Lysine & Arginine Protein Post-translational Modifications by Enhanced DIA Libraries: Quantification in Murine Liver Disease

Aaron E. Robinson,[1] Aleksandra Binek,[1] Vidya Venkatraman,[1] Brian C. Searle,[2] Ronald J. Holewinski,[1] George Rosenberger,[3] Sarah J. Parker,[1] Nathan Basisty,[4] Xueshu Xie,[4] Peder J. Lund,[5] Gautam Saxena,[6] José M. Mato,[7] Benjamin A. Garcia,[5] Birgit Schilling,[4] Shelly C. Lu,[8] Jennifer E. Van Eyk,[1]*

[1]Advanced Clinical Biosystems Research Institute, The Smidt Heart Institute, Cedars Sinai Medical Center, Los Angeles, CA 90048, USA; [2]Department of Genome Sciences, University of Washington, Seattle, WA, USA; [3]Department of Systems Biology, Columbia University, New York, NY, USA; [4]Buck Institute for Research on Aging, Novato, CA 94945, USA; [5]Department of Biochemistry and Biophysics, Epigenetics Institute, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA; [6]DeepDIA, USA [7]CIC bioGUNE, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (Ciberehd), Technology Park of Bizkaia, 48160 Derio, Bizkaia, Spain; [8]Division of Digestive and Liver Diseases, Cedars-Sinai Medical Center, Los Angeles, CA, 90048, USA

Corresponding author:

Jennifer E. Van Eyk

Jennifer.VanEyk@cshs.org

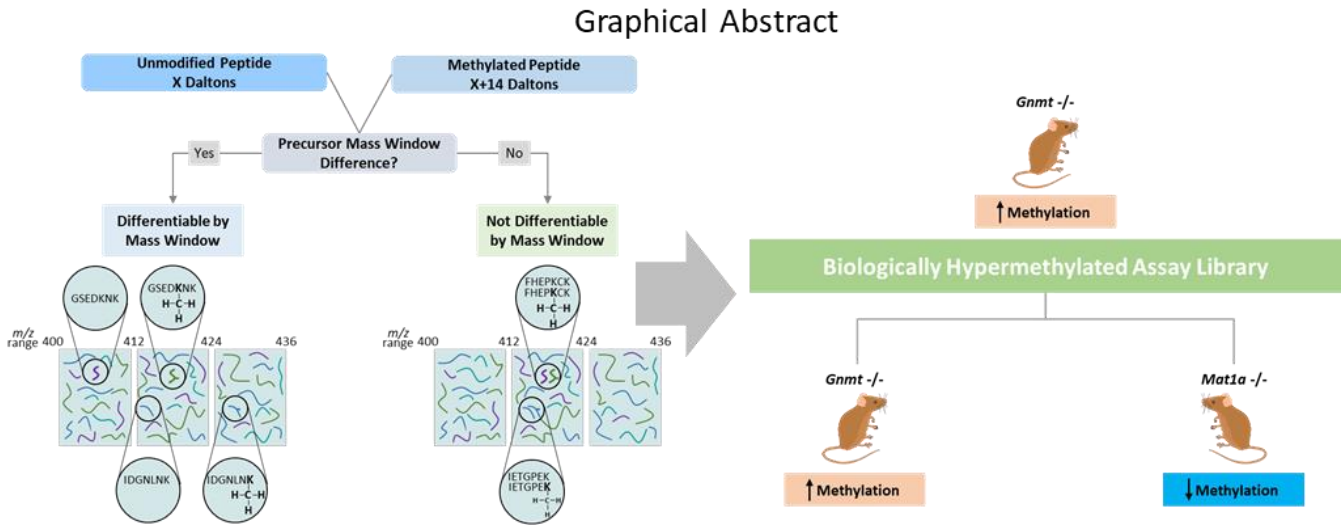
127 S. San Vicente Blvd, Advanced Health Sciences Pavilion, 9302, Los Angeles, CA 90048

Abstract

Arg and Lys undergo multiple enzymatically driven post-translational modifications (PTMs), including methylation, which may be linked to key cellular processes. Today there is not a method that simultaneously quantifies proteins and the methylome at the site level. We have developed a method that can differentiate an unmodified peptide from its mono-, di- or tri-methylated Arg or Lys counterpart through a data independent acquisition approach that is suitable for both triple TOF and Orbitrap mass spectrometers. This method was further expanded to include Lys acetylation and succinylation, which in conjunction with methylation allow for simultaneous quantification of Lys PTMs in a single measurement. Our approach leverages small precursor mass windows to physically separate the various methylated species from each other during MS2 acquisition and shows linearity in the quantitation of low abundant peptides. We did this by first creating a biologically hyper-methylated peptide assay library, for mono-, di- and tri-methylated Lys and mono- and di-methylated Arg to which each experimental sample is compared. We further expanded the peptide assay library to include Lys acetylation and succinylation, providing the opportunity to multiplex five Lys and two Arg modification without the need for sample enrichment. To assess the validity of PTMs quantified with this method, we added an algorithm to determine the false localization rate for each potential modified amino acid residue. To evaluate its biological relevance, we applied this method to complex liver lysate from differentially methylated in-vivo non-alcoholic steatohepatitis mouse models and show that differential methylation potential altered not only protein quantity but also acetylation and to some degree succinylation. Of interest, acetylation data together with total protein changes drive strong novel hypothesis of a regulatory function of post-translational modifications in protein synthesis and mRNA stability. In conclusion: our method, which is suitable for different mass spectrometers along with the corresponding publicly available mouse liver PTM enriched peptide assay library, provides an easily adoptable framework needed for studying global protein methylation, acetylation, and succinylation in any proteomic experiment when total protein

quantification is being carried out. Wide adoption of this easy approach will allow more rapid integration of PTM studies and knowledge.

Graphical Abstract



Introduction

Reversible post-translational modifications (PTMs) of proteins such as phosphorylation of Ser, Thr, and Tyr and methylation of Arg and Lys are of great interest as they act as dynamic regulators of cellular signaling.[1–3] While there have been substantial advances in biochemical methods for the detection and subsequent mass spectrometry (MS) to quantify phosphorylation (e.g., metal affinity and immunoaffinity enrichments), detection of protein methylation has lagged behind.[4,5] PhosphositePlus, a depository of PTMs identified via MS has reported 290,117 unique phosphosites while annotating only 18,394 unique methyl sites.[6] The low stoichiometry of PTM peptidoforms has been a persistent limitation to their detection in shotgun MS, particularly when data dependent precursor selection favors the more abundant peptides in a complex sample. Technological advancements in MS instrument sampling speed have expanded the dynamic range of precursor sampling for fragmentation by DDA, and concomitantly enabled reductions in precursor window widths for data independent acquisition (DIA) based fragmentation while still maintaining reasonable cycle times. The added benefit of DIA based sampling to reduce missing data and thereby improve the consistency of peptide detection and quantitation promises the continued advancement for the MS-based study of low abundant PTMs. In addition, recently published algorithms PIQUed,[7] Peptide Collapse plugin for Perseus,[8] Inference of Peptidoforms (IPF),[9] and Thesaurus,[10] claim reduced false site localization and greater PTM recovery providing an alternative to the classical approach to elucidate PTM dynamics in complex biological systems and a more complete understanding of the complexity of cellular signaling, which we have harnessed.[4,5,9,11–13]

In proteins, methylation generally occurs on Arg and Lys residues with the ε-amino group of Lys being either mono, di, or trimethylated by protein lysine methyltransferases while Arg can be mono or di-methylated by protein arginine methyltransferases. In mammalian cells, S-Adenosylmethionine (SAMe) is the exclusive methyl donor for both Lys and Arg methylation.[3] The majority of research in protein methylation has focused on histone methylation and its role in epigenetics. Histone methylation of Lys and Arg is an integral part of the histone code known to alter the physical properties of chromatin

subsequently altering gene expression.[3] Conversely, methylation of non-histone proteins has only recently emerged as a biologically relevant PTM and recent research has shown its importance in cellular signaling and function.[14–18] This information has in part been obtained from global protein methylation proteomic analysis which has recently been enabled by the development of labeling and enrichment techniques like hM SILAC, high and low pH SCX fractionation, and utilizing peptide immunoprecipitations of methylated Lys and Arg and subsequent mass spectrometry.[19–22] However, antibody based methyl-enrichments have a number of challenges and limitations: (1) they require a large amount of sample and multiple antibodies to cover all methyl forms; (2) they often contain a sequence bias for antigen recognition, thus selectively assaying a portion of the methylome while masking an unknown but likely significant portion of methyl-sites; and (3) by nature these approaches do not enrich the unmodified form of the peptide, hindering quantification of all peptidoforms together and thus impairing the ability to study the relative abundance of PTMs.[23,24]

To overcome these limitations and facilitate a comprehensive and technically straightforward quantitative survey of the methylome in biological samples, we leveraged small precursor acquisition window DIA-MS, coupled with careful curation of a combined low pH SCX fractionated and methyl peptide immunoprecipitation enriched peptide library consisting of both unmodified and methylated peptides against which each experimental sample is extracted and compared. In addition to methylation, lysine residues can be modified by various other PTM's, including acetylation and succinylation.[25–29] To understand the extent of lysine post translational modification, we expanded our DIA assay library to include succinylated and acetylated peptides, thus providing an opportunity to quantify methylated, acetylated, succinylated, and unmodified peptides from the same sample and acquisition run. This not only allows for classical cellular protein quantification, but also provides the opportunity to quantify peptidoforms containing Lys methylation, succinylation, acetylation, and Arg methylation simultaneously from a single DIA acquisition. Traditional DDA and DIA based PTM-immunoprecipitation enrichment approaches rely on antibody enrichment for the modified peptide and by nature are unable to quantify

unmodified peptides in the same acquisition or enrich at the single modified protein which has limited scope.[4,28] In our method, we are able to quantify both modified and unmodified peptides of seven different PTMs using DIA acquisitions with small precursor mass window requiring only five percent of starting material necessary to perform PTM immunoprecipitations. The important distinction with using small precursor mass windows is that we are able to physically trap each modified peptidoform away from other modified peptidoforms in addition to the unmodified peptide, so that all mono-methylated peptides are in a different precursor mass window than the unmodified peptidoform as well as the di- the tri-methylated peptidoforms. Combined with an algorithm which provides localization scores and false localization rate (FLR) calculations for each modified amino acid residue, our approach provides both protein quantification and high probability quantitative data simultaneously on seven PTMs.

To highlight biological insight that can be gained using our small precursor mass window DIA method, two mouse models of nonalcoholic steatohepatitis (NASH) where employed, each chosen due to their extreme methylation potential. Glycine N-Methyltransferase knockout (*Gnmt* -/-) mice have increased SAMe, have higher methylation potential, and are biologically hypermethylated while methionine adenosyltransferase A1 knockout (*Mat1a* -/-) mice are SAMe deficient, have reduced methylation potential, and are therefore hypomethylated.[30–32] Both animal models spontaneously develop NASH and resemble major subtypes of human disease.[30,33] Interestingly, disease in both animal models can be treated by normalization of SAMe levels highlighting the importance of SAMe and methylation capacity in NASH. [30,34,35]

Using the 4 Da DIA-MS method with FLR algorithm decreases limitations for PTM analysis as relative quantification of both total protein and seven different PTMs can be obtained in the same mass spectrometry run. This allows for a technically and analytically superior comparison of proteomic and relative PTM abundance change in complex lysate across different biological conditions for these two animal models of NASH. We hypothesized that these two NASH models would highlight the role that

differential methylation plays on altered proteins involved in common pathways. Nevertheless, we also encountered less expected connections, such as in protein acetylation showing similar signatures in both models, that could be potentially explained by two different mechanisms.

Materials and Methods

Methylated Peptide Synthesis: 100 synthesized unmodified and mono-, di-, tri-methylated and acetylated peptides (JPT Peptide Technologies GmbH, Berlin, Germany) were supplied as lyophilized powder. Peptides were resuspended in 0.1% Formic Acid and pooled at 1:1 ratio, aliquoted, and stored at –80°C until used. Stabile Isotopically labeled (SIL) methyl-peptides were synthesized and were supplied as a lyophilized powder. The supplier provided product characterization (MALDI-TOF and HPLC traces) as proof of MW and purity accuracy. The peptides were of >95% purity (New England Peptide). Peptides were solubilized in 0.1% Formic Acid and pooled at 1:1 ratio, aliquoted, and stored at –80°C until used.

Liver Tissue Sample Preparation: Livers were obtained from ten-month old methionine *Mat1a* -/- male mice in a C57Bl/6 background with hepatic lipid accumulation and their aged-matched wild-type (WT) male sibling littermates,[30] and three-month old *Gnmt* -/- mice in a C57Bl/6 background with hepatic lipid accumulation with age matched WT littermates.[31,32] Animals were bred and housed in the CIC bioGUNE animal unit, accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC). Animals were fed with standard commercial chow animal diet (Ref. 2914, Envigo, Barcelona, Spain).

Frozen mouse livers (n=6/condition) were ground to powder under liquid N2 in a cryohomogenizer (Retsch). Tissue powder was thawed and lysed in 8M urea and 100mM TRIS-HCL, pH 8.0 and samples were ultrasonicated (QSonica) at 4 ºC for 10 min in 10 second repeating on/off intervals of 10 seconds and centrifuged at 16,000 x g for 10min at 4 ºC. The protein concentration of soluble supernatant was determined via Bicinchoninic Acid Assay (Thermo). 100µg of protein was reduced with DTT (15mM) for 1h at 37 ºC, alkylated with iodoacetamide (30mM) for 30min at room temperature in the dark, diluted to a final concentration of 2M Urea with 100mM TRIS-HCL, pH 8.0 and digested for 16 hours on a shaker at 37 ºC with a 1:40 ratio of Trypsin/Lys-C mix (Promega). Each sample was de-salted using HLB plates (Oasis HLB 30µm, 5mg sorbent, Waters).

To establish a broad methylation peptide assay library, 500µg of lysate from mouse livers (n=1/condition) was digested as described above. Each sample were de-salted (Oasis HLB, 10 mg sorbent cartridges) and Strong Cation Exchange (SCX) fractionated as described in *Kooij et al.*[36] with some modifications. De-salted peptides were dried and re-suspended in SCX buffer A (7 mM KH_2PO_4, 30% ACN, pH 2.65). The SCX columns were prepared by adding 100 mg of SCX bulk media (polysulfoethyl aspartamide, Nest Group) suspended in 1ml of 30% ACN to an empty 1 ml cartridge (Applied Separations), allowing the slurry to settle and sealing the column bed with an additional frit. The column was wet with 3 ml of 80% ACN, followed by a 3 ml H_2O wash and then equilibrated with 6 ml of SCX buffer A. Samples were loaded on the columns and the flow through was collected. Columns were washed two more times with SCX buffer A and the flowthroughs were combined as the first fraction. Samples were sequentially eluted with 1 ml SCX buffer A containing four different salt concentrations (10mM KCL, 40 mM KCl, 60mM KCL, and 150 mM KCl). All 5 fractions were dried, re-suspended in 1ml 1% FA, and desalted on using HLB plates (Oasis HLB 30µm, 5mg sorbent, Waters).

MS acquisition: MS data were acquired on either 6600 TripleTOF (Sciex) or Orbitrap Fusion Lumos (Thermo) mass spectrometers operating in DDA-MS or DIA-MS mode. iRT Standards (Biognosys) were added to each sample before acquisition.

