bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Quantifying dynamic protein acetylation using quantitative stoichiometry

Authors: Josue Baeza[1,2], Alexis J. Lawton[1,2], Jing Fan[2,3], Michael J. Smallegan[4,5], Ian Lienert[6], Tejas Gandhi[6], Oliver M. Bernhardt[6], Lukas Reiter[6], John M. Denu[1,2,7]

Author Affiliations:
[1]Department of Biomolecular Chemistry, University of Wisconsin-Madison, Madison, WI
[2]Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI
[3]Morgridge Institute for Research, University of Wisconsin-Madison, Madison, WI
[4]BioFrontiers Institute, University of Colorado, Boulder, CO
[5]Molecular Cellular and Developmental Biology Department, University of Colorado, Boulder, CO
[6]Biognosys AG, Wagistrasse 25, CH-8952 Schlieren, Switzerland
[7]Corresponding author. Email- john.denu@wisc.edu

Abstract:
Protein acetylation is a widespread post-translational modification that is implicated in many cellular processes. Emergent technologies such as mass spectrometry has catalogued thousands of sites throughout the cell, however identifying regulatory acetylation marks has proven to be a daunting task. Here, we reveal the dynamic nature of site-specific acetylation in response to serum stimulation. An improved method of quantifying acetylation stoichiometry was developed and validated, providing a detailed landscape of dynamic acetylation stoichiometry within cellular compartments. The nuclear compartment displayed significantly higher median stoichiometry, in line with the existence of known acetyltransferases. Site-specific, temporal acetylation using a serum starved-refed cell culture model uncovered distinct clusters of acetylations with diverse kinetic profiles. Of particular note, dynamic acetylation sites on protein translational machinery suggests a major regulatory role.

Introduction: Reversible lysine acetylation is a regulatory modification which has emerged as a widespread modification, rivaling phosphorylation in scope (1). Protein acetylation was first discovered on the N-terminal lysine residues of histone proteins localized to the nucleus (2, 3) and has since been identified throughout the cell including: cytoplasm (4), mitochondria (5), endoplasmic reticulum (6), peroxisomes (7), etc. In the nucleus, histone acetylation is associated with active gene expression, acting in part to open chromatin and allowing access to transcriptional machinery. Bromodomain containing proteins recognize and bind acetyl-lysine residues for recruitment of larger multisubunit complexes and activation of gene transcription (8, 9). Acetylation of cytoplasmic proteins affects diverse cellular processes which include cell migration, cytoskeleton dynamics, metabolism, and aging. Mitochondrial protein acetylation has been linked to metabolic regulation, oxidative stress, OXPHOS, and mitochondrial gene expression (10).

Reversible acetylation is catalyzed by the enzymatic activity of lysine acetyltransferases (KATs) and deacetylases (KDACs), however mounting evidence suggests that a significant proportion of acetyl-lysine sites are a result of a nonenzymatic mechanism (11–13). We previously measured second order rate constants for nonenzymatic acetylation and found that mimicking the chemical conditions of the mitochondrial matrix *in vitro* could explain a nonenzymatic mechanism, especially for those *in vivo* sites with low acetylation stoichiometry (14). With the number of *in vivo* acetylation sites reaching over 20,000 in human cells (15), a way to prioritize the functional and regulatory acetylation sites is needed, whether or not acetylation is enzyme catalyzed. Therefore, quantifying acetylation stoichiometry would provide critical information to understand regulatory acetylation networks.

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Lysine acetylation has not been associated with signaling cascades, where the acetylation event of one acetyltransferase leads to acetylation of a second acetyltransferase to transmit a biological signal. Such cascades are used to amplify and propagate a signal down a signal transduction pathway. Rather, evidence suggests that acetylation can modulate protein-protein and protein-DNA interactions, cellular localization, enzyme activity and stability (10). In mitochondria, protein acetylation generally serves as an inhibitory mark for enzymatic function. In this regard, acetylation appears to function as a rheostat to modulate the degree of a biochemical process. However, it is not known what level of stoichiometry is needed for a biological effect.

Cells adjust their signaling networks in response to environmental changes. Phosphorylation cascades are well known examples of this response. Cells also respond to external stimuli with dynamic reversible acetylation. For example, upon glucose stimulation, the acetyltransferase, p300, acetylates cytosolic phosphoenolpyruvate carboxykinase (PCK1), which increases the conversion of phosphoenolpyruvate to oxaloacetate, the anaplerotic reaction of PCK1 (16). Activation of cytoplasmic signal transducer and activator of transcription 3 (STAT3) by CBP/p300 upon growth factor and insulin stimulation regulates mitochondrial metabolism (17).

