevented its widespread applications. Moreover, there is a need to develop a high-throughput strategy for comparative glycoproteomics in a large cohort of samples. Herein, we developed isobaric *N*,*N*-dimethyl Leucine derivatized ethylenediamine (DiLeuEN) tags to increase the charge states of glycopeptides, thereby improving the fragmentation efficiency and allowing for in-depth intact glycopeptide analysis, especially sialoglycopeptides. Moreover, the unique reporter ions of DiLeuEN-labeled glycopeptide generated in tandem MS spectra enable relative quantification of up to four samples in a single analysis, which represents a new high-throughput method for quantitative glycoproteomics.

A novel chemical probe featuring cleavable DiLeu-Biotin-Azide (cDBA) was introduced in Chapter 6, allowing a versatile chemical proteomics platform to be established. The idea of this project stems from expanding the analytical capability of traditional isotopic tandem orthogonal proteolysis–activity-based protein profiling (isoTOP-ABPP) strategy. By incorporating isobaric labeling concept and our DiLeu tag design, several candidates of cDBA tag were designed and synthesized. After evaluating the performance of these candidates including labeling/enrichment efficiency and fragmentation pattern in tandem mass spectrometry (MS/MS), the best candidate was selected. The usage of cDBA tag was further explored due to the exquisite design of its chemical structure and unique features of biorthogonal reactions. By altering alkyne probe to selectively label different PTM sites, various PTMs can be derivatized, captured, enriched, and analyzed. The versatility of the application of cDBA tag makes it an ideal pan-PTM high throughput quantitative analysis platform. Furthermore, the combination of different PTM

constitutes the next level of complexity for proteomics research and the study of this PTM crosstalk can also benefit from our pan-PTM proteomics platform.

Future Directions

There are several extensions and ongoing projects derived from the research described in this dissertation, and the depth and scope of their applications can be further explored. For the N-glycomics study, the 3-plex mdSUGAR tag is being further developed into 6-plex set, adding another three channels where the average mass difference between each channel is around 10mDa. The updated 6-plex mdSUGAR tag doubles the throughput comparing to original 3-plex set and allows a novel strategy called triplex labeling to be implemented. In triplex labeling strategy, samples are divided into three equal parts before performing chemical derivatization. After labeling with three different channels of mdSUGAG tag, all samples are pooled together before doing LC-MS/MS analysis. In this way, the same glycans within the sample shall have a triple-peak under ultra-high mass resolution (commonly > 1M). This unique feature enables tremendous sensitivity for the identification of glycans even at the single-cell level based on the full mass scan. The MRM or PRM analysis can then be conducted based on the initial full mass screening, and signature glycan fragments (such as *m/z* 138, *m/z* 168, *m/z* 186 and *m/z* 204) as well as patterns (+162 m/z interval) can be further utilized for the confirmation of the targeted peaks.

It is reported that PMP can not only be used for the labeling of O-glycans, but also applicable to derivatize the N-glycans. This triggers us to develop our DiLeuPMP as a universal isobaric tag for quantitative analysis of N-glycans and O-glycans simultaneously. Currently, the quantification of N-glycans and O-glycans sequentially relies on different labeling strategies and

different tags, limiting the direct comparison of the N-glycans and O-glycans within the same sample. We reasoned that the exact same chemical structure (DiLeuPMP) incorporation would provide similar levels of ionization efficiency, thus making direct analysis of the signal between N-glycans and O-glycans possible. Nevertheless, delicate sample preparation workflow and quantitative quality control need to be further evaluated.

Taking the advantages of cDBA tag and based on the isoBOP-ABPP platform, various chemical proteomics studies can be achieved. The tetrafunctional architecture of cDBA tag is highly versatile, capable of isobaric labeling, click chemistry conjugation and selective enrichment at one time. Since the quantification is based on MS2, high-throughput analysis can be readily implemented. No similar commercially available tags ever utilize MS2 based quantification strategy before, and they are not able to achieve more than duplex analysis in a single experiment, while our DBA tag can easily analyze 12 or 21 samples simultaneously.

cDBA tag serves as a platform for various PTM analysis which is benefited from its modular design. Through activity-based protein profiling (ABPP) strategy, special alkyne probes can target various PTMs including but not limited to cysteine oxidation, protein carbonylation and citrullination, where all of those can be labeled by DBA tag via biorthogonal click chemistry. We demonstrated its application to S-nitrosylation (SNO), S-sulphenation (SOH) and protein carbonylation using peptide standards. The detailed project plans are directed to the “future direction” section of Chapter 6.