The Orbitrap LUMOS mass spectrometer (Thermo Scientific) was equipped with an EasySpray ion source and connected to Ultimate 3000 nano LC system (Thermo Scientific). Peptides were loaded onto a PepMap RSLC C18 column (2 µm, 100 Å, 150 µm i.d. x 15 cm, Thermo) using a flow rate of 1.4 µL/min for 7 min at 1% B (mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1 % formic acid in acetonitrile) after which point they were separated with a linear gradient of 5-20%B for 45 minutes, 20-35%B for 15 min, 35-85%B for 3 min, holding at 85%B for 5 minutes and re-equilibrating at 1%B for 5 minutes. Each sample was followed by a blank injection to both clean the column and re-

equilibrate at 1%B. The nano-source capillary temperature was set to 300 °C and the spray voltage was set to 1.8 kV.

For DDA analysis MS1 scans the AGC target was set to $4x10^5$ ions with a max fill time of 50 ms. MS2 spectra were acquired using the TopSpeed method with a total cycle time of 3 seconds and an AGC target of $5x10^4$ and a max fill time of 22 ms, and an isolation width of 1.6 Da in the quadrapole. MS1 scans were acquired in the Orbitrap at a resolution of 60,000 FWHM from mass range 400-1000m/z. Precursor ions were fragmented using HCD with a normalized collision energy of 30% and analyzed in the Orbitrap at 15,000K resolution. Monoisotopic precursor selection was enabled and only MS1 signals exceeding 50000 counts triggered the MS2 scans, with +1 and unassigned charge states not being selected for MS2 analysis. Dynamic exclusion was enabled with a repeat count of 1 and exclusion duration of 15 seconds.

For DIA analysis, MS1 scans were acquired in the Orbitrap at a resolution of 60,000 FWHM from mass range 400-1000m/z. For MS1 scans the AGC target was set to $3x10^5$ ions with a max fill time of 50 ms. DIA MS2 scans were acquired in the Orbitrap at a resolution of 15000 FWHM with fragmentation in the HCD cell at a normalized CE of 30. The MS2 AGC was set to 5e4 target ions and a max fill time of 22ms. DIA was performed using 4 Da (150 scan events, 6.46 second cycle time with 1% standard deviation), 6 Da (100 scan events), or 12 Da (50 scan events) windows over the precursor mass range of 400-1000 m/z and the MS2 mass range was set from 100-1500 m/z. Multiplexed precursor mass window DIA acquisitions were acquired using the approach outlined in *Sidoli et al.*,[37] using 4 Da and 6 Da precursor mass windows outlined above.

The 6600 TripleTOF (Sciex) was connected to Eksigent 415 LC system that was operated in micro-flow mode. The mobile phase A was comprised of 0.1% aqueous formic acid and mobile phase B was 0.1% formic acid in acetonitrile. Peptides were pre-loaded onto the trap column (ChromXP C18CL 10 x 0.3 mm 5 µm 120Å) at a flow rate of 10 µL/min for 3 min and separated on the analytical column (ChromXP C18CL 150 x 0.3mm 3 µm 120Å) at a flow rate of 5 µL/min using a linear A-B gradient composed of 3-

35% A for 60 min, 35-85% B for 2 min, then and isocratic hold at 85% for 5 min with re-equilibrating at 3% A for 7 min. Temperature was set to 30°C for the analytical column. Source parameters were set to the following values: Gas 1 = 15, Gas 2 = 20, Curtain Gas = 25, Source temp = 100, and Voltage = 5500V.

DDA MS1 scans were acquired at 45,000 FWHM, using a dwell time of 250 ms in the mass range of 400-1250 m/z of 100 counts per second were selected for fragmentation. DDA MS2 scans were acquired in high-sensitivity mode at 15,000 FWHM with dynamic accumulation with a dwell time of 25 ms for ions ranging from +2 to +5 using rolling collision energy and a collision energy spread of 5. Ions were excluded for fragmentation after one occurrence for a duration of 15 seconds.

DIA MS1 scans were acquired using a dwell time of 250 ms in the mass range of 400-1250 m/z at 45,000 FWHM. DIA MS2 scans were acquired in high-sensitivity mode at 15,000 FWHM over the precursor range of 400-1250 m/z with the MS2 range of 100-1800 m/z using 100 variable windows with a dwell time of 30ms. Additionally, the same samples were acquired in DIA using 4 Da fixed windows over a precursor range of 400-1000 m/z with a dwell time of 22ms.

Methyl-Enrichment Sample Preparation: A cell pellet (~1 ml volume) of a hepatoma cell line derived from a *Gnmt* -/- mouse[35] was resuspended in 4 ml urea lysis buffer (8 M urea, 100 mM NaCl, 50 mM Tris pH 8) supplemented with 1X protease inhibitors (Halt, Thermo) and sonicated on ice with a probe sonicator (Fisher Sonic Dismembrator Model 100, setting 2, 3 x 15 sec). Whole cell extract was then cleared by centrifugation for 10 mins at 4°C and 14,000 rpm (15,996 x g). A total of 10 mg of protein was reduced with 10 mM dithiothreitol at 56°C for 30 mins and alkylated with 50 mM iodoacetamide for 40 mins at room temperature in the dark. This incubation protocol is unique to this set of samples, while the rest of the study's samples were subjected to reduction at 37°C. After adding 4 volumes of 25 mM Tris pH 8, proteins were digested with trypsin (1/200 ratio by mass) overnight at 37°C. The digest was

acidified by addition of trifluoroacetic acid (TFA) to 1% and then desalted with SepPak cartridges (3cc, 200 mg). Dried peptides were resuspended in 1 ml of phosphate-buffered saline (PBS) and dibasic sodium phosphate was added to neutralize the pH (~ 50 mM final concentration). Insoluble debris was pelleted by centrifugation for 10 mins at 17,000 x g at room temperature and the supernatant was used for immunoprecipitation of methylated peptides. Peptide concentration was estimated based on UV absorbance.

Anti-mono-methyl lysine and anti-di-methyl lysine antibodies were produced by Proteintech Group Inc. using a synthetic peptide library immunogen consisting of CX6KX6, where X is any amino acid except cysteine and the central lysine is unmodified, mono-methylated, or di-methylated. The provided anti-serum was further affinity purified in house. A total of 50 µg of each antibody per 1 mg of input peptides were combined and incubated with magnetic protein A beads (GE, 70 µl slurry per 112 µg antibody) for several hours on a rotator at 4°C. The charged beads were washed 4 x 1 ml with PBS and then incubated with 1 ml (~3.8 mg) of peptides overnight at 4°C with rotation. The beads were washed 1 x 1 ml with PBS, 1 x 1 ml with PBS supplemented with 0.8 M NaCl, 1 x 1 ml with PBS, and 1 x 1 ml with water. Captured peptides were eluted by incubation with 100 µl of 1% TFA for 5 mins at room temperature and then desalted with C18 stage tips prior to analysis by LC-MS/MS.

Methyl-Enrichment MS Acquisition: Peptides were analyzed by LC-MS/MS using a nano EasyLC 1000 system (Thermo) fitted with a fused silica column (Polymicro Tech, 75 µm i.d. x 20 cm) packed with ReproSil-Pur 120 C18-AQ (3 µm, Dr. Maisch GmbH) and positioned in line with an Orbitrap Fusion mass spectrometer (Thermo). The chromatography run consisted of a 120 min gradient increasing from 2% to 40% solvent B over 105 mins, 40% to 80% solvent B over 15 mins, and then holding at 80% solvent B for 10 mins. Water containing 0.1% formic acid and acetonitrile containing 0.1% formic acid served as solvents A and B, respectively. The mass spectrometer was programmed to perform MS1 scans in positive profile mode in the Orbitrap at 120,000 resolution covering a range of 300-1500 m/z with a

maximum injection time of 100 ms and AGC target of 1e6. MS2 scans on the most intense ions, fragmented by HCD, were performed in the Orbitrap in positive centroid mode at 15,000 resolution with a maximum injection time of 200 ms, target of 1e5, and NCE of 27. Quadrupole isolation was enabled with an isolation window of 2 m/z. Dynamic exclusion was set to 30 sec. MS2 scans were performed over a 3 sec cycle time.

Acetyl and Succinyl Enrichment Sample Preparation: Mitochondrial extracts and total protein lysate were prepared from mouse tissue as previously described.[28] Briefly, crude mitochondrial fractions were isolated from the livers of SIRT5-/- (C57BL/6) mice by differential centrifugation. Total protein lysate was also isolated from the livers of WT (C57BL/6) mice. Protein concentration of mitochondrial extracts and total protein lysate was determined by BCA assay and 1 mg of protein per sample were brought to equal volumes in 8M urea in 50 mM TEAB buffer and vortexed for 10 min. Samples were reduced with 20 mM Dithiothreitol (DTT) for 30 min at 37C, alkylated in 40 mM iodoacetamide for 30 min at room temperature, diluted 10-fold in 50mM TEAB, and digested overnight at 37C with trypsin used at 1:50 enzyme protein. Digestion was quenched with formic acid and samples were desalted with Oasis HLB 10 mg Sorbent Cartridges, vacuum concentrated to dryness, and resuspended in 1.4mL immunoaffinity purification (IAP) buffer. Immunoaffinity enrichments of modified peptides were done using the PTMScan® Acetyl-Lysine Motif [Ac-K] Kit (#13416) and PTMScan® Succinyl-Lysine Motif [Succ-K] Kit (#13764) (Cell Signaling Technology, Danvers, MA), using the 'one-pot' affinity enrichment previously described.[28] For one-pot enrichments, 1 mg peptide digests were incubated in a tube containing equal parts (~62.5 µg immobilized antibody) of both the succinyl- and acetyl-lysine antibodies.

Acetyl and Succinyl Enrichment MS Acquisition: Samples were analyzed by reverse-phase HPLC-ESI-MS/MS using the Eksigent Ultra Plus nana-LC 2D HPLC system (Dublin, CA) combined with a cHiPLC System, and directly connected to a quadrupole time-of-flight SCIEX TripleTOF 5600 or a TripleTOF 6600 mass spectrometer (SCIEX, Redwood City, CA). Typically, mass resolution in precursor scans was ~35,000 (TripleTOF 5600) or 45,000 (TripleTOF 6600), while fragment ion resolution was

~15,000 in 'high sensitivity' product ion scan mode. Injected peptide mixtures initially flow into a C18 pre-column chip (200 µm x 6 mm ChromXP C18-CL chip, 3 µm, 300 Å, SCIEX) and washed at 2 µl/min for 10 min with the loading solvent (H2O/0.1% formic acid) for desalting. Subsequently, peptides flow to the 75 µm x 15 cm ChromXP C18-CL chip, 3 µm, 300 Å, (SCIEX), and eluted at a flow rate of 300 nL/min using a 3 or 4 hr gradient using aqueous and acetonitrile solvent buffers.

For the creation of a spectral library for analysis by the library-based DIA workflow, data-dependent acquisitions (DDA) were carried out on all PTM enrichments to obtain MS/MS spectra for the 30 most abundant precursor ions (100 ms per MS/MS) following each survey MS1 scan (250 ms), yielding a total cycle time of 3.3 sec. For collision induced dissociation tandem mass spectrometry (CID-MS/MS), the mass window for precursor ion selection of the quadrupole mass analyzer was set to ± 1 m/z using the Analyst 1.7 (build 96) software.

Generation of Spectral Libraries and Database Searches: Assay libraries were generated as described by *Parker et al.* 2016[38] with some modifications. Briefly, DDA files were converted to mzXML and searched through the Trans Proteomic Pipeline (TPP) using 3 algorithms, (1) Comet;[39] (2) X!tandem! Native scoring;[40] and (3) X!tandem! K-scoring[41] against a reviewed, mouse canonical protein sequence database, downloaded from the Uniprot database on January 24th, 2019, containing 17,002 target proteins and 17,002 randomized decoy proteins. Precursor and fragment mass tolerance for all of these search algorithms were set to 10ppm. Peptide probability modeling was performed using the TPP peptide prophet "xinteract" and the results searches were combined using the TPP "interprophet parser". Further filtering was done using Mayu to select peptide spectral match probability values consistent with a 1% peptide false discovery rate (FDR) and a spectral library was generated using the TPP SpectraST tool. Retention times were then normalized to 'indexed' retention time space[42] using the custom python script spectrast2spectrast_irt.py, publicly available via the MSPROTEOMICSTOOLS python package (https://github.com/msproteomicstools/msproteomicstools). Biognosys internal retention time reference

peptides were added to each sample immediately before acquisition and were used for retention time (RT) alignment. As needed, the expanded Common internal Retention Time standards reference peptides[43] was used to align RT in fractionated library samples, removing any outliers as described in *Parker et al. 2016*.[38,42] Retention time normalized splib files were consolidated to a single spectrum entry for each modified peptide sequence using the "consensus spectrum" function in Spectrast. These spectral libraries were then filtered to generate peptide assay libraries by selecting the top twenty most intense b or y ion fragments. The resulting file was imported into Skyline for DIA analysis as described in *Egertson et al.*[44] and was simultaneously converted to the OpenSWATH library input TraML format with decoys appended. The Skyline and OpenSWATH assay libraries normalized to 'indexed' retention time space and all raw files used in this study will be available at Panorama (https://panoramaweb.org/methylation_methods_1.url, Proteome Exchange ID: PXD012621) as a public resource to the proteomics community.

Quantitation of Individual Specimen by DIA-MS: Peak group extraction and FDR analysis was done as outlined in *Parker et al, 2016.*[38] Briefly, raw intensity data for peptide fragments was extracted from DIA files using the open source OpenSWATH workflow[45] against the sample specific peptide assay. Then, retention time prediction was made using the Biognosys iRT Standards spiked into each sample. Target and decoy peptides were then extracted. scored and analyzed using the mProphet algorithm[46] to determine scoring cut-offs consistent with 1% FDR. Peak group extraction data from each DIA file was combined using the 'feature alignment' script, which performs data alignment and modeling analysis across an experimental dataset.[47] Finally, all duplicate peptides were removed from the dataset to ensure that peptide sequences are proteotypic to a given protein in our FASTA database.

Site Localization Scoring in DIA-MS

We performed five individual data searches with Thesaurus search engine (version 0.9.4)[10] presetting a different modification (acetylation, succinylation, mono-methylation, di-methylation or tri-methylation)

in each analysis. For the Thesaurus searches, we used a filtered (modified peptidoforms only) version of the PTM enriched DIA Assay Library generated in this study. Localization strategy parameter was set to recalibrated (peak width only). Library precursor and fragment mass tolerance were both set to 10 ppm. Peptidoforms were selected for false localization rate (FLR or localization FDR) analysis if they passed user-specified localization p-value ($p>0.05$) and localization score (>2) thresholds. All the detected peptidoforms in the individually processed DIA-MS RAW files (at 5% global peptide FDR) were posteriorly integrated and processed with Percolator 3.01 (with the global peptide FDR filter set at 1%). Once all five modification-specific analyses were generated, five output results files (.elib) were integrated into one output and processed with Percolator (at 1% global peptide FDR filter). Localization p-values for passing peptidoforms are independently corrected using Benjamini-Hochberg FDR test (localization FDR). Peptidoforms meeting FLR <1% criteria were used for further biological analyses.