How dynamic acetylation is regulated at the proteome level remains unknown. In this study, we improved upon our method to quantify acetylation stoichiometry at the proteome level, which was benchmarked using cellular proteomes with defined stoichiometry. We quantified steady state as well as dynamic acetylation stoichiometry in response to growth factor stimulation in MCF7 cells, which ranged from less than 1% to 99%. Acetylation stoichiometry distribution across cellular compartments was significantly different. The nuclear compartment, which contains the majority of documented lysine acetyltransferases, displays the highest acetylation stoichiometry, while mitochondrial and cytoplasmic proteins displayed the lowest stoichiometry. RNA binding proteins including ribosomal proteins were enriched in high acetylation stoichiometry. To understand how acetylation and proteome dynamics are concomitantly modulated in cells, MCF7 cells were synchronized by serum depletion/refeeding and monitored for changes in the proteome and site specific acetylation, revealing rapid and dynamic changes in acetylation and protein expression profiles. Quantifying acetylation stoichiometry dynamics will be a critical tool for prioritizing the ever-increasing number of detected lysine acetylation sites and towards a deeper understanding of this regulatory modification.

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

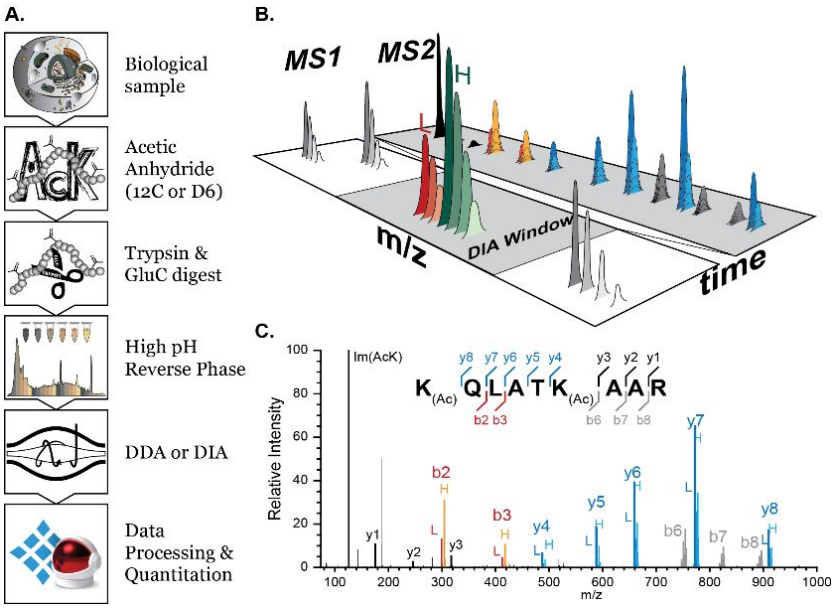
Results

DIA acetylation stoichiometry method optimization

We previously reported a method to determine lysine acetylation stoichiometry across an entire proteome (18). This method employs an isotopic chemical acetylation approach to label all unmodified lysine residues within a sample and, upon proteolytic digestion coupled to LC-MS/MS, has been utilized to quantify proteome-wide acetylation stoichiometry in various biological conditions (18–21) as well as in *in vitro,* nonenzymatic acetylation kinetics (14). Here, we utilize an improved method to quantify acetylation stoichiometry using peptide prefractionation coupled with data-independent acquisition (DIA) mass spectrometry to quantify global acetylation stoichiometry (**Figure 1A**). Briefly, a sample is chemically acetylated using isotopic acetic anhydride and digested with trypsin and GluC. The sequential digestion of the acetyl-proteome generates shorter peptides for MS analysis. Peptides are then prefractionated offline using high pH reverse phase (HPRP) chromatography, analyzed using nano-LC-MS/MS in DIA mode and analyzed using Spectronaut™ v11 (22–25). A project specific spectral library is generated from a ^{12}C-acetic anhydride labeled sample and acquired in data-dependent acquisition (DDA) mode. In order for the spectral library to include both the light and heavy acetyl-lysine fragment ions, a novel, standalone software was developed to be used with Spectronaut™ which generates the heavy labeled fragment ion *in silico* from the ^{12}C-AcK spectral library. This process assures that for a given peptide all heavy and light combinations are contained within the spectral library. Combining offline prefractionation and DIA analysis addresses unique limitations of the original study as previously discussed (21). HPRP prefractionation reduces interferences caused by coeluting peptides and has the added benefit of increasing the depth of the acetylome coverage. DIA is used to measure multiple light and heavy fragment ions of a precursor, which are used for stoichiometry calculation, thus allowing for multiple measurements of a unique lysine site even when multiple lysines are present on a peptide (**Figure 1B, C**).