Generally, the chemical probes developed in this thesis have proved great potential for the analysis of various biomolecules such as glycans, peptides and proteins. We anticipated these chemical tools can benefit the biomedical research in the filed and facilitate the discovery of biomarkers and exploration of biological mechanisms.

Appendix

Academical Summary

Peer-Reviewed Publications

- **Li, M.**; Huang, J.; Ma, M.; Shi, X.; Li, L. Selective Enrichment of Sialylglycopeptides Enabled by Click Chemistry and Schiff-base Covalent Exchange. *Anal. Chem.* **2022**, 94, 18, 6681–6688.
- **Li, M.**[#]; Gu, T.[#]; Lin, X.; Li, L. DiLeuPMP: a multiplexed isobaric labeling method for quantitative analysis of O-glycans. *Anal. Chem.* **2021**, 93, 28, 9845–9852.
- Feng, Y.[#]; **Li, M.**[#]; Lin, Y.; Chen, B.; Li, L. Multiplex Quantitative Glycomics Enabled by Periodate Oxidation and Triplex Mass Defect Isobaric Multiplex Reagents for Carbonyl Containing Compound Tags (mdSUGAR). *Anal. Chem.* **2019**. 91, 18, 11932-11937.
- Hou, L.; Li, Y.; Wu, Q.; **Li, M.**; Older, EA.; Tang, X.; Nagarkatti, P.; Nagarkatti, M.; Liu, Y.; Li, L.; Fan, D.; Bugni, TS.; Shang, Z.; Li, J. Discovery of anti-infective adipostatins through bioactivity-guided isolation and heterologous expression of a type III polyketide synthase. *Bioorg. Chem.* **2021**, 112, 104925.
- Zhong, X.; Frost, DC.; Yu, Q.; **Li, M.**; Gu, T.; Li, L. Mass Defect-Based DiLeu Tagging for Multiplexed Data-Independent Acquisition. *Anal. Chem.* **2020**, 92, 16, 11119–11126.
- Zhang, H.; Shi, X.; Vu, NQ.; Li, G.; Li, Z.; Shi, Y.; **Li, M.**; Wang, B.; Welham, NV.; Patankar, MS.; Weisman, P.; Li, L. On-Tissue Derivatization with Girard's Reagent P Enhances N-Glycan Signals for Formalin-Fixed Paraffin-Embedded Tissue Sections in MALDI Mass Spectrometry Imaging. *Anal. Chem.* **2020**, 92, 19, 13361–13368.
- Pan, P.; Chen, J.; Li, X.; **Li, M.**; Yu, H.; Zhao, J. J.; Ni, J.; Wang, X.; Sun, H.; Tian, S.; Zhu, F.; Liu, F.; Huang, Y.; Hou, T. Structure-Based Drug Design and Identification of H_2O-Soluble and Low Toxic Hexacyclic Camptothecin Derivatives with Improved Efficacy in Cancer and Lethal Inflammation Models in Vivo. *J. Med. Chem.* **2018**, 61, 19, 8613-8624.
- Ma, M.; Yu, Q.; Delafield, G D.; Westmark, R P.; Cui, Y.; Li, Z.; **Li, M.**; Ma, G.; Wu, W.; Shi, X.; Gutierrez, A.; Gao, A.; Xu, M.; Xu, W.; Westmark, J C.; Li, L. On-tissue spatial proteomics integrating MALDI-MS imaging with shotgun proteomics reveals soy consumption-induced Candidate Markers in a fragile X syndrome mouse model. *Under review.*
- Wang, D.; Ma, M.; Huang, J.; Gu, T.; Cui, Y.; **Li, M.**; Wang, Z.; Zetterberg, H.; Li, L. Boost-DiLeu: Enhanced Isobaric *N,N*-Dimethyl Leucine Tagging Strategy for Comprehensive Quantitative Glycoproteomic Analysis. *Under review.*
- **Li, M.**[#]; Zhong, X.[#]; Feng, Y.; Li, L. Novel Isobaric Tagging Reagent Enabled Multiplex Quantitative Glycoproteomics via Electron-Transfer/Higher-Energy Collisional Dissociation (EThcD) Mass Spectrometry. *Ready for submission.*
- Feng, Y.[#]; **Li, M.**[#]; Li, L. 12-plex SUGAR tag enabled pancreatic cancers cells glycomics. *Ready for submission.*
- Donahue, K.; Xie H.; **Li, M.**; Gao, A.; Wang, Y.; Tipton, R.; Belmar, J.; Semanik, N.; Park, H B.; Li, S.; Li, L.; Tang, W.; Xu, W. Diptoindonesin G is a novel HSP90 modulator for cancer treatment. *Ready for submission.*
- **Li, M.**[#]; Ma, M.[#]; Li, L. Cleavable DiLeu-Biotin-Azide (cDBA) Tag Enabled isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP) Chemical Proteomics Platform for High-throughput Quantitative pan-PTM Analysis. *Manuscript in preparation.*
- **Li, M.**[#]; Ma, M.[#]; Li, L. DiLeuC enabled complementary ion based quantitative proteomics. *Manuscript in preparation.*
- Ma, M. [#]; **Li, M.**[#]; Wang, Z.; Li, L. 6-plex mdSUGAR tag enabled single-cell glycomics. *Manuscript in preparation.*
- Wang, Z.; **Li, M.**; Li, L. Multiplex quantitative analysis of fatty acids enabled by 12-plex