Normalization and Quantitation: The total ion current (TIC) associated with the MS2 signal across the chromatogram was calculated for normalization using in-house software. This 'MS2 Signal' of each file was used to adjust the transition intensity of each peptide in a corresponding file. To obtain a measure of total protein quantity, all modified (methylated, acetylated, and succinylated) peptides were removed and the remaining peptides analyzed separately. Normalized transition-level data of the unmodified peptides were subsequently processed using mapDIA to obtain protein level quantitation.[48] Normalized transition-level data of the all modified (methylated, acetylated, and succinylated) peptides were then analyzed individually through mapDIA to obtain peptide level quantitation of PTM containing peptides. Finally, to obtain a PTM/Total ratio, the intensity of the PTM containing peptide was divided by the intensity of its respective protein, calculated using mapDIA from an aggregate of all unmodified peptides for that protein.

Data Visualization and Validation: The aligned OpenSWATH output was imported into Skyline[44] which was used to visualize and manually validate methylated peptides. The Skyline documents

containing DIA acquisitions extracted against their respective peptide assay libraries are available at Panorama (https://panoramaweb.org/methylation_methods_1.url, Proteome Exchange ID: PXD012621).

Metabolomic Analysis: Liver metabolic profiles were semi quantified using an ultra-high performance liquid chromatography (UHPLC)-Time of Flight-MS based platform as described previously in *Alonso et al. 2017*.[30] This platform provided coverage over polar metabolites, such as vitamins, nucleosides, nucleotides, carboxylic acids, coenzyme-A derivatives, carbohydrate precursors/derivatives, and redox-electron-carriers was used. Briefly, proteins were precipitated from the liver tissue (15 mg) by adding methanol spiked with metabolites not detected in unspiked cell extracts (internal standards) and samples were homogenized using a Precellys 24 homogenizer (Bertin Technologies, Montigny-le-Bretonneux, France) at 6500 rpm for 23 seconds x 1 round, and centrifuged at 18,000 x g for 10 minutes at 4 ºC. Then, 500 µl were collected, mixed with chloroform and vortexed. After 10 minutes of agitation, samples were centrifuged at 18,000 x g for 15 minutes at 4 ºC. Supernatants were collected, dried under vacuum, reconstituted in water and resuspended with agitation for 15 minutes. After centrifugation at 18,000 x g for 5 minutes at 4 ºC, samples were transferred to plates for UHPLC-MS analysis as described in *Barbier-Torres et al. 2015*.[49]

Data were pre-processed using the TargetLynx application manager for MassLynx 4.1 software (Waters Corp., Milford, USA). Metabolites were identified prior to the analysis. Peak detection, noise reduction and data normalization were performed as previously described in *Martinez-Arranz et al. 2015*.[50]

Results

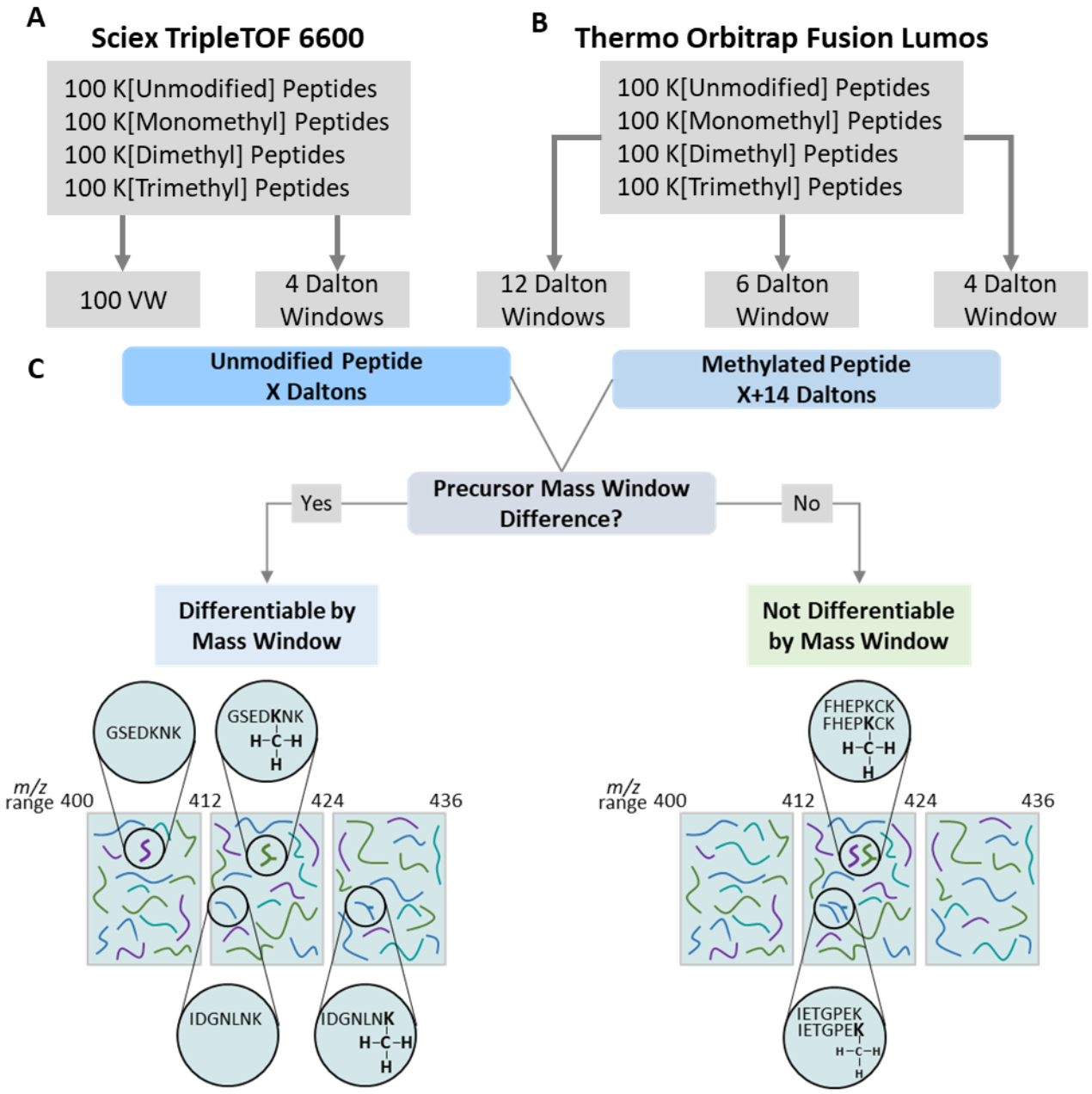


Figure 1. Schematic Overview of Experimental Design:
(A) An equimolar pool of 400 synthesized peptides containing K[Unmodified], K[Monomethyl], K[Dimethyl], and K[Trimethyl] residues were acquired in DIA on the Sciex TripleTOF 6600 with 4 Dalton and 100 Variable Window (VW) precursor mass windows. (B) Same peptide pool was acquired in DIA on the Thermo Orbitrap Fusion Lumos with 4, 6, and 12 Dalton precursor mass windows. (C) Strategy to differentiate an unmodified peptide from it's methylated form based off of precursor mass window difference. An unmodified precursor will be fragmented in a different mass window than its methylated form making it differentiable by precursor mass window.

Synthesized Methylated Peptides

The overall experimental design for quantifying protein methylation is shown in Figure 1. Although we initially focus on Lys methylation to develop the workflow, this approach is applicable to Arg methylation or any other methylated residue. To establish DIA-MS parameters to allow for differentiating a methylated peptide from an unmodified peptide, first a DDA-MS library was built on a Sciex 6600 Triple TOF based on the DDA acquisition of a pool of 400 synthesized peptides (400 femtomoles/peptide) containing either K[Unmodified], K[Monomethyl], K[Dimethyl], or K[Trimethyl]. DIA acquisitions of 400 femtomoles/peptide of the methyl peptide pool were then acquired using a 4 Dalton precursor mass windows and 100 variable mass windows (Figure 1A). A DDA library was also created from the methyl peptide pool acquired on the Orbitrap Fusion Lumos (200 femtomoles/peptide). DIA acquisitions of 200 femtomoles/peptide were then acquired with 4, 6, and 12 Dalton fixed precursor mass windows respectively (Figure 1B). Generally, the unmodified peptide can be physically separated from its methylated form due to precursor mass difference and thus its fragmentation in a different precursor mass window allowing for differentiation of the unmodified and methyl peptide by precursor mass window alone. However, this is not always the case and depending on precursor mass window size and the charge on the peptide the methylated peptide will fragment in the same precursor mass window as its unmodified form. In this case, a site-specific transition or a co-eluting precursor trace is necessary to correctly identify the post-translational modification (Figure 1C, Table S1). A difference in retention time cannot reliably be used to differentiate an unmodified peptide and its methyl peptidoform as 58.5% of monomethyl peptides elute within 30 seconds of each other in a 60 minute gradient while 92.7% elute within 60 seconds making separation through LC alone unpractical for DIA-MS workflows (Figure S1). Due to this, we next looked into methods to physically separate peptidoforms using precursor mass window widths.

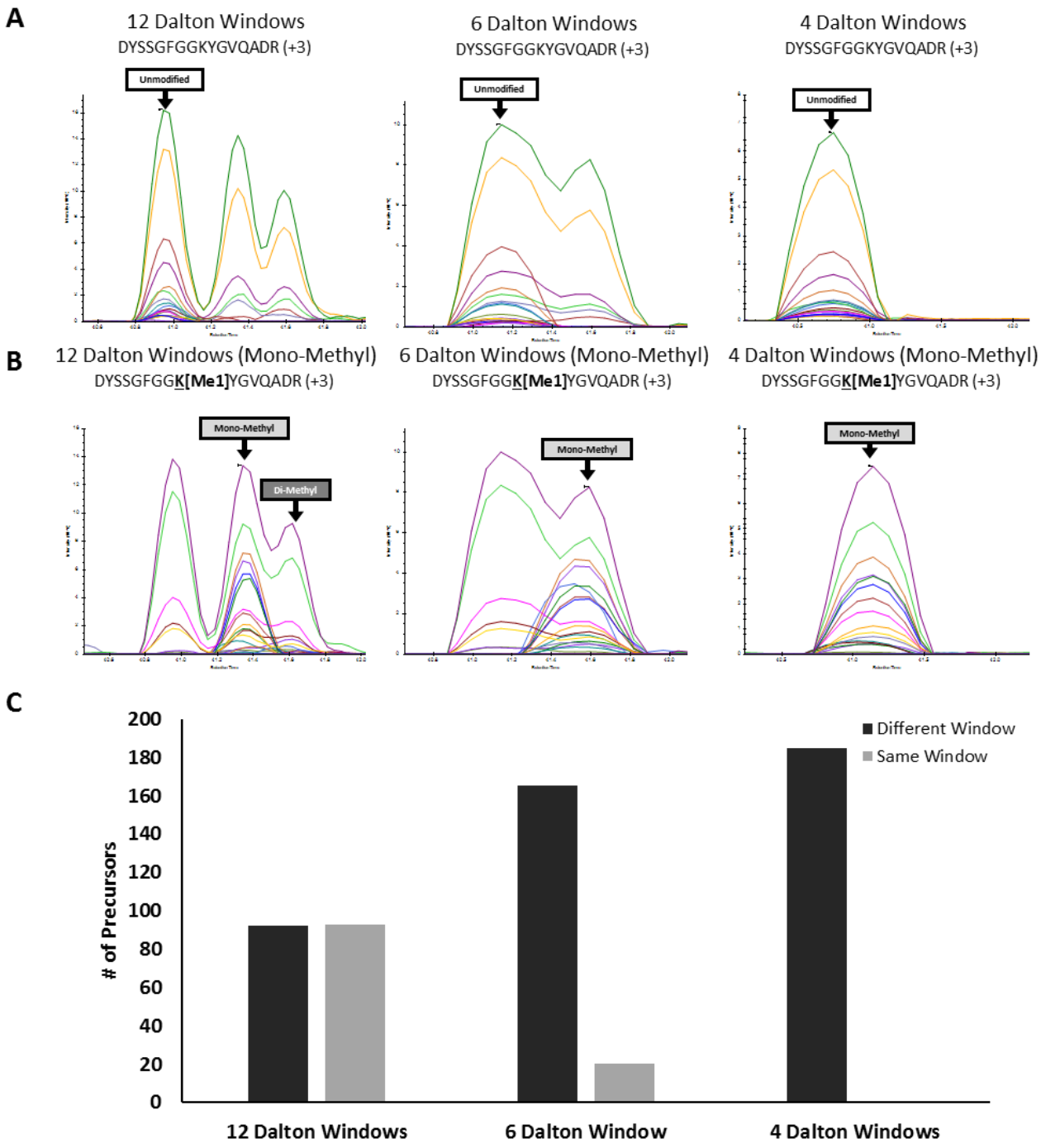


Figure 2. Synthesized Methylated Peptides: Presence of synthesized methylated and unmodified peptides acquired on the Thermo Orbitrap Fusion Lumos with 4, 6, and 12 Dalton precursor mass windows were visualized in Skyline (A) Quantified transitions of DYSSGFGGKYGVQADR (B) Quantified transitions of DYSSGFGG**K[Me1]**YGVQADR (C) Graphical representation of unmodified and monomethyl +2 and +3 precursors which can be differentiated from their unmodified forms using 4, 6, and 12 Dalton precursor mass windows. Black represents peptides eluting in different precursor mass windows while grey represents peptides co-eluting in the same precursor mass window.

First, we tested whether narrowing the MS1-precursor isolation window width affected the separation of different methylated peptidoforms from their unmethylated counterparts. For example, when the methyl peptide pool was acquired on the Orbitrap Lumos Fusion, we can see that with 12 Dalton precursor mass windows three species of the same peptide have co-fragmented in the same precursor mass window. With 6 Dalton precursor mass windows there are two species of peptide which have co-fragmented in the same window. With 4 Dalton windows, the precursor mass window is small enough that there is only one species of the peptide found (Figure 2 A, B). With the smallest acquisition mass window schema (4 Da) 100% of methylated peptides were separated from their corresponding unmodified forms. Increasing the mass window just 2 Da reduced the segregation slightly, and increasing mass window to 12 Da resulted in only ~ 50% segregation of methylated peptides from unmethylated. (Figure 2C). The 4 Dalton precursor mass windows separate 100% of both +2 and the smaller m/z +3 charged precursors from their methylated forms, whereas the commonly used 100 variable precursor mass window acquisition schema has multiple cases where 2 or more species of peptide co-fragment in the same window (Figure S2, Table S2).

In-vivo PTM DIA Assay Library

Moving next to a complex cellular lysate (liver). Two mouse models of NASH where employed, each chosen due to their extreme methylation potential. *Gnmt* -/- mice have increased SAMe and are biologically hypermethylated while *Mat1a* -/- mice are SAMe deficient and therefore hypomethylated.[30,31] First, we created a hypermethylated DIA assay library by DDA acquisitions of 5 SCX fractions from *Gnmt* -/- mouse livers. We supplemented this library with a DDA acquisition of a methyl-Lys immunoprecipitation from a hepatoma cell line derived from *Gnmt* -/- mice.[35] This gave us an assay library of 808 methyl Arg and Lys peptides. We further supplemented the library with DDA acquisitions of a 'one-pot' acetyl-Lys and succinyl-Lys immunoaffinity enrichment of mitochondrial extract and whole liver lysate isolated from WT mouse liver as described in *Basisty et al*[28] which provided us with 2428 succinylated peptides and 2547 acetylated peptides.