Figure 1: DIA acetylation stoichiometry workflow

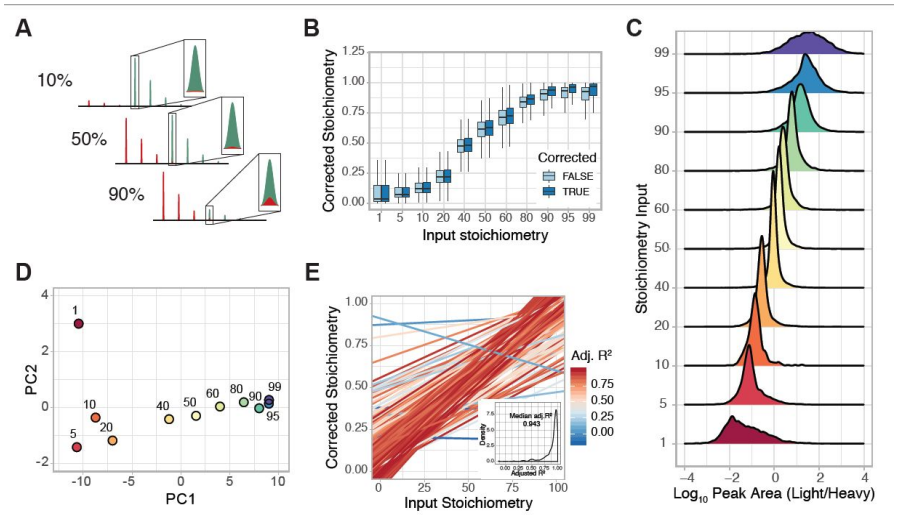
(A) Workflow for measuring acetylation stoichiometry from whole cell lysate. (B) Diagram representing DIA quantitation of precursor and fragment ions across the elution profile. Peak areas for precursor and fragment ions are quantified. (C) Diagram illustrating MS2 spectra for the Histone H3 peptide containing lysine K18Ac and K23Ac. The fragments b2-b3 are specific for K18 and y4-y8 are specific for K23. The fragments b6-b8 are ambiguous as they contain K18Ac and K23Ac, while the fragments y1-y3 contain no acetyl lysine information.



bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To assess the accuracy and precision of our improved workflow, we generated a proteome-wide stoichiometry curve encompassing 1 – 99% acetylation stoichiometry. A whole cell lysate was labeled with light (^{12}C-) and heavy (D_6-) acetic anhydride, mixed at varying ratios, trypsin digested, and subjected to our DIA workflow. Due to the 3 Da mass shift between the light and heavy acetyl peptides, high acetylation stoichiometry will lead to an underestimation of stoichiometry due to increased intensity of the M+3 natural abundance isotopic peak. To account for this, we incorporated a natural abundance correction to all heavy acetyl lysine fragment ions. This correction was performed by subtracting the M+3 isotopic peak of the light acetyl lysine fragment ion from the M+0 isotopic peak of the heavy acetyl lysine fragment ion (**Figure 2A**). This global correction improved the precision of the stoichiometry quantification, especially in the higher stoichiometry values (**Figure 2B**). An alternative method to assess the precision of the quantification is to measure the ratio of the light and heavy fragment ions. Quantification of the light/heavy ratios corresponding to stoichiometry profiles between 20 and 80% displayed the highest precision (**Figure 2C**). This is due to the abundance values of the light and heavy fragment ions near 1:1 ratio. In contrast, stoichiometries at the extreme ends of the curve (< 5% and > 95%) displayed the lowest precision since quantitation in these conditions require the measurement of fragment ions greater than 20-fold difference (**Figure 2C**). Principal component analysis (PCA) also detects the increase in stoichiometry across the experimental conditions (**Figure 2D**). The first principal component (PC1) represents the variability associated with the input stoichiometry, and, as expected, increases in a linear fashion. The stoichiometry curve analysis quantified stoichiometry for ~1400 acetyl lysine sites. The number of acetyl sites quantified in at least nine conditions was 616. Linear regression analysis using acetyl lysine sites quantified in at least nine conditions shows high reproducibility of our method (**Figure 2E**) with a median R^2 of 0.94, after correction for multiple regression analysis (**Figure 2E inset**). This global analysis with well defined input stoichiometries highlights the quantitative nature of this method and is applicable to query acetylation stoichiometry of an entire proteome.