DiLeuEN and capillary electrophoresis. *Manuscript in preparation*.

- Wang, D.[#]; **Li, M.[#]**; Li, L. 12-plex SUGAR tag enabled by quantitative analysis of sialoglycopeptides. *Manuscript in preparation*.

Conference Presentations

Oral Presentation

- **Li, M.**; Li, L., "Novel 7-plex DiLeuC isobaric tags based on complementary-ion quantification enables accurate, precise, and sensitive high-throughput quantitative proteomics" Oral presentation. 70th ASMS Annual Conference on Mass Spectrometry and Allied Topics. Minneapolis, MN. June 2022.

Poster Presentation

- **Li, M.**; Li, L. Selective Enrichment of Sialylglycopeptides Enabled by Click Chemistry and Schiff-base Covalent Exchange. Poster. 69th Annual ASMS Conference on Mass Spectrometry and Allied Topics. Philadelphia, PA. October 2021.
- **Li, M.**; Zhong, X.; Li, L. High Throughput Fine Mapping of Glycosylation Enabled by DiLeuEN and ETD-MS. Poster. 68th Annual ASMS Conference on Mass Spectrometry and Allied Topics. Online. June 2020.
- **Li, M.**; Feng, Y.; Li, L. Quantitative Glycomics with Improved Multiplexing Performance by Mass-Defect SUGAR Tags and Both-ends Labeling. Poster. 67th Annual ASMS Conference on Mass Spectrometry and Allied Topics. Atlanta, GA. June 2019.
- **Li, M.**; Feng, Y.; Li, L. Amino Acid-Based Mass Defect Chemical Tags for Glycomics Analysis. Poster. 66th Annual ASMS Conference on Mass Spectrometry and Allied Topics. San Diego, CA. June 2018.

Patents

- Li. L., **Li. M.** Ma. M. Multiplexed DiLeu-Biotin-Azide (DBA) Tag enabled isobaric tandem orthogonal proteolysis activity-based protein profiling (isoBOP-ABPP) platform for high-throughput quantitative pan-PTM analysis. Application Number: P220286US01.
- Li. L., Feng, Y., Chen, B., Yu, Q., Zhong, X, **Li, M.** Novel 4-plex isobaric aldehyde reactive N, N-Dimethyl Leucine derivative tags for quantitative glycomics analysis. Publication number: 20190225559.

Honors and Awards

- 2022 GSFLC Travel Grant, Department of Chemistry, University of Wisconsin Madison.
- 2020 Graduate Student Travel Award, 68th ASMS Conference.
- 2020 Hay Fellowship, Department of Chemistry, University of Wisconsin Madison.
- 2019 Schrag Travel Grant, Department of Chemistry, University of Wisconsin Madison.