Accuracy of Peptide Quantitation in 4 Dalton DIA Method

To establish the accuracy the peptide quantitation obtained from our 4 Da precursor mass window DIA workflow, we performed a dilution series of 71 SIL peptides in complex liver lysate using both 60- and 120-minute gradients. Observation of enough MS2 spectra across the chromatographic elution profile, also termed 'points across the peak' is critical for estimating the area-under-the-curve (AUC) of intensity and elution time, and thus, the accuracy of the quantitative data for a given peptide. This is especially true for low abundant peptides which may contain PTMs. The more windows sampled, the more time spent in a given duty cycle, potentially sacrificing quantitative accuracy. The 4 Da DIA method used in this study has a median cycle time of 6.46 seconds with 1.0% standard deviation. We determined that the 4 Da precursor mass windows can detect an average of 6.94 points across a peak in a 60 minute LC gradient while expanding the LC gradient to 120 minutes, and thus widening chromatographic peaks, detected an average of 10.24 points across the peak (Figure S3 A,B). We next estimated the accuracy of quantification by examining linearity of SIL peptide AUC across the dilution series. We next determined the accuracy of quantitation using both 60- and 120-minute gradients by quantifying the SIL peptides in our dilution series. We have linearity in quantification ($r^2 > 0.95$) of 87/91 (95.6%) of transitions corresponding to precursors for the SIL peptides in the 60-minute gradient and linearity in quantification ($r^2 > 0.95$) for 91/91 (100%) of the transitions corresponding to precursors for the SIL peptides in the 120-minute gradient. We additionally looked at the lower end of our dilution curve (7 fmol to 100 fmol on column) to see how our method performs on lower abundant peptides, which are a better representation of the biological concentration of PTMs in complex lysate. We found a ($r^2 > 0.95$) for 84/91 (92.3%) and 85/91 (93.4%) of the transitions corresponding to SIL precursors for the 60-minute gradient and the 120-minute gradient respectively. Furthermore, we performed a dilution series of the same 71 SIL peptides in complex liver lysate with 12 Da precursor mass windows in a 120-minute gradient and found linearity in quantification ($r^2 > 0.95$) of 87/91 (95.6%) of transitions corresponding to precursors for the SIL peptides. Unexpectedly, looking at the lower end of our dilution curve for 12 Da precursor mass window acquisitions, we found only 59/91 (64.8%) of the transitions corresponding to SIL precursors show

linearity in quantification showing that our 4 Dalton precursor mass window DIA acquisition method provides more accurate quantitation of lower abundant peptides than a 12 Dalton precursor mass window method (Figure S3C, Table S3).

Performance of Alternate DIA Data Analysis Methods

MS1 centric workflows like DIA-Umpire[51] and the "DirectDIA" function in Spectronaut (Biognosys) use of MS1 masses in a "spectral-centric" workflow have the ability to differentiate a modified peptide from an unmodified peptide by the precursor mass alone. However, these tools rely on a coeluting precursor trace to accurately identify a peptide. Using the SIL peptides dilution series to mimic low abundant peptides in complex liver lysate and visualizing the results in Skyline we found cases whereas we lowered the concentration of SIL peptides, a coeluting precursor trace was lost while transitions remained. Although 100% of SIL peptides had a coeluting precursor trace at 2250 femtomoles (fm) on column in both the 60- and 120-minute gradient, as the concentration of SIL peptides decreased so did their coeluting precursor traces. When we approached 7 fm on column, only 36% had coeluting precursor traces in the 60 min gradient and 46% had coeluting precursor traces in 120 minute gradient (Figure S3 E,F). Accordingly, using MS1 traces of SIL peptides to determine linearity of their quantitation, we found a sharp decrease in accurate quantitation looking at the lower end of our dilution curve in 4 Da and 12 Da DIA acquisition methods (Figure S3D, Table S3). Furthermore, looking at endogenous modified peptides in these same acquisitions from our DIA assay library, methylated, acetylated, or succinylated, we found 60.4% do not have a coeluting precursor trace, making them undetectable by "MS1-centric" approaches and in the case of methylated peptides undifferentiable from their unmodified peptidoform without 4 Da precursor mass windows (Figure S4 A,B). To further this point, we next analyzed DIA acquisitions of complex liver lysate using different data analysis software. We chose to use Spectranaut's software suite as Spectranaut has the ability to analyze DIA data with a targeted library based method as well as DirectDIA, a MS1-centric library free approach. Assaying complex liver lysate DIA acquisitions against our PTM enriched DIA assay library resulted in quantitation of 188 endogenous modified peptides from

our DIA assay library; methylated, acetylated, or succinylated. DirectDIA analysis of the same DIA acquisitions resulted in 28 endogenous modified peptides. We also assayed the same complex liver lysate acquired with 12 Da windows using DirectDIA which resulted in the 35 endogenous PTM containing peptides (Figure S4C, Table S4). Instead of relying on a coeluting precursor mass to identify a modified peptide and differentiate a methylated peptide from unmodified peptide, our method physically separates the unmodified and methylated peptides by acquiring their precursors into separate mass windows. In addition, using a PTM enriched DIA assay library, our method allows for the quantitation of low abundant modified peptides without a MS1 trace.

Total Protein Normalization Allows for More Accurate Quantification of Methylation in NASH

After extracting 4 Dalton precursor mass window DIA acquisition maps of *Gnmt* -/- and *Mat1a* -/- mouse liver cellular lysate against the hypermethylated DIA assay library, we quantified 78 methylated peptides in OpenSWATH with unmodified peptides quantifiable for the same protein. To determine the relative abundance of PTMs, having total protein quantification is required. Since we are using a DIA assay library containing both methylated and unmodified peptides extracted against an unenriched complex lysate, we have the ability to quantify both methylated Arg and Lys containing peptides and unmodified peptides in a single DIA acquisition. For example, a unique mono-methylated peptide corresponding to EF1A1 K165(mono-methyl) was found to be significantly increased in the hypermethylated *Gnmt* -/- NASH model compared to WT ($p<0.0005$), whereas the same site was significantly decreased in the hypomethylated *Mat1a* -/- NASH model compared to WT ($p<0.05$) (Figure 3A). After normalizing to total protein, by taking the ratio of the intensity of the methylated peptide compared to the intensity of an aggregate of all unmodified EF1A1 peptides as determined by MapDIA, *Mat1a* -/- hypomethylation compared to WT was more significant ($p<0.0005$) (Figure 3B). By being able to normalize to the aggregate protein intensity derived from unmodified peptides in the same MS acquisition we have the ability to look at Methyl/Total ratio which allows for easy normalization of the methylated peptide to the aggregate protein abundance in one DIA acquisition, providing similar data as a protein immunoblot for a

PTM and total protein, which can further the biological insight that one can take away from this type of data.

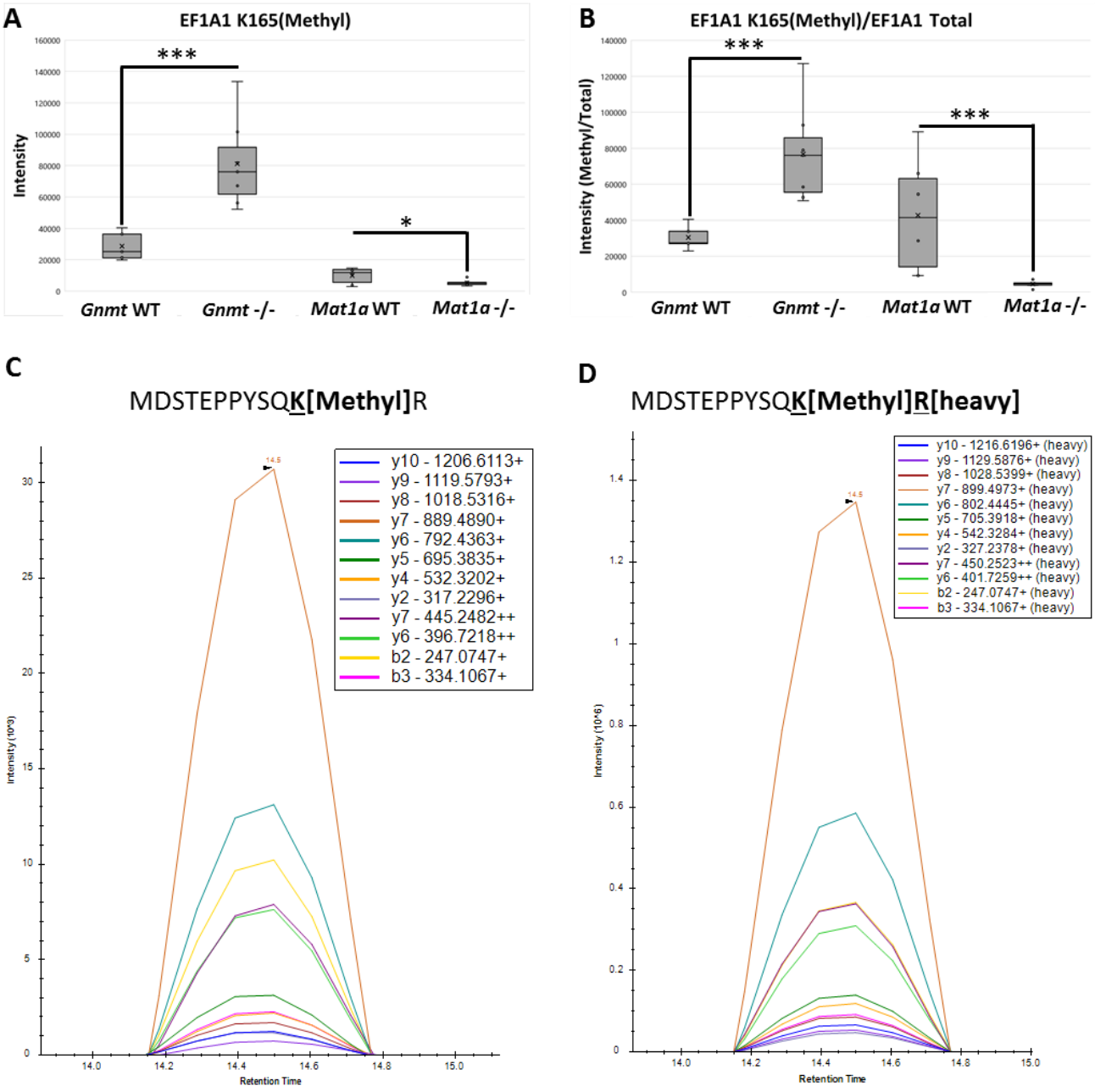


Figure 3. *In-vivo* Methylation: (A) Quantification of the MS2 TIC normalized intensity of peptide containing EF1A1 K[165] Mono-Methyl in *Gnmt* -/- and wild-type littermates and *Mat1a* -/- and wild-type littermates Data are box and whisker plots of six biological replicates. Two-tailed Student's t-test, $^{***}P < 0.005$, $^{*}P < 0.05$ (B) Quantification of the intensity of the MS2 TIC normalized peptide containing EF1A1 K[165] Mono-Methyl divided by the intensity of EF1A1 total protein, determined by EF1A1 unmodified peptides using MAPDIA in *Gnmt* -/- and wild-type littermates and *Mat1a* -/- and wild-type littermates Data are box and whisker plots of six biological replicates per condition. Two-tailed Student's t-test, $^{***}P < 0.0005$ (C) Skyline Visualization of the XIC of the peptide MDSTEPPYSQK[Methyl]R which corresponds to EF1A1 K[165] Mono-Methyl in *Gnmt* -/- complex liver lysate (D) Skyline Visualization of the XIC of a stably isotopically labeled peptide MDSTEPPYSQK[Methyl]R[heavy] which corresponds to EF1A1 K[165] Mono-Methyl spiked into *Gnmt* -/- complex liver lysate

To validate our methyl library, we synthesized heavy methyl-peptides to match three proteins localized in different cellular compartments which were present in complex liver lysate from *Gnmt* -/- mice with NASH. Next, these heavy peptides were added to the *Gnmt* -/- complex liver lysate and 4 Dalton precursor mass window DIA-MS was carried out. We found co-eluting peaks from all of the endogenous methyl-peptides and their corresponding synthetic heavy peptides (Figure 3D, Figure S5). From our SCX fractionated peptide assay library in this study, 53.7% of unmodified peptide precursors are +2 and 38.5% of precursors are +3. However, looking at only methylated precursors, only 22.2% are +2 while 58.1% of the methylated precursors are +3 (Figure S6, Table S5). This is likely due to more missed cleavages due to methylation inhibiting trypsin's ability creating longer peptides with a higher charge.[52] We can differentiate 80.3% of methylated precursors in our library from their unmodified forms due to precursor charge state of +2 and +3. When looking further into the 19.7% of peptides that remain with a charge of +4 or higher and calculating the respective precursor mass of the methylated peptide and its corresponding unmodified peptide, we can differentiate 99.1% of the remaining methylated peptides from their unmodified form by which mass window each precursor falls in leaving 0.24% (2/808) unresolvable by 4 Dalton precursor mass windows (Table S5).

Ambiguous Methylation Quantitation in DIA-MS

Another important consideration is the extent of ambiguous transitions and their impact on methyl-peptide quantitation. In other words: peptides that are in the same precursor mass window as their unmodified form, and where transitions from the peak of a co-eluting unmodified peptide are being assigned as methylated. Looking at complex cellular liver lysate from *Gnmt* -/- mice acquired with 12 Da precursor mass windows, we have an example of an ambiguous identification; the same peak is being called for the methylated peptide and the unmodified peptide with a single site-specific transition (b7) for the methyl-site identified (Figure 4A,B). Looking at the same sample acquired with 4 Da precursor mass windows, although we can still detect the unmodified peptide, we no longer can detect the methylated peptide, leading us to believe that the quantified intensity of the methylated peptide with 12 Da windows

is incorrect (Figure 4C,D). Additionally, due to this co-eluting site-specific transition, FLR algorithms will incorrectly assign a high probability to this ambiguous modified peptidoforms. 4 Dalton precursor mass windows separate 100% of +2 and +3 charged precursors from their methylated forms leading to zero modified peptides with ambiguous transitions leading to incorrect quantitation in complex lysate while 12 Da precursor mass windows have over 50% of methylated peptides which fall in the same precursor mass window and are likely to contain transitions which will lead to ambiguous FLR assignment and quantitation. (Figure 4E). In addition to ambiguous quantitation, false positive identification also occurs for methyl and dimethyl peptides where the unmodified peak was being called as methylated based on the same criteria but in this case zero site specific transitions are observed (Figure S7).