Figure 2: Benchmarking the DIA stoichiometry workflow

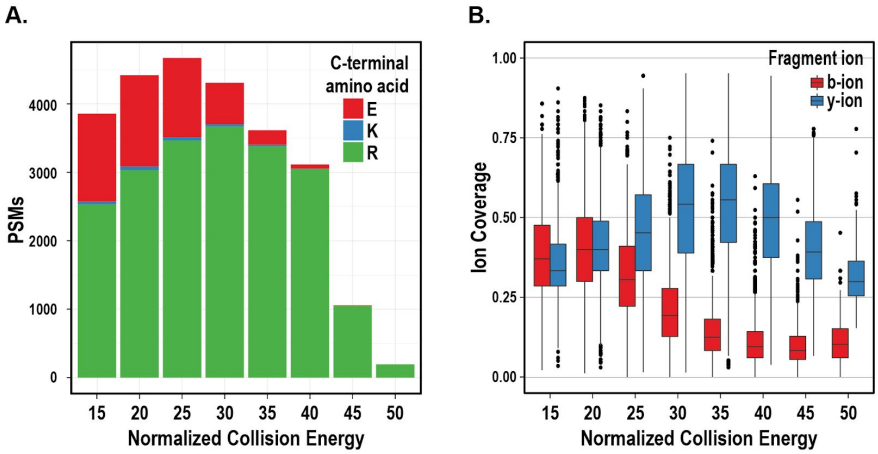


(A) Correction of natural abundance isotopic envelope between light and heavy acetyl peptides. Isotopic overlap is much more pronounced at high stoichiometry. (B) Boxplot of the proteome-wide stoichiometry distribution before and after natural abundance isotope correction. (C) Density plot of the Log_{10}-transformed (light / heavy) peak area for each input stoichiometry condition. (D) Principal Components Analysis of stoichiometry curve. (E) Linear regression analysis of the all acetyl lysine sites (> 9 data points, n = 616) with the color of each line corresponding to the R^2 value. This analysis reveals the high correlation between the measured and expected stoichiometry as the majority of AcK sites display high R^2 values (inset).

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Our improved stoichiometry workflow enables the quantitation of different lysines from a single peptide, removing any ambiguity of site quantification. As an example, the histone H3 peptide (containing K18 & K23), $K_{Ac}QLATK_{Ac}AAR$, has fragment ions that are unique to each lysine site. K18 is quantified by the fragment ions b2-b3, while y4-y8 are specific for K23 (**Figure 1C**). Obtaining high quality and high coverage of b- and y-ions will provide quantification of multiple lysines on the same peptide. Therefore, we optimized the normalized collision energy (NCE) for higher energy collisional dissociation (HCD) fragmentation (26) and evaluated the number of peptide spectral matches (PSMs) (**Figure 3A**) as well as the global b- and y-ion coverage (**Figure 3B**) across a wide range of NCEs (15-50 in 5 unit increments) using a chemically acetylated, trypsin and GluC digested proteome. To determine the global b- and y-fragment ion coverage, we counted each fragment ion identified for a given PSM and normalized to the peptide length. As y-ions increase with higher NCE, the proportion of b-ions begin to decline at a similar rate (**Figure 3B**). The NCE 25 was used to balance the frequency of b-ions, y-ions, as well as the number of PSMs (**Figure 3A**). A negligible amount of PSMs with a c-terminal lysine are seen (blue bar) demonstrating our ability to chemically acetylate the entire proteome to near completion.

Figure 3: Optimization of instrument parameters for increased fragment ion coverage
(A) Peptide spectral matches (PSMs) at varying normalized collision energy (NCE) for trypsin (c-terminal arginine and lysine) and GluC (c-terminal glutamate) digested peptides. The data consists of an MCF7 whole cell lysate labeled with ^{12}C-acetic anhydride, digested with trypsin and GluC and analyzed in DDA mode using varying NCE settings. (B) Global fragment ion coverage for b- and y-ions of the data in A.



bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Subcellular distribution of acetylation stoichiometry