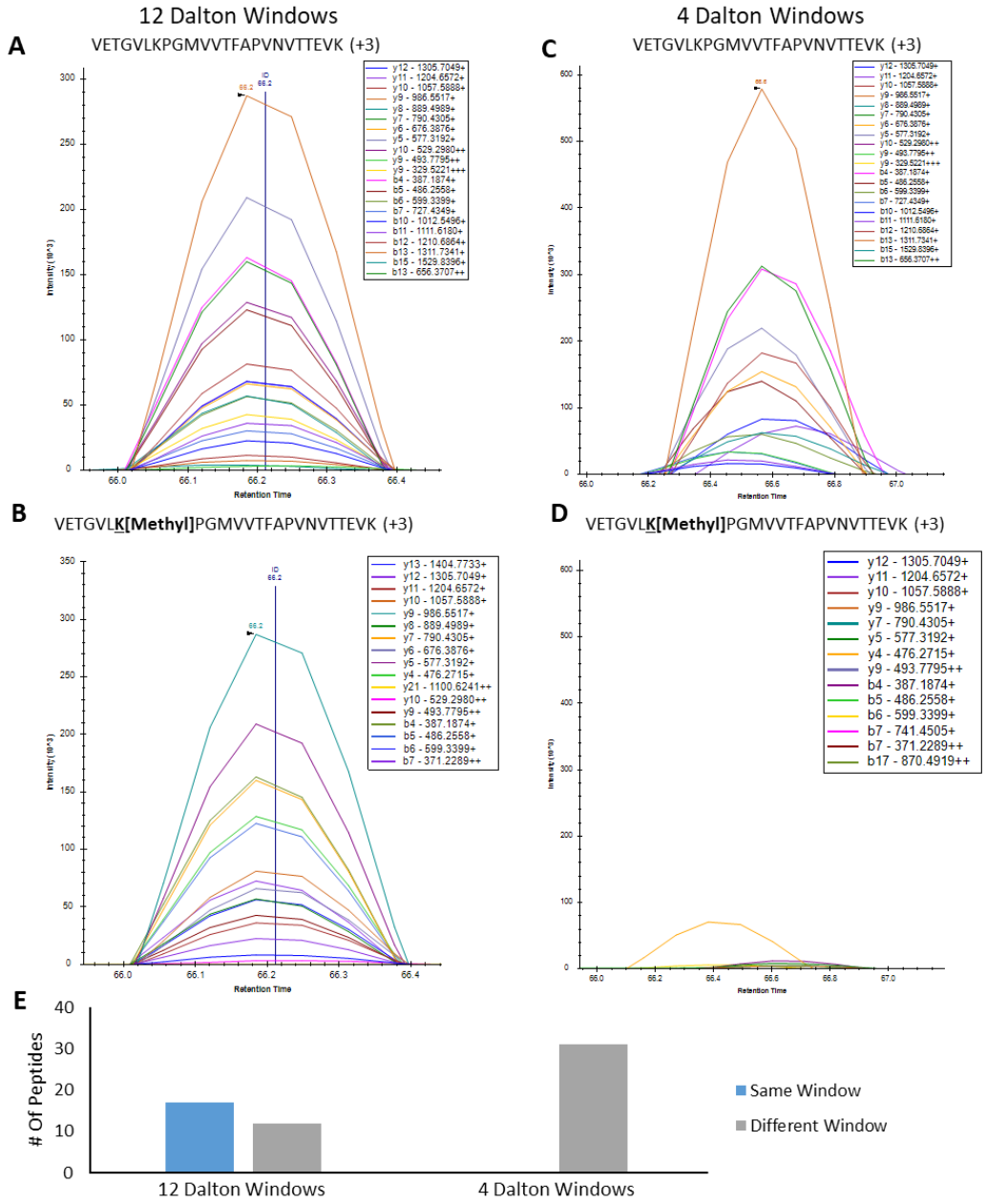


Figure 4. Ambiguous Methylation Quantitation: Complex cellular liver lysate from *Mat1a* -/- acquired on the Thermo Orbitrap Fusion Lumos with 4 and 12 Dalton precursor mass windows visualized in Skyline (A) Quantified transitions of VETGVLKPGMVVTFAPVNVTTEVK acquired with 12 Dalton precursor mass window (B) Quantified transitions of VETGVL**K[Methyl]**PGMVVTFAPVNVTTEVK acquired with 12 Dalton precursor mass windows. This precursor falls in the same mass window as its unmodified form, elutes at the same time, and contains one site specific transition (b7). (C) Quantified transitions of VETGVLKPGMVVTFAPVNVTTEVK acquired with 4 Dalton precursor mass window (D) Quantified transitions of VETGVL**K[Methyl]**PGMVVTFAPVNVTTEVK acquired with 4 Dalton precursor mass windows. There is no discernable peak; leading us to believe that the quantitation of the methylated peptide found with 12 Dalton precursor mass windows is incorrect due to many ambiguous transitions.(E) Graphical representation of unmodified, monomethyl, and dimethyl +2 and +3 precursors which can be differentiated from their unmodified forms using 4 and 12 Dalton precursor mass windows. Grey represents peptides eluting in different precursor mass windows while blue represents peptides co-eluting in the same precursor mass window.

Acetylation and Succinylation are Differentially Changed in NASH Subtypes

Next, looking at acetylation and succinylation, we quantified 176 acetylated and 59 succinylated peptides with OpenSWATH in the DIA-MS runs on the mouse tissue which had unmodified peptides quantifiable of the same protein. As with methylation, to determine relative abundance of the modified peptidoform, having total protein quantification is required. By using a DIA assay library containing both modified and unmodified peptides and assaying against complex lysate, we have the ability to quantify both acetylated and succinylated Lys containing peptides and unmodified peptides in a single DIA acquisition.

A unique succinylated peptide corresponding to mitochondrial malate dehydrogenase (MDHM) K165(Succinyl) was significantly increased in the *Gnmt* -/- NASH model compared to WT ($p<0.005$) while the *Mat1a* -/- NASH model compared to WT was not significantly changed (Figure 5A). When normalizing to total protein, by taking a ratio of the intensity of the succinylated peptide compared to the intensity of unmodified MDHM peptides as determined by MapDIA, *Mat1a* -/- compared to WT became significantly decreased in the *Mat1a* -/- NASH model compared to WT ($p<0.005$) while *Gnmt* -/- NASH model compared to WT remained significant ($p<0.05$) (Figure 5B). Furthermore, a peptide proteotypic for Histone H4 (H4) with three acetyl sites, K8, K12 and K16 respectively was found in the DIA-MS datasets. This peptide was not significantly changed in either the *Gnmt* -/- NASH model compared to WT or the *Mat1a* -/- NASH model compared to WT (Figure 5D). When normalizing to total protein, by taking a ratio of the intensity of the acetylated peptide compared to the intensity of unmodified H4 peptides, *Gnmt* -/- compared to WT became significantly decreased versus WT ($p<0.05$) and *Mat1a* -/- became significantly increased compared to WT ($p<0.005$) (Figure 5E). As with methylation, being able to normalize to the protein intensity derived from unmodified peptides in the same MS acquisition we have the ability to look at Succinyl/Total or Acetyl/Total ratio which provides an additional piece of information, similar to a protein immunoblot for total protein, which allows for a representation of the biological concentration of the PTM in complex lysate (Figure 5F).

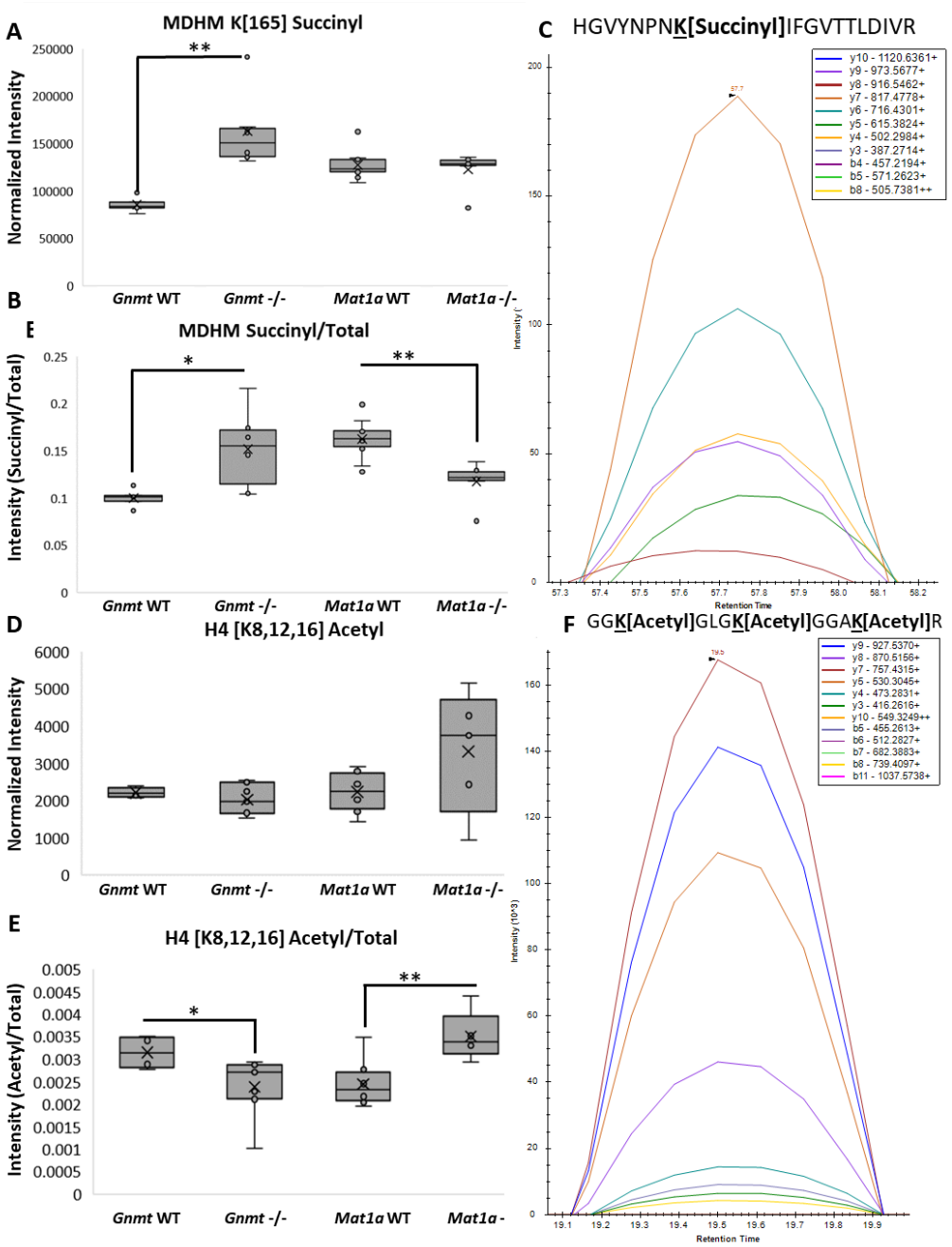


Figure 5. In-vivo Succinylation and Acetylation: (A) Quantification of the MS2 TIC normalized intensity of peptide containing MDHM K[165] Succinyl in *Gnmt* -/- and wild-type littermates and *Mat1a* -/- and wild-type littermates (n=6/condition) Data are box and whisker plots of six biological replicates per condition. Two-tailed Student's t-test, $^{**}P < 0.005$ (B) Quantification of the intensity of the MS2 TIC normalized peptide containing MDHM K[165] Succinyl divided by the intensity of MDHM total protein, determined by MDHM unmodified peptides using MAPDIA. Data are box and whisker plots of six biological replicates per condition. Two-tailed Student's t-test, $^{**}P < 0.005$, $^{*}P < 0.05$ (C) Skyline Visualization of the XIC of the peptide containing MDHM K[165] Succinyl (D) Quantification of the MS2 TIC normalized intensity of a proteotypic peptide mapping to H4 K[8,12,16] Acetyl in *Gnmt* -/- and wild-type littermates and *Mat1a* -/- and wild-type littermates. Data are box and whisker plots of six biological replicates per condition. (E) Quantification of the intensity of the MS2 TIC normalized peptide containing H4 K[8,12,16] Acetyl divided by the intensity of H4 total protein, determined by H4 unmodified peptides using MAPDIA in *Gnmt* -/- and wild-type littermates and *Mat1a* -/- and wild-type littermates. Data are box and whisker plots of six biological replicates per condition. Two-tailed Student's t-test, $^{**}P < 0.005$, $^{*}P < 0.05$ (F) Skyline Visualization of the XIC of the peptide containing H4 K[8,12,16] Acetyl

We confirmed the presence of these succinylated and acetylated peptides by showing the highest abundant transitions found in DDA and DIA acquisitions of 'one-pot' acetyl-Lys and succinyl-Lys immunoaffinity enrichment from WT mouse liver were also found in complex *Gnmt* -/- liver lysate (Figure S8, S9)**.** Furthermore, for the tri-acetylated H4 K[8,12,16] peptide, the relative abundance of these transitions were constant between DIA acquisitions acquired on the Orbitrap Fusion Lumos and the Sciex TripleTOF 6600. These acquisitions were run on instruments located at two different research institutes, and yet yielded the same confident identifications of an acetylated and succinylated peptide respectively.

Modified Peptidoform Localization With 4 Dalton DIA Method

We first tested the ability of Thesaurus (version 0.9.4) to correctly localize modified amino acid residues in a pool of 400 synthesized peptides (400 femtomoles/peptide) containing either K[Monomethyl], K[Dimethyl], K[Trimethyl], or K[Acetyl] on the same a non-terminal Lys residue of 100 peptide backbones. After creating a DDA Library containing the MS2 spectra of synthesized peptides, we performed peptidoform localizations of 4 Da precursor mass window DIA acquisitions using Thesaurus. Out of peptidoforms which were detected with less than 1% global peptide FDR, we were able to localize with less than 1% FLR 62/76 (81.6%) of mono-methylated peptides, 60/77 (78.0%) of di-methylated peptides, 52/65 (80%) of tri-methylated peptides, and 61/76 (80.0%) of acetylated peptides (Figure 6A, Table S6). Overall, 42 peptide backbones were correctly localized with all 4 modifications (Figure 6B) across 3 replicates of DIA-MS acquisitions.

After assessing Thesaurus's performance with acetylated and methylated residues in synthesized peptides DIA-MS dataset, we moved onto PTM localizations in a complex lysate. Using our PTM enriched DIA Assay Library and performing analysis in Thesaurus against the 4 Da precursor mass window DIA acquisitions of complex cellular lysate from *Gnmt* -/- and *Gnmt* WT livers, we were able to identify with >1% global peptide FDR in ≥4/12 (≥33%) of samples and localize with >1% FLR 14 mono-methyl peptides, 161 di-methyl peptides, 28 tri-methyl peptides, 141 acetylated peptides, and 32 succinylated peptides (Figure 6C). The same cohort of samples was assayed against the PTM enriched DIA Assay

Library in OpenSWATH 3.2.0 with >1% FDR, identification in ≥4/12 (≥33%) of samples, and >1% FLR localization in Thesaurus. This resulted in localization and quantitation of 9 mono-methyl peptides, 17 di-methyl peptides, 13 tri-methyl peptides, 76 acetylated peptides, and 14 succinylated peptides (Figure 6C). Additionally, we were able to localize 128/171 (74.6%) of the modified peptidoforms quantified in OpenSWATH with >1% FLR in Thesaurus and found 121 quantifiable peptidoforms in common between Thesaurus and OpenSWATH (Figure 6D).