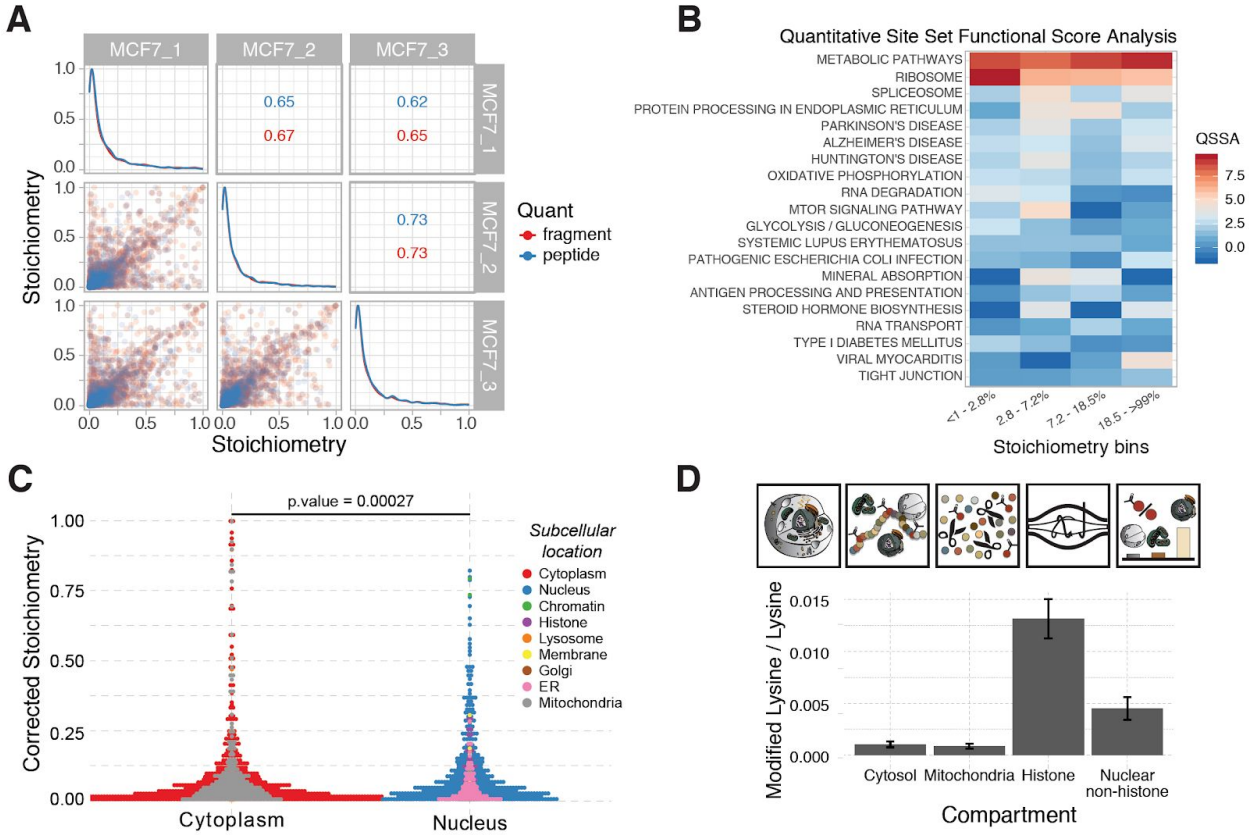
Studies measuring (or estimating) acetylation stoichiometry in mammalian systems are limited and most have focused on the mitochondrial or nuclear compartments of the cell (20, 27). Thus, there has not been a clear picture of the acetylation stoichiometry distribution across the cell. To address this, we utilized our quantitative stoichiometry approach using MCF7 cells, a breast cancer cell line and quantified a wide range of stoichiometry (< 1% up to 97%) with high correlation between acetyl lysine fragment ions (red) and peptides (blue) between the three biological replicates (**Figure 4A**). To identify biological processes that are enriched in acetylation, we applied a biological pathway analysis tool that we recently developed termed quantitative site set functional score analysis (QSSA) to our acetylation stoichiometry dataset (28). This tool was developed for PTM pathway enrichment analysis taking into account the number of modified sites in a given pathway as well as the relative change across conditions (fold-change) and was adapted for acetylation stoichiometry datasets. The stoichiometry from MCF7 cells was divided into quartiles (for stoichiometry ranges, see Materials and Methods). Each quartile was used as input for the QSSA. Gene Ontology processes that were enriched in this study include Metabolic Pathways, Ribosome, Spliceosome, and Protein Processing in Endoplasmic Reticulum (**Figure 4B**). Enrichment of metabolic pathways is a hallmark of previous acetylation studies (3, 28, 29) and processes associated with RNA protein complexes have been previously identified in a wide range of organisms including *E. coli*, *B. subtilis*, *P. falciparum*, and *S. cerevisiae* (18, 30–32), confirming our QSSA results.

Mass spectrometry is inherently biased towards quantifying high abundant peptides and proteins. Therefore, we wanted to quantify the levels of acetylation where high as well as low abundant proteins were measured, specifically within subcellular compartments. To do this, we performed a differential centrifugation and acid extraction on MCF7 cells to enrich for histone proteins, non-histone nuclear fraction, mitochondria, and the cytosolic fraction. Each fraction was treated with a combination of proteases in order to digest the proteins of each subcellular compartment to individual amino acids. The relative abundance of acetyl-lysine and unmodified lysine can be measured using mass spectrometry (33) (**Figure 4C**). This approach reduces sample complexity and provides a global and unbiased view of the entire pool of modified and unmodified lysine residues. Acetylation is significantly enriched on histone proteins and the nuclear fraction compared to the cytoplasm and mitochondrial fractions (**Figure 4D**). We then asked whether our acetylation stoichiometry data agreed with the targeted lysine amino acid quantification results. The number of acetyl-lysine containing peptides corresponding to each of the subcellular fractions analyzed were disproportionately measured. We therefore grouped each protein accession into a cytoplasmic or nuclear fraction. Using this grouping and a non-parametric analysis, the nuclear fraction contains more acetylation sites with a higher stoichiometry compared to the cytoplasmic fractions ($p = 0.00027$) (**Figure 4E**).

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4: Nuclear-localized proteins have highest lysine PTM abundances

(A) Scatterplot matrix of acetylation stoichiometry across the three biological replicates of MCF7 cells. Pearson correlation is displayed for each pairwise comparison of fragment ions (red) acetyl-lysine site stoichiometry (blue). (B) QSSA heatmap showing enriched KEGG pathways. Acetylation stoichiometry was binned into quartiles and used as input for the QSSA. (C) Acetylation stoichiometry distribution across cellular compartments. Compartments were grouped *in silico* as cytoplasmic or nuclear. Statistical analysis was performed using the posthoc kruskal nemenyi test. Colored circles represent individual subcellular location assignment based on Uniprot location and mitocarta database (34, 35). (D) Workflow for quantifying single amino acid abundances of modified and unmodified lysines of enriched subcellular fractions. Each fraction is digested with a cocktail of proteases to generate single amino acids. Amino acids are analyzed by MS to quantify modified lysine / unmodified lysine as a measure of global subcellular stoichiometry. (E) acetyl-lysine:lysine ratio across cytoplasmic, mitochondrial, histone, and non-histone proteins.



bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acetylation and proteome dynamics

Protein acetylation can be modulated by the enzymatic activity of acetyltransferases, deacetylases and, non-enzymatically, by the levels of acetyl-CoA (10, 29, 36). Additionally, changes in acetylation stoichiometry can be observed in conjunction with changes in protein abundance. Therefore, understanding dynamics of not only protein acetylation, but also proteome dynamics will be important to understand the interplay between these two processes. The majority of studies quantifying acetylation dynamics utilize an antibody based workflow to enrich for the acetyl-peptides (37, 38). Using an enrichment strategy, it is necessary to account for changes in protein abundance in order to accurately report changes in acetylation by analyzing a sample of the proteome which was not subjected to the immuno-enrichment procedure. Our protocol does not utilize an enrichment step. Instead, we chemically modify all free lysine residues using acetic anhydride, a step which is analogous to the alkylation of cysteine residues with iodoacetamide. Therefore, precursor abundance data collected from our acetylation stoichiometry workflow can also be used to estimate protein abundance using label free quantification techniques.

To understand cellular protein and acetylation dynamics, a serum starve-replete cell culture model system was used which has the benefits of synchronizing the cells upon serum starvation followed by robust changes in signaling pathways that occur upon serum replenishment (17, 39, 40). MCF7 cells were grown to 70-80% confluency and serum starved for 24 hours. Cells were then replenished with fresh media containing serum and harvested at 0, 1, 2, 4 hours (**Figure 5A**). To verify activation of major signaling pathways, we monitored the level of phosphorylation of ribosomal protein S6 (**Figure 5B, C**), as a proxy for mTOR signaling activation (39).

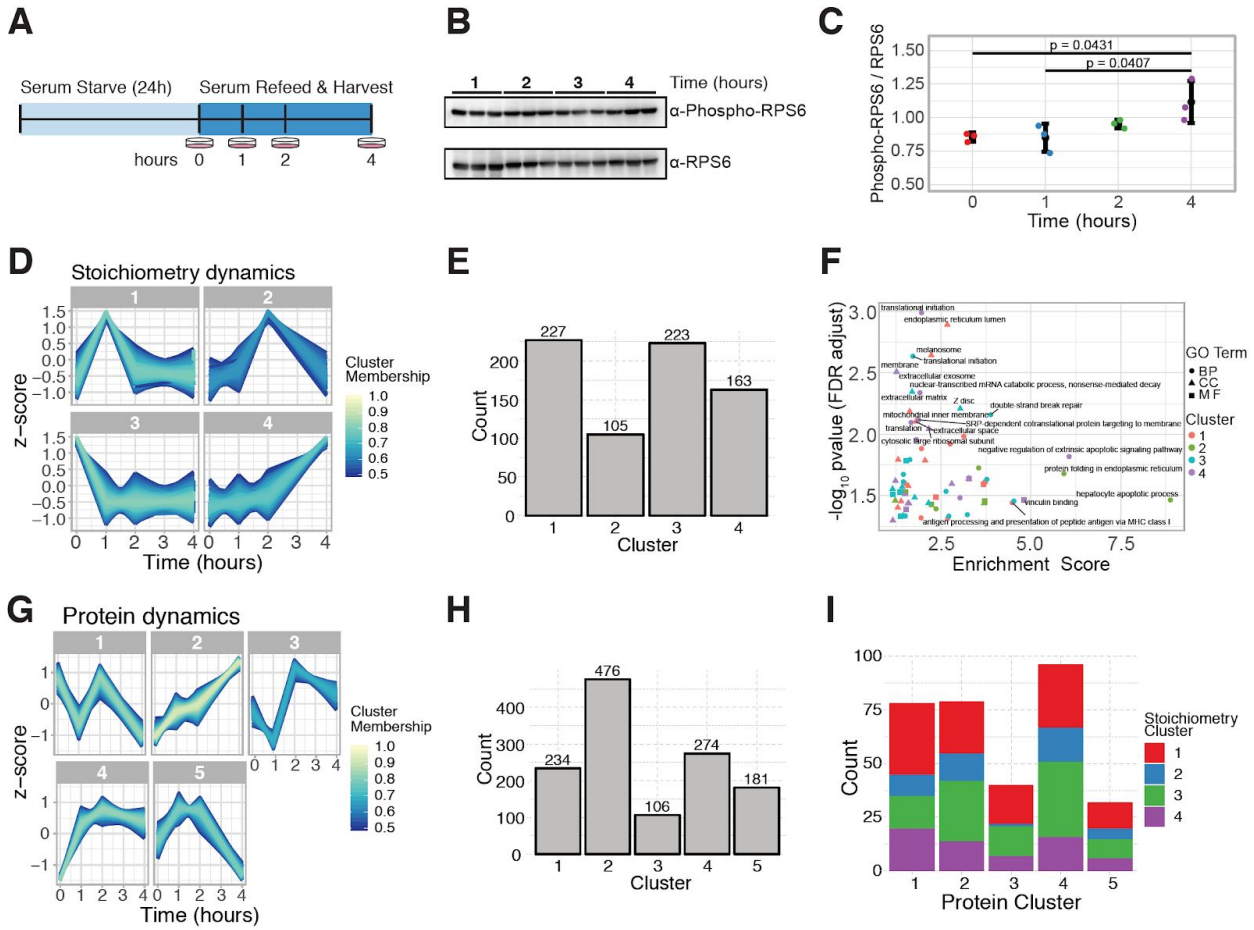
Acetylation stoichiometry was quantified using our DIA-MS approach followed by a pattern recognition analysis using fuzzy c-means clustering (41). This analysis identified four distinct clusters where acetylation is dynamic with varied profiles (**Figure 5D, E**). Over two thirds of the acetylation sites identified in this clustering analysis were found in clusters 1 and 3 which display rapid changes upon growth factor stimulation. These clusters correspond to acetylation levels that rapidly increased and returned to pretreatment baseline levels (Cluster 1) as well as trends where acetylation rapidly decreased and remained at a consistent level (Cluster 3), respectively. Gene ontology (GO) enrichment analysis was performed on the acetylation stoichiometry clusters demonstrating an enrichment of biological processes associated with cell growth such as transcription and translation (**Figure 5F**). Protein abundance was determined by label free quantification using MSstats with our chemically acetylated proteome sample (42) followed by the clustering analysis using fuzzy c-means (**Figure 5G, H**) (41).