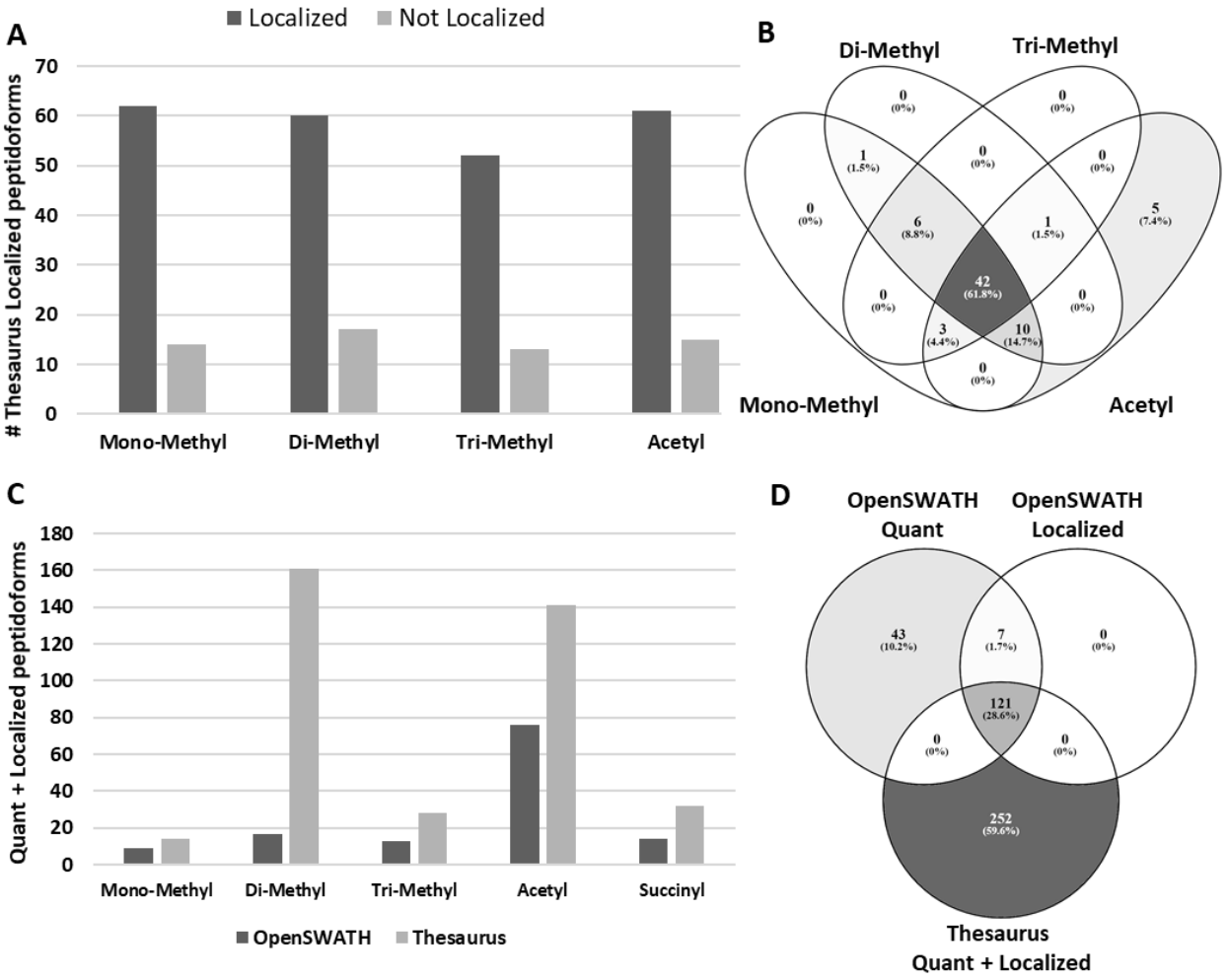


Figure 6. Modified Peptidoform Localization: **(A)** Localization results of modified peptidoforms using Thesaurus 0.9.4 from a mixture of synthesized peptides containing the same peptide backbone but with modifications (mono-, di-, tri-methylated and acetylated) on the same a non-terminal Lys residue. **(B)** Venn diagram showing the overlap of Thesaurus localization results from each modification type from mixture of synthesized peptides **(C)** Quantified and localized modified peptidoform from 4da DIA acquisitions of complex cellular lysate from Gnmt -/- and Gnmt WT livers assayed against the PTM enriched DIA Assay Library in either OpenSWATH 3.2.0 or Thesaurus 0.9.4 and identified with >1% FDR in ≥4/12 (≥33%) of samples in each respective approach in addition to >1% FLR localization calculated in Thesaurus 0.9.4 and applied to both the OpenSWATH and Thesaurus results. **(D)** Venn diagram showing the overlap of quantifiable peptidoforms from OpenSWATH, quantifiable peptidoforms from OpenSWATH which were localized using Thesaurus, and peptidoforms which were both quantifiable and localization from Thesaurus.

Simultaneous Total Protein and PTM Quantification

Total protein quantification was also obtained at the same time as the PTM data because our DIA assay library was built with SCX fractionated complex cellular lysate, containing both modified peptides as well as unmodified peptides. Focusing on the total protein data from *Gnmt* -/- mice with NASH and comparing the 12 Dalton precursor mass window data to the 4 Dalton precursor mass window data, we achieve marginally better results when searching the 4 Dalton precursor mass window data through OpenSWATH. Assaying against the same SCX fractionated DDA library, 5,111 proteins, 26,293 peptides, and 341,241 transitions where identified with our 4 Dalton Precursor mass Windows. With the 12 Dalton Precursor Mass Window acquisitions, we matched 4,595 proteins, 22,782 peptides, and 304,365 transitions. (Figure S10).

Regulatory Role of Protein Methylation During Translation: mRNA Stability in NASH

Two NASH models which knock out Mat1a and Gnmt (*Mat1a* -/- and *Gnmt* -/-, respectively) were analyzed to determine the role that differential methylation potential plays in NASH by manipulating SAMe concentration. Since cellular methylation potential is decreased overall in *Mat1a* -/- and increased in *Gnmt* -/-, we focused on methylated residues and proteins which show this trend. We found differentially methylated peptides in *Mat1a* -/- and *Gnmt* -/- which are associated with protein translation in addition to differential cytosolic and mitochondrial ribosomal protein levels, the protein machinery needed for translation. As an example, there is a cluster of differential methylated peptides at known amino acid residues in *Mat1a* -/- and *Gnmt* -/- which are proteotypic for three RNA binding proteins involved in translation (EF1A1), RNA stability (poly(A)-binding protein (PABP1)), and transcription and RNA stability (Heterogeneous nuclear ribonucleoprotein U (HNRPU)) (Figure 7A). Based on differential methylation patterns of these residues in both mutants we hypothesize that decreased methylation in these proteins will result in decreased messenger RNA stability and subsequent decreased protein translation which should be observed as differentially expressed proteins between *Gnmt* -/- and *Mat1a* -/- NASH. 627 proteins which were decreased in *Mat1a* -/- (FDR < 0.01, log2 FC < -0.32) and 735 proteins which

were significantly increased in *Gnmt* -/- (FDR < 0.01, log2 FC > 0.32) (Table S7). Of these differentially dysregulated proteins we found significant enrichment in: mitochondrial electron transport chain complexes (ETC, OxPhos) (Figure 7C,D), mitochondrial ribosomes (MRPLs) and cytosolic ribosomes (Figure 7B). We also analyzed our previously published microarray data[30,31] from the same models of NASH as a measure for mRNA stability of transcripts from the pathways which were differentially altered at the protein level (Figure S11 A, B). The trends in mRNA levels of the altered pathways (decrease in mRNA in *Mat1a* -/- and increase of mRNA in *Gnmt* -/-) are concordant with our hypothesis that PABP1 methylation effects their mRNA stability. Quantifying methylation and total protein from the same DIA acquisition allows for a much better understanding of how a pathway can be regulated by PTMs from a single experiment.

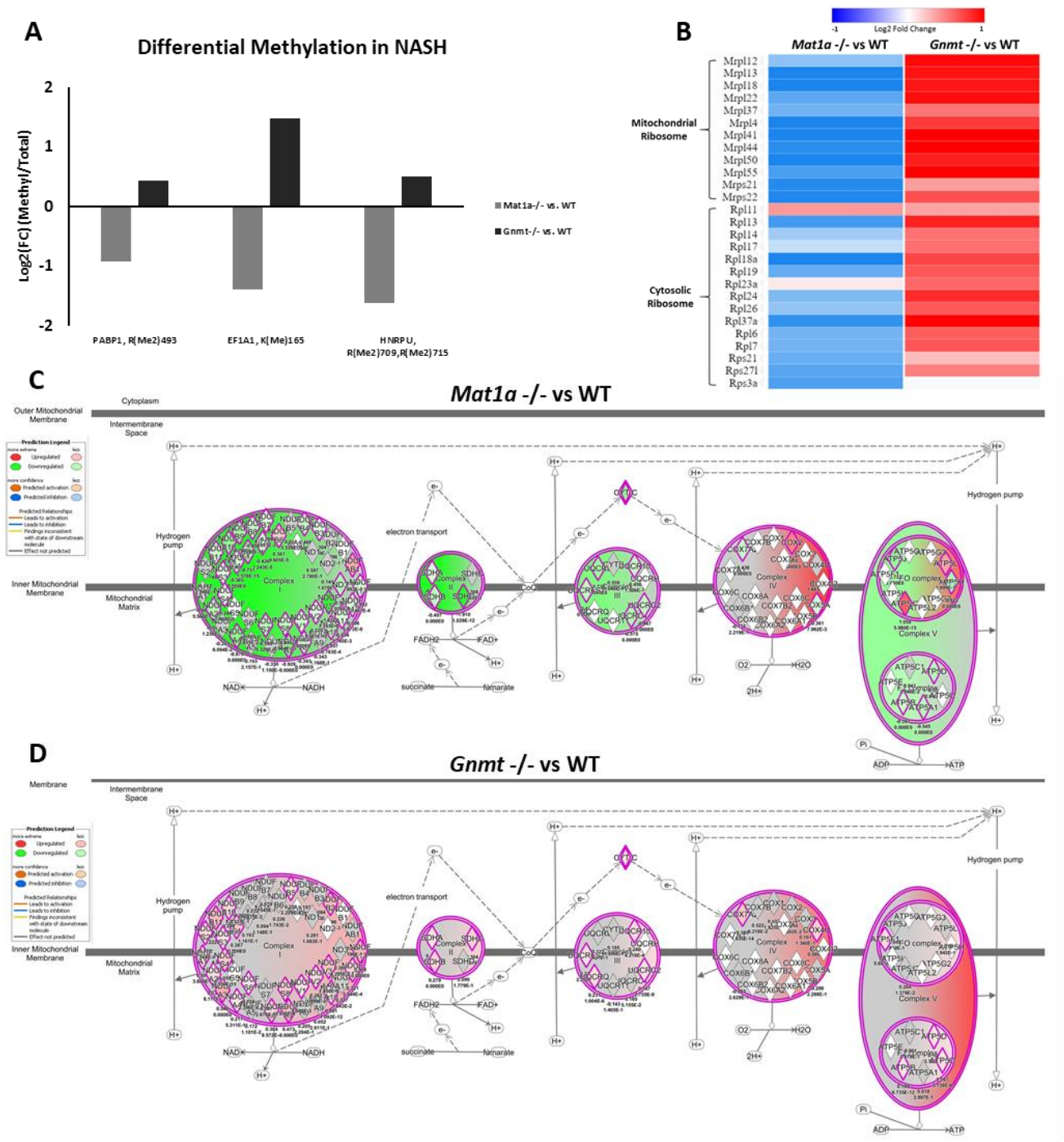


Figure 7. Methylation's Differential Effect on Translation: (A) Quantification of the intensity of the MS2 TIC normalized methylated peptide divided by the intensity of their respective total protein, determined by unmodified peptides using MAPDIA in Gnmt -/- and wild-type littermates and Mat1a -/- and wild-type littermates **(B)** Total protein quantitation of mitochondrial ribosomes and cytosolic ribosomes determined using MAPDIA in Gnmt -/- and wild-type littermates and Mat1a -/- and wild-type littermates**.** All data has a FDR<0.01. **(C)** Predicted inhibition of oxidative phosphorylation complex 1,2,3, and 5 in *Mat1a* -/- vs WT from total protein using Ingenuity Pathway Analysis (activation z-score -4.61) **(D)** Predicted activation of oxidative phosphorylation complex 1,2,3, 4, and 5 in *Gnmt* -/- vs WT (activation z-score 6.403) All data has FDR<0.01.

Disparate Mechanisms Drive Decreased Acetylation in Murine NASH Models

Protein acetylation was decreased in both models of NASH (Figure S12A) compared to controls. This could be due to two different mechanisms which should also be observed by changes in the proteome. First, metabolomic analysis of Acetyl-CoA levels from livers from *Mat1a* -/- and wild-type littermates show a 54% decrease in Acetyl-CoA (p-value = 0.0002) in *Mat1a* -/-, (Figure S12D) supporting globally decreased acetylation in *Mat1a* -/-. Interestingly, there is no change in the total protein quantity of sirtuins SIRT2 and SIRT5, a class of protein de-acetylases, in the proteomics analysis of *Mat1a* -/-. Second, in *Gnmt* -/-, there is a 61% increase of the deacetylase SIRT3 at the total protein level (FDR = 5.4E-18). In addition to increased SIRT3 levels in *Gnmt* -/- mice, we also find significant increases in total protein abundance of IDH2 and SOD2, both of which are known to be activated via SIRT3 mediated deacetylation[53–55] (Figure S12B). In our acetyl-peptide dataset, we found a known target of SIRT3 deacetylation, ECHA K411-Acetyl, significantly decreased in *Gnmt* -/-, in agreement with SIRT3 activation in *Gnmt* -/- (Figure S12C). In addition at the total protein level, we see substantial increase of all five complexes of the electron transport chain in *Gnmt* -/-, also consistent with downstream effects of SIRT3 mediated deacetylation[53–55](Figure 7D).

Discussion

In this study, we have developed a method utilizing 4 Dalton precursor mass windows that is able to differentiate methylated peptidoforms (mono-, di- and tri-methylated Lys and mono- and di-methylated Arg) from each other and their corresponding unmodified amino acid sequence while still allowing for quantification of unmodified peptides required for total protein quantification. A previous study by *Sidoli et. al.*, also has applied small precursor mass windows to analyze methylation from histone enriched lysate using multiplexed 6 Dalton precursor mass windows which have the ability to differentiate a +2 precursor from its unmodified form but not a +3 precursor.[37]

We compared multiplexed 4 and 6 Dalton Precursor mass windows using the approach outlined in *Sidoli et al*., to non-multiplexed windows on complex liver lysate. Although we were able to find methylated peptides, the intensity was on average 10-fold lower than our 4 Dalton Precursor mass window method. The fragment ions do not perfectly overlap causing issues with identification and quantitation of the peaks (Figure S13). Combined with the long time needed for data analysis to de-multiplex windows with current tools, we chose to not use that method for this study. We have shown that by using synthesized unmodified and methylated peptides that we have the ability to physically separate 100% of the observable +2 and +3 methylated precursors from their unmodified forms. Furthermore, we applied our method to complex cellular lysate and have shown that our method can quantify protein methylation. In addition to protein methylation, we have shown that our method can be expanded to quantify Lys acetylation and succinylation. The ability to assay multiple biologically relevant Lys PTMs not only provides an opportunity to quantify methylated, acetylated, and succinylated peptides from the same sample and acquisition run but also gives insight into the importance of Lys PTM crosstalk.

Lys residues can be modified by multiple PTMs and crosstalk between these PTMs constitutes a major regulatory mechanism of protein function.[25–27] Despite its importance in biology, PTM crosstalk is difficult to study and currently not possible without enrichment of each respective PTM. Using our

approach, we can assay 946 Lys residues which contain multiple PTMs; 24 Lys residues contain multiple methyl forms (Figure S14A) and 922 Lys residues where crosstalk between acetylation, succinylation, or methylation occurs, 872 of which have crosstalk between acetylation and succinylation (Figure S14B). Having the ability to simultaneously identify and quantify peptides with Lys PTM crosstalk in one DIA run will further our understanding of the role of PTM crosstalk in biological systems.