To investigate the co-trends of protein and acetylation dynamics, we compared the trends of acetylation with protein clusters (**Figure 5I**). Observed trends in protein acetylation can be the result of several biochemical processes such as: acetylation, deacetylation, changes in protein abundance or any combination of these processes. For example, an increase in protein abundance without lysine acetylation (or deacetylation) would result in a trend showing a decrease in acetylation stoichiometry. In fact, we do observe this relationship when comparing acetylation cluster 3 and protein clusters 2 and 4. Lastly, acetylation stoichiometry cluster 4 represents acetylation sites which increase over time, which follows the same trend as the phosphorylation of ribosomal protein S6 (**Figure 5C**).

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5: Acetylation and protein dynamics

(A) Diagram of experimental approach to quantify acetylation stoichiometry dynamics. (B) Western blot of phosphorylated Ribosomal Protein S6 (top) and Ribosomal Protein S6 (bottom). (C) Quantitation of the western blot. Statistical analysis was performed using an ANOVA. (D) Time-course clusters of acetylation stoichiometry dynamics by fuzzy c-means clustering, and (E) number of AcK sites per cluster. (F) GO Term enrichment analysis of each stoichiometry cluster, represented is fold enrichment versus significance. (G) Time-course clusters of protein abundance dynamics by fuzzy c-means clustering, and (H) number of proteins per cluster. (I) Barplot quantifying the frequency of acetylation dynamic clusters within each acetylation dynamics cluster.



bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

Integrated stoichiometry workflow

We previously developed a method to quantify acetylation stoichiometry in an entire proteome, where a protein extract is chemically modified with an isotopic acetic anhydride, digested, and analyzed by high resolution mass spectrometry (18). In this study, we integrated our workflow for quantifying acetylation stoichiometry with Spectronaut™ for the targeted analysis of the light and heavy acetyl fragment ions using a novel spectral library. A spectral library which contains all light and heavy acetyl-lysine feature pairs was generated. The samples for the project specific spectral library were generated by chemically labeling a sample with ^{12}C-acetic anhydride followed by proteolytic digestion and analysis by DDA-MS. The heavy acetyl fragment ions were generated in silico using a 3 Dalton shift. This step ensures that all light and heavy acetyl lysine pairs are used for the targeted analysis of acetylation stoichiometry.

Acetylation stoichiometry distribution in the cell

Using our integrated stoichiometry workflow, we quantified the distribution of acetylation stoichiometry across MCF7 cells, observing a significantly higher proportion of proteins with higher acetylation stoichiometry within the nuclear compartment. Binning the stoichiometry into equal sized quartiles and performing a PTM pathway enrichment analysis (28), we found that metabolic pathways were enriched for all levels of acetylation. Additionally, KEGG pathways associated with RNA binding such as the Ribosome, Spliceosome and RNA Degradation display a high QSSA enrichment score. It was recently shown that the lysine acetyltransferases, NAT8/NAT8B, are localized in the lumen of the endoplasmic reticulum and function to acetylate properly folded proteins as proteins traverse through the secretory pathway (43). NAT8/NAT8B use acetylation to signal correctly folded proteins. In this context, acetylation stoichiometry would be expected to be higher due to the localization of the acetyltransferases. Indeed, our QSSA results identify higher acetylation stoichiometry in this cellular compartment. Additionally, amino acid quantification from digested subcellular proteomes agreed with the site-specific acetylation distribution. Interestingly, well-characterized acetyltransferases reside in the nuclear compartment, highlighting the strong regulatory role of epigenetic and gene expression systems.