Growing interest in studying residue specificity of PTMs using DIA approaches has created a need for tools that can accurately localize modified amino acid residues. It has recently become possible to computationally assess the strength of PTM localization of peptidoforms which contain positional isomers using tools which score distinguishing fragment ions between isomers to differentiate and assess the correct site of modification.[7,9,10] However, these tools were all tested and optimized for phospho-peptide positional isomer localization using synthesized peptides and designed to be used on phospho-peptide enrichments; a significantly less complex sample matrix than complex cellular lysate. We chose to use Thesaurus 0.9.4 for modified peptidoform localizations, a graphical interface easily operated by an end-user which contains the ability to score and localize the Lys and Arg PTM's used in this study increasing the confidence of PTMs identified with our approach.

We have shown that our 4 Da precursor mass window DIA method provides linear protein quantitation in both the 60- or 120-minute LC gradient and that the quantitation of the method significantly outperforms current DIA acquisition methods for low abundant peptides, and therefore will be more useful to quantify low abundant PTMs in complex lysate. We recommend using a 120-minute LC gradient for future studies where quantitation is important, as with our method's cycle time; we achieve more points across the chromatographic peak. Furthermore, we have shown that our method outperforms MS1 centric workflows for low abundant peptides, including endogenous low abundant and modified peptides. In the majority of cases, as the abundance of a peptide decreases, the co-eluting precursor trace is lost. Instead of relying on a coeluting precursor trace to differentiate a methylated peptide from an unmodified peptide, our method

physically separates the unmodified and methylated peptides by acquiring their precursors into separate mass windows allowing for the quantitation of low abundant modified peptides without the presence of a MS1 trace.

To better understand the role of PTMs in NASH, we have tested this methodology in two differentially methylated mouse models where the primary methyl donor, SAMe, is genetically manipulated.[32,34] As mentioned previously, two major findings were changes in protein translation was a major differentially dysregulated pathway based on total protein quantity in *Gnmt* -/- and *Mat1a* -/- NASH. We also linked differential methylation potential and altered acetylation together with total protein changes to propose a hypothesis of regulatory function of post-translational modifications on proteins involved in mRNA stability, protein synthesis, and mitochondrial function between the *Mat1a* -/- and *Gnmt* -/- NASH models. For example, differential methylation patterns detected in amino acid residues of RNA binding proteins (EF1A1, PABP1) indicate that decreased methylation could impair messenger RNA stability and protein translation. It has previously been reported that PABP1 is controlled by post-translational modifications and that the proline-rich linker region is crucial for its polymerization, for mature mRNA nuclear export, protein synthesis, and as well as its capacity to bind and stabilize mRNA.[56–58] The di-methyl site (R493), identified in this study on PABP1, is a known methyl site which resides within this critical proline-rich linker region but has not been attributed to known function or previously described in NASH. This study's method has established a link between significantly changed methylated peptides and differentially expressed proteins between *Gnmt* -/- and *Mat1a* -/- NASH. Differential methylation of PABP1 via destabilized mRNA could be responsible for dysregulated protein synthesis in observed significantly altered pathways (Figure 7B-D).

While the cellular methylation potential in these two animal models is solely controlled by SAMe concentration, acetylation is not genetically manipulated in either of these models of NASH and has multiple points of control. Globally decreased protein acetylation observed in both NASH models could

be due to two different mechanisms indicative of methylation/acetylation cross talk. Acetylation is thought to mainly occur non-enzymatically,[28,59] so limited cellular concentration of the main acetyl group donor (Acetyl-CoA) is one of possible mechanisms explaining observed changes in protein acetylation in Mat1a -/-. Alternatively, downregulation of protein acetylation could be driven by activation of sirtuins, enzymes that deacetylate proteins.[60,61] While there is no increase in observable sirtuins abundance in hypomethylated background (Mat1a -/-), we observe targets concordant with (SIRT3) activation in hypermethylated Gnmt -/-. SIRT3 is known to have a broad interactome network,[59,60] and in our dataset we found alterations in downstream targets known to be activated upon SIRT3 deacetylation.[53–55] Although we don't detect SIRT3 mediated de-acetylation in proteins involved in Oxidative Phosphorylation, we do see a substantial increase in the total protein quantity of all five mitochondrial complexes in *Gnmt -/-*, which could partially be regulated by increased SIRT3 activity in addition to PABP1 hypermethylation.

There are some limitations with our approach. DIA-MS relies on a pre-existing peptide assay library and compares the MS data obtained from each experimental sample to this library. As such, PTMs not contained within this assay library are not assayed in experimental samples using our approach and the library is species specific and often sample (e.g. cell type) specific. We have shown the ineffectiveness of current library free approaches at detecting PTMs (Figure S4C). To this point, this is why so much work went into maximizing the PTMs in our murine PTM enriched DIA assay library. As with all MS analysis including DDA analysis of complex samples, the proteome depth is limited unless fractionation of some form is used. The hereby proposed methodology is limited to the dynamic range of unfractionated complex lysate, and we are only able to detect the highest abundant methylated, succinylated, and acetylated peptides. However, DIA-MS approaches display a number of favorable properties, by systematically fragmenting all precursor ions in precisely defined (peptidoform-tailored) width of mass to charge windows, and concurrently overcome the stochastic ion selection of data-dependent-acquisition (DDA), providing larger completeness to the dataset. The ability to quantitate highly abundant modified

peptides as well as total protein quantitation from the same DIA acquisition provides a complementary and analytically valuable alternative despite the limited depth. In addition, we have shown that our method performs quite well with peptide of low abundance and in fact outperforms existing DIA acquisition methods for low abundant peptides at concentrations similar to biological PTMs in complex lysate (Figure S3).

Another caveat with our strategy is that although the unmodified peptide is always in a different mass window as the methylated peptide, if the unmethylated counterpart sequence contains a methionine, there is a chance that the oxidized methionine species will end up in the same precursor mass window as the methylated peptide. In this situation, quantification of a site-specific transition or a co-eluting precursor trace is needed to verify correct precursor mass. In this case, a retention time difference is a valuable method for further peptidoform differentiation, as peptides containing an oxidized methionine elute before their corresponding unmodified peptide by 2.37% ACN 'gradient units'.[62] By looking at either the retention time difference from the unmodified peptide, a co-eluting precursor trace, or by having a site specific transition, one can easily differentiate a methylated peptide from an oxidized peptide. In addition, methylated peptides can co-elute with unmodified peptides. Using 4 Da precursor mass windows gives the advantage of not only relying on site-specific transitions to quantify a modified peptide because the methylated and unmodified peptides are physically separated in the mass spectrometer allowing for the use of all transitions for quantification of the respective peptidoforms. Using large precursor mass windows can lead to ambiguity, especially in cases where there is only one site specific transition. Additional consideration should be taken when comparing tri-methylated and acetylated peptides. Acetylation of ε-amino groups of lysine results in shifted retention time with respect to their unmodified counterpart sequences whereas tri-methylation does not.[63] Furthermore, high mass accuracy instrumentation can distinguish these isobaric modifications.[64] Therefore the combination of relative retention time shifts with site-specific transitions allows for the unequivocal determination of tri-methylation vs. acetylation.

Finally, a methylated Lys or Arg often inhibits trypsin's ability to cleave.[52] In order to determine the stoichiometric site specific ratio of a modified peptidoform, we ideally would use an enzyme that would give a methylated peptidoform without any missed cleavages and normalize to the intensity of the unmodified peptidoform which would give a better representation of the quantity of the modified peptidoform than our approach. However, to make our method for methylation quantitation more useful for the proteomics community as a whole, we decided to use Trypsin for our study as it is currently the most commonly used endoproteinase in the field. However, missed cleavages of unmodified peptides do still occur even with Trypsin/LysC mix 5-10% of the time.[65] An unmodified peptide with a missed cleavage can occur due to incomplete digestion, although generally it would be at a much lower abundance than the tryptic version of the peptide. Since the PTMs in this study are also low abundant, it would be possible to misidentify an incompletely digested unmodified peptide containing a missed cleavage with a PTM containing peptide especially if it had only few low-abundant site-specific transitions. However, this is not the case when we apply our 4 Da window method which physically separates the unmodified from the modified version within the MS instrument. Of importance, there are also plenty of methylated peptides that do not have a missed cleavage in our database, and which have been reported. For example, in Geoghegan et al, 2015[66] they find (682/5781) 11.8% of methylated peptides without a missed cleavage (tryptic peptides). In these situations, one would run into the issue where the tryptic unmodified and methylated peptide would occur and co-elute in the same precursor mass window, unless they applied 4 Da window DIA acquisitions

The small precursor mass window DIA approach outlined in this study which is suitable for both triple TOF and Orbitrap mass spectrometers along with the corresponding publicly available hypermethylated, acetylated, and succinylated mouse liver methyl peptide assay library provides an easily adoptable framework needed for studying global protein methylation, acetylation, and succinylation as well as total protein quantification in any DIA proteomic experiment without the need for enrichment of each experimental sample. This approach only requires enough sample for one DIA acquisition which opens

up the possibility for studying PTMs from small experimental samples like tissue biopsies where there isn't enough sample for traditional PTM immunoprecipitations. Wide adoption of this easy approach will allow more rapid integration of methylation studies in DIA proteomic experiments and increase the knowledge of the role of methylation in complex biological systems.

Conclusion:

In this study we have developed a DIA-based method utilizing 4 Dalton precursor mass windows that is able to differentiate all forms of methylated peptides from their unmodified forms using synthesized methylated peptides and validated the method for accurate quantitation of low abundant peptides. We included an algorithm to assess PTM site localization providing the ability to quantify high quality PTMs. We then interrogated this novel method on in-vivo complex cellular lysate using a PTM-enriched mouse liver DIA assay library which is publicly available and can be used in future studies as a community resource. We have shown that we are able to detect and quantify protein methylation from our approach in two differentially methylated in-vivo mouse models of NASH as well as quantify lysine acetylation, succinylation, and total protein from the same sample in a single acquisition. At the same time, we can obtain both site specific methylation, acetylation, succinylation, and total protein quantity allowing for the determination of relative ratio for each modified peptidoform. We applied our method to two differentially methylated proteomes and were able to link observed protein modification changes to significantly altered protein pathways. We show that this proteome and PTM analysis approach is an extremely powerful tool for novel hypothesis generation, demonstrating potential regulatory function of methylation in mRNA stability and translation as well as the unexpected interplay of protein acetylation in NASH. Overall, our study describes a novel approach to simultaneously quantify protein methylation, acetylation, succinylation, and total protein in in-vivo complex cellular lysate in one DIA acquisition without the need for enrichment of each biological sample and has highlighted the importance of the regulatory function and interplay of PTMs in NASH.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgment

The authors thank Helen Choi Robinson for helping with figure design. This work was supported by USA National Institutes of Health (NIH) grants R01DK107288 (SC Lu and J Van Eyk), NIH GM110174 and AI118891 (B Garcia), and Agencia Estatal de Investigación MINECO SAF 2017-88041-R, ISCiii PIE14/00031 CIBERehd-ISCiii, and Severo Ochoa Excellence Accreditation SEV-2016-0644) (J Mato).

ABBREVIATIONS

PTMs	Post translational modifications
MS	Mass spectrometry
LC	Liquid Chromatography
DIA	Data independent acquisitions
SAMe	S-Adenosylmethionine
NASH	Non-alcoholic steatohepatitis
Mat1a	Methionine Adenosyltransferase A1
Gnmt	Glycine *N*-methyltransferase
SCX	Strong Cation Exchange
SIL	Stabile Isotopically Labeled
DDA	Data Dependent Acquisition
FDR	False discovery rate
TIC	Total Ion Current
fm	femtomoles
AUC	area-under-the-curve

FIGURE LEGENDS

Figure 1. Schematic Overview of Experimental Design: (A) An equimolar pool of 400 synthesized peptides containing K[Unmodified], K[Monomethyl], K[Dimethyl], and K[Trimethyl] residues were acquired in DIA on the Sciex TripleTOF 6600 with 4 Dalton and 100 Variable Window (VW) precursor mass windows. (B) Same peptide pool was acquired in DIA on the Thermo Orbitrap Fusion Lumos with 4, 6, and 12 Dalton precursor mass windows. (C) Strategy to differentiate an unmodified peptide from it's methylated form based off of precursor mass window difference. An unmodified precursor will be fragmented in a different mass window than its methylated form making it differentiable by precursor mass window.

Figure 2. Synthesized Methylated Peptides: Presence of synthesized methylated and unmodified peptides acquired on the Thermo Orbitrap Fusion Lumos with 4, 6, and 12 Dalton precursor mass windows were visualized in Skyline (A) Quantified transitions of DYSSGFGGKYGVQADR (B) Quantified transitions of DYSSGFGG**K[Me1]**YGVQADR (C) Graphical representation of unmodified and monomethyl +2 and +3 precursors which can be differentiated from their unmodified forms using 4, 6, and 12 Dalton precursor mass windows. Black represents peptides eluting in different precursor mass windows while grey represents peptides co-eluting in the same precursor mass window.

Figure 3. *In-vivo* Methylation: (A) Quantification of the MS2 TIC normalized intensity of peptide containing EF1A1 K[165] Mono-Methyl in *Gnmt* -/- and wild-type littermates and *Mat1a* -/- and wild-type littermates Data are box and whisker plots of six biological replicates. Two-tailed Student's t-test, $^{***}P < 0.005$, $^{*}P < 0.05$ **(**B) Quantification of the intensity of the MS2 TIC normalized peptide containing EF1A1 K[165] Mono-Methyl divided by the intensity of EF1A1 total protein, determined by EF1A1 unmodified peptides using MAPDIA in *Gnmt* -/- and wild-type littermates and *Mat1a* -/- and wild-type littermates Data are box and whisker plots of six biological replicates per condition. Two-tailed Student's t-test, $^{***}P < 0.0005$ (C) Skyline Visualization of the XIC of the peptide MDSTEPPYSQK[Methyl]R which corresponds to EF1A1 K[165] Mono-Methyl in *Gnmt* -/- complex liver lysate (D) Skyline

Visualization of the XIC of a stably isotopically labeled peptide MDSTEPPYSQK[Methyl]R[heavy] which corresponds to EF1A1 K[165] Mono-Methyl spiked into *Gnmt* -/- complex liver lysate

Figure 4. Ambiguous Methylation Quantitation: Complex cellular liver lysate from *Mat1a* -/- acquired on the Thermo Orbitrap Fusion Lumos with 4 and 12 Dalton precursor mass windows visualized in Skyline (A) Quantified transitions of VETGVLKPGMVVTFAPVNVTTEVK acquired with 12 Dalton precursor mass window (B) Quantified transitions of VETGVL**K[Methyl]**PGMVVTFAPVNVTTEVK acquired with 12 Dalton precursor mass windows. This precursor falls in the same mass window as its unmodified form, elutes at the same time, and contains one site specific transition (b7). (C) Quantified transitions of VETGVLKPGMVVTFAPVNVTTEVK acquired with 4 Dalton precursor mass window (D) Quantified transitions of VETGVL**K[Methyl]**PGMVVTFAPVNVTTEVK acquired with 4 Dalton precursor mass windows. There is no discernable peak; leading us to believe that the quantitation of the methylated peptide found with 12 Dalton precursor mass windows is incorrect due to many ambiguous transitions. (E) Graphical representation of unmodified, monomethyl, and dimethyl +2 and +3 precursors which can be differentiated from their unmodified forms using 4 and 12 Dalton precursor mass windows. Grey represents peptides eluting in different precursor mass windows while blue represents peptides co-eluting in the same precursor mass window.