Integrated analysis of protein and acetylation dynamics

Understanding how acetylation is regulated in the cell is an active area of investigation and quantifying the dynamics is a critical component to understand this regulatory modification. To understand dynamic acetylation and signaling events, we quantified acetylation and protein dynamics upon serum starvation followed by growth factor stimulation. Acetylation and protein groups were clustered using fuzzy c-means, which allows a protein (or lysine site) to belong to multiple groups, reflecting the multiple functions of proteins within a cell. We identified several patterns in acetylation dynamics upon serum replenishment with the following trends. Acetylation cluster 1, acetyl sites with baseline stoichiometry, increase rapidly upon serum repletion followed by a return to baseline. Acetylation cluster 3 are acetyl sites that begin elevated, decrease rapidly and remain low. Finally, acetylation cluster 4 are acetyl sites which begin at a baseline and increase steadily. We quantified an acetylation site of ribosomal protein S6 (K64Ac) within cluster 4. Interestingly, the phosphorylation level of ribosomal protein S6 also appears very similar to the profile of acetylation cluster 4 (**Figure 5C**). Whether there is any cross-talk between acetylation and phosphorylation of ribosomal protein S6 is unclear. The method described in this study provides a valuable way to decipher regulatory acetylation

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

events. Determining proteome-wide acetylation kinetics will help decode functional vs. spurious acetylation events and prioritize those regulatory lysine sites.

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Cell Culture conditions

MCF7 cells were grown using DMEM supplemented with 10% FBS. For global acetylation stoichiometry and single amino acid analysis, MCF7 cells were harvested at ~80% confluency. Four hours prior to harvesting, cells were washed with PBS and replaced with fresh media. MCF7 cells were cultured for a total of 48 hours.

Sample preparation

Protein chemical acetylation and digestion

Equal amount of protein (200 µg) was resuspended into 25-30 µL of urea buffer (8 M urea (deionized), 500 mM ammonium bicarbonate pH = 8.0, 5 mM DTT). Incubation steps throughout the sample preparation are carried out using the Eppendorf ThermoMixer® C. Sample was incubated at 60 °C for 20 minutes while shaking at 1500 RPM. Cysteine alkylation was carried out with 50 mM iodoacetamide and incubating for 20 minutes. Chemical acetylation of unmodified lysine residues was performed as previously described (14, 18, 33). Briefly, ~20 µmol of the "light" ^{12}C-acetic anhydride (Sigma) or "heavy" D_6-acetic anhydride (Cambridge Isotope Laboratories) was added to each sample and incubated at 60 °C for 20 minutes at 1500 RPM. The pH of each sample was raised to ~8 using ammonium hydroxide and visually checked with litmus paper. Two rounds of chemical acetylation were performed for each sample to ensure near complete lysine acetylation. To hydrolyze any O-acetyl esters formed during the chemical acetylation, the pH of the sample raised to ~8.5 and each sample was incubated at 60 °C for 20 minutes at 1500 RPM. For protein digestion, the urea concentration of each sample was diluted to ~2 M by adding 100 mM ammonium bicarbonate pH = 8.0 followed by addition of trypsin (Promega) at a final ratio of 1:100. The sample was digested at 37 °C for 4 hours while shaking at 500 RPM. If a second digestion using gluC (Promega) occurred, the urea concentration was further diluted to ~1 M using 100 mM ammonium bicarbonate pH = 8.0 and digested with gluC (1:100) at 37 °C overnight while shaking at 500 RPM. Each sample was acidified by addition of 15 µL of acetic acid.

Digesting protein sample to single amino acids

For complete digestion of proteins, which converts all unmodified lysine residues to free lysine, and all N-ε-acetylated lysine residue to acetyl-lysine, 20 µg of sample was diluted into 50 µL of digestion buffer (50 mM ammonium bicarbonate, pH 7.5, 5 mM DTT, in LC–MS grade water). A sample with 50 µL digestion buffer without protein was also included as procedural blank. The samples were digested to single amino acids by treatment with three enzymes sequentially: First, samples were treated with 0.4 µg Pronase and incubated for 24 hr at 37°C. Then the Pronase activity was stopped by heating to 95°C for 5 min. After cooling down to ambient temperature, samples were then treated with 0.8 µg aminopeptidase and incubated at 37°C for 18 hr. Aminopeptidase activity was again stopped by heating samples to 95°C for 5 min and cooling down. Finally, samples were digested with 0.4 µg prolidase and incubated at 37°C for 3 hr. To extract the resulted single amino acids, 200 µL LC–MS grade acetonitrile (ACN) was added to each sample. The mixture was vortexed for 5 sec, spun at maximal speed for 5 min, and supernatant was saved for analysis by LCMS.

Offline High pH Reverse Phase (HPRP) Prefractionation

Chemically acetylated