Figure 5. In-vivo Succinylation and Acetylation: (A) Quantification of the MS2 TIC normalized intensity of peptide containing MDHM K[165] Succinyl in *Gnmt* -/- and wild-type littermates and *Mat1a* -/- and wild-type littermates (n=6/condition) Data are box and whisker plots of six biological replicates per condition. Two-tailed Student's t-test, $^{**}P < 0.005$ **(**B) Quantification of the intensity of the MS2 TIC normalized peptide containing MDHM K[165] Succinyl divided by the intensity of MDHM total protein, determined by MDHM unmodified peptides using MAPDIA. Data are box and whisker plots of six biological replicates per condition. Two-tailed Student's t-test, $^{**}P < 0.005$, $^{*}P < 0.05$ (C) Skyline Visualization of the XIC of the peptide containing MDHM K[165] Succinyl **(**D) Quantification of the MS2 TIC normalized intensity of a proteotypic peptide mapping to H4 K[8,12,16] Acetyl in *Gnmt* -/- and

wild-type littermates and *Mat1a -/-* and wild-type littermates. Data are box and whisker plots of six biological replicates per condition. **(**E) Quantification of the intensity of the MS2 TIC normalized peptide containing H4 K[8,12,16] Acetyl divided by the intensity of H4 total protein, determined by H4 unmodified peptides using MAPDIA in *Gnmt* -/- and wild-type littermates and *Mat1a -/-* and wild-type littermates. Data are box and whisker plots of six biological replicates per condition. Two-tailed Student's t-test, $^{**}P < 0.005$, $^{*}P < 0.05$ (F) Skyline Visualization of the XIC of the peptide containing H4 K[8,12,16] Acetyl

Figure 6. Modified Peptidoform Localization: (A) Localization results of modified peptidoforms using Thesaurus 0.9.4 from a mixture of synthesized peptides containing the same peptide backbone but with modifications (mono-, di-, tri-methylated and acetylated) on the same a non-terminal Lys residue. **(B)** Venn diagram showing the overlap of Thesaurus localization results from each modification type from mixture of synthesized peptides **(C)** Quantified and localized modified peptidoform from 4 Da DIA acquisitions of complex cellular lysate from Gnmt -/- and Gnmt WT livers assayed against the PTM enriched DIA Assay Library in either OpenSWATH 3.2.0 or Thesaurus 0.9.4 and identified with >1% FDR in ≥4/12 (≥33%) of samples in each respective approach in addition to >1% FLR localization calculated in Thesaurus 0.9.4 and applied to both the OpenSWATH and Thesaurus results. **(D)** Venn diagram showing the overlap of quantifiable peptidoforms from OpenSWATH, quantifiable peptidoforms from OpenSWATH which were localized using Thesaurus, and peptidoforms which were both quantifiable and localization from Thesaurus.

Figure 7. Methylation's Differential Effect on Translation: (A) Quantification of the intensity of the MS2 TIC normalized methylated peptide divided by the intensity of their respective total protein, determined by unmodified peptides using MAPDIA in Gnmt -/- and wild-type littermates and Mat1a -/- and wild-type littermates **(B)** Total protein quantitation of mitochondrial ribosomes and cytosolic ribosomes determined using MAPDIA in Gnmt -/- and wild-type littermates and Mat1a -/- and wild-type littermates**.** All data has a FDR<0.01. **(C)** Predicted inhibition of oxidative phosphorylation complex 1,2,3, and 5 in *Mat1a* -/- vs WT from total protein using Ingenuity Pathway Analysis (activation z-score -

4.61) **(D)** Predicted activation of oxidative phosphorylation complex 1,2,3, 4, and 5 in *Gnmt* -/- vs WT (activation z-score 6.403) All data has FDR<0.01.

REFERENCES

1. Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* **17**, 666–672 (2010).
2. Murn, J. & Shi, Y. The winding path of protein methylation research: milestones and new frontiers. *Nat. Rev. Mol. Cell Biol.* **18**, 517–527 (2017).
3. Biggar, K. K. & Li, S. S.-C. Non-histone protein methylation as a regulator of cellular signalling and function. *Nat. Rev. Mol. Cell Biol.* **16**, 5 (2014).
4. Ballif, B. A., Carey, G. R., Sunyaev, S. R. & Gygi, S. P. Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. *J. Proteome Res.* **7**, 311–318 (2008).
5. Montoya, A., Beltran, L., Casado, P., Rodríguez-Prados, J.-C. & Cutillas, P. R. Characterization of a TiO2 enrichment method for label-free quantitative phosphoproteomics. *Methods San Diego Calif* **54**, 370–378 (2011).
6. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520 (2015).
7. PIQED: automated identification and quantification of protein modifications from DIA-MS data | Nature Methods. https://www.nature.com/articles/nmeth.4334.
8. Bekker-Jensen, D. B. *et al.* Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nat. Commun.* **11**, 1–12 (2020).
9. Rosenberger, G. *et al.* Inference and quantification of peptidoforms in large sample cohorts by SWATH-MS. *Nat. Biotechnol.* **35**, 781–788 (2017).
10. Searle, B. C., Lawrence, R. T., MacCoss, M. J. & Villén, J. Thesaurus: quantifying phosphopeptide positional isomers. *Nat. Methods* **16**, 703–706 (2019).
11. Mann, M. & Jensen, O. N. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21**, 255–261 (2003).
12. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355 (2016).
13. Collins, B. C. *et al.* Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system. *Nat. Methods* **10**, 1246–1253 (2013).
14. Paik, W. K., Paik, D. C. & Kim, S. Historical review: the field of protein methylation. *Trends Biochem. Sci.* **32**, 146–152 (2007).
15. Boisvert, F.-M., Rhie, A., Richard, S. & Doherty, A. J. The GAR motif of 53BP1 is arginine methylated by PRMT1 and is necessary for 53BP1 DNA binding activity. *Cell Cycle Georget. Tex* **4**, 1834–1841 (2005).
16. Dhayalan, A., Kudithipudi, S., Rathert, P. & Jeltsch, A. Specificity analysis-based identification of new methylation targets of the SET7/9 protein lysine methyltransferase. *Chem. Biol.* **18**, 111–120 (2011).
17. Swiercz, R., Cheng, D., Kim, D. & Bedford, M. T. Ribosomal protein rpS2 is hypomethylated in PRMT3-deficient mice. *J. Biol. Chem.* **282**, 16917–16923 (2007).
18. Wang, K. *et al.* Antibody-Free Approach for the Global Analysis of Protein Methylation. *Anal. Chem.* **88**, 11319–11327 (2016).
19. Cao, X.-J., Arnaudo, A. M. & Garcia, B. A. Large-scale global identification of protein lysine methylation in vivo. *Epigenetics* **8**, 477–485 (2013).
20. Uhlmann, T. *et al.* A Method for Large-scale Identification of Protein Arginine Methylation. *Mol. Cell. Proteomics* **11**, 1489–1499 (2012).

21. Wang, Q., Liu, Z., Wang, K., Wang, Y. & Ye, M. A new chromatographic approach to analyze methylproteome with enhanced lysine methylation identification performance. *Anal. Chim. Acta* **1068**, 111–119 (2019).
22. Larsen, S. C. *et al.* Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Sci. Signal.* **9**, rs9 (2016).
23. Gayatri, S. *et al.* Using oriented peptide array libraries to evaluate methylarginine-specific antibodies and arginine methyltransferase substrate motifs. *Sci. Rep.* **6**, (2016).
24. Cornett, E. M. *et al.* A functional proteomics platform to reveal the sequence determinants of lysine methyltransferase substrate selectivity. *Sci. Adv.* **4**, eaav2623 (2018).
25. Horita, H., Law, A. & Middleton, K. Utilizing Optimized Tools to Investigate PTM Crosstalk: Identifying Potential PTM Crosstalk of Acetylated Mitochondrial Proteins. *Proteomes* **6**, (2018).
26. Yang, X.-J. & Seto, E. Lysine Acetylation: Codified Crosstalk with Other Posttranslational Modifications. *Mol. Cell* **31**, 449–461 (2008).
27. Xu, H.-D., Wang, L.-N., Wen, P.-P., Shi, S.-P. & Qiu, J.-D. Site-Specific Systematic Analysis of Lysine Modification Crosstalk. *Proteomics* **18**, e1700292 (2018).
28. Basisty, N., Meyer, J. G., Wei, L., Gibson, B. W. & Schilling, B. Simultaneous Quantification of the Acetylome and Succinylome by 'One-Pot' Affinity Enrichment. *Proteomics* **18**, (2018).
29. Stastna, M. & Van Eyk, J. E. Posttranslational modifications of lysine and evolving role in heart pathologies-recent developments. *Proteomics* **15**, 1164–1180 (2015).
30. Alonso, C. *et al.* Metabolomic Identification of Subtypes of Nonalcoholic Steatohepatitis. *Gastroenterology* **152**, 1449-1461.e7 (2017).
31. Varela-Rey, M. *et al.* Fatty liver and fibrosis in glycine N-methyltransferase knockout mice is prevented by nicotinamide. *Hepatol. Baltim. Md* **52**, 105–114 (2010).
32. Luka, Z., Capdevila, A., Mato, J. M. & Wagner, C. A glycine N-methyltransferase knockout mouse model for humans with deficiency of this enzyme. *Transgenic Res.* **15**, 393–397 (2006).
33. Barr, J. *et al.* Liquid Chromatography-Mass Spectrometry (LC/MS)-based parallel metabolic profiling of human and mouse model serum reveals putative biomarkers associated with the progression of non-alcoholic fatty liver disease. *J. Proteome Res.* **9**, 4501–4512 (2010).
34. Lu, S. C. *et al.* Methionine adenosyltransferase 1A knockout mice are predisposed to liver injury and exhibit increased expression of genes involved in proliferation. *Proc. Natl. Acad. Sci.* **98**, 5560–5565 (2001).
35. Martínez-López, N. *et al.* Hepatoma Cells from Mice Deficient in Glycine N-Methyltransferase Have Increased RAS Signaling and Activation of Liver Kinase B1. *Gastroenterology* **143**, 787-798.e13 (2012).
36. Kooij, V., Holewinski, R. J., Murphy, A. M. & Van Eyk, J. E. Characterization of the Cardiac Myosin Binding Protein-C Phosphoproteome in Healthy and Failing Human Hearts. *J. Mol. Cell. Cardiol.* **60**, 116–120 (2013).
37. Sidoli, S., Fujiwara, R. & Garcia, B. A. Multiplexed data independent acquisition (MSX-DIA) applied by high resolution mass spectrometry improves quantification quality for the analysis of histone peptides. *Proteomics* **16**, 2095–2105 (2016).

38. Parker, S. J., Venkatraman, V. & Van Eyk, J. E. Effect of peptide assay library size and composition in targeted data-independent acquisition-MS analyses. *Proteomics* **16**, 2221–2237 (2016).
39. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: an open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24 (2013).
40. Craig, R. & Beavis, R. C. TANDEM: matching proteins with tandem mass spectra. *Bioinforma. Oxf. Engl.* **20**, 1466–1467 (2004).
41. MacLean, B., Eng, J. K., Beavis, R. C. & McIntosh, M. General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. *Bioinforma. Oxf. Engl.* **22**, 2830–2832 (2006).
42. Escher, C. *et al.* Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **12**, 1111–1121 (2012).
43. Parker, S. J. *et al.* Identification of a Set of Conserved Eukaryotic Internal Retention Time Standards for Data-independent Acquisition Mass Spectrometry. *Mol. Cell. Proteomics MCP* **14**, 2800–2813 (2015).
44. Egertson, J. D., MacLean, B., Johnson, R., Xuan, Y. & MacCoss, M. J. Multiplexed Peptide Analysis using Data Independent Acquisition and Skyline. *Nat. Protoc.* **10**, 887–903 (2015).
45. Röst, H. L. *et al.* OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat. Biotechnol.* **32**, 219–223 (2014).
46. Reiter, L. *et al.* mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* **8**, 430–435 (2011).
47. Röst, H. L. *et al.* TRIC: an automated alignment strategy for reproducible protein quantification in targeted proteomics. *Nat. Methods* **13**, 777–783 (2016).
48. Teo, G. *et al.* mapDIA: Preprocessing and statistical analysis of quantitative proteomics data from data independent acquisition mass spectrometry. *J. Proteomics* **129**, 108–120 (2015).
49. Barbier-Torres, L. *et al.* Stabilization of LKB1 and Akt by neddylation regulates energy metabolism in liver cancer. *Oncotarget* **6**, 2509–2523 (2015).
50. Martínez-Arranz, I. *et al.* Enhancing metabolomics research through data mining. *J. Proteomics* **127**, 275–288 (2015).
51. DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics | Nature Methods. https://www.nature.com/articles/nmeth.3255.
52. Ong, S.-E., Mittler, G. & Mann, M. Identifying and quantifying *in vivo* methylation sites by heavy methyl SILAC. *Nat. Methods* **1**, 119–126 (2004).
53. Ogura, Y. *et al.* Renal mitochondrial oxidative stress is enhanced by the reduction of Sirt3 activity, in Zucker diabetic fatty rats. *Redox Rep. Commun. Free Radic. Res.* **23**, 153–159 (2018).
54. Qiu, X., Brown, K., Hirschey, M. D., Verdin, E. & Chen, D. Calorie restriction reduces oxidative stress by SIRT3-mediated SOD2 activation. *Cell Metab.* **12**, 662–667 (2010).
55. Someya, S. *et al.* Sirt3 mediates reduction of oxidative damage and prevention of age-related hearing loss under caloric restriction. *Cell* **143**, 802–812 (2010).
56. Brook, M. *et al.* The multifunctional poly(A)-binding protein (PABP) 1 is subject to extensive dynamic post-translational modification, which molecular modelling suggests plays an important role in co-ordinating its activities. *Biochem. J.* **441**, 803–812 (2012).
57. Sawazaki, R. *et al.* Characterization of the multimeric structure of poly(A)-binding protein on a poly(A) tail. *Sci. Rep.* **8**, (2018).

58. Shan, P. *et al.* SIRT1 Functions as a Negative Regulator of Eukaryotic Poly(A)RNA Transport. *Curr. Biol.* **27**, 2271-2284.e5 (2017).
59. Peterson, B. S. *et al.* Remodeling of the Acetylproteome by SIRT3 Manipulation Fails to Affect Insulin Secretion or β Cell Metabolism in the Absence of Overnutrition. *Cell Rep.* **24**, 209-223.e6 (2018).
60. Bheda, P., Jing, H., Wolberger, C. & Lin, H. The Substrate Specificity of Sirtuins. *Annu. Rev. Biochem.* **85**, 405–429 (2016).
61. Chang, H.-C. & Guarente, L. SIRT1 and other sirtuins in metabolism. *Trends Endocrinol. Metab.* **25**, 138–145 (2014).
62. Lao, Y. W. *et al.* Chromatographic behavior of peptides containing oxidized methionine residues in proteomic LC–MS experiments: Complex tale of a simple modification. *J. Proteomics* **125**, 131–139 (2015).
63. Yang, L. *et al.* Unambiguous Determination of Isobaric Histone Modifications by Reversed-Phase Retention Time and High-Mass Accuracy. *Anal. Biochem.* **396**, 13–22 (2010).
64. Zhang, L., Eugeni, E. E., Parthun, M. R. & Freitas, M. A. Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma* **112**, 77–86 (2003).
65. Saveliev, S. *et al.* Trypsin/Lys-C protease mix for enhanced protein mass spectrometry analysis. *Nat. Methods* **10**, 1134–1134 (2013).
66. Geoghegan, V., Guo, A., Trudgian, D., Thomas, B. & Acuto, O. Comprehensive identification of arginine methylation in primary T cells reveals regulatory roles in cell signalling. *Nat. Commun.* **6**, 6758 (2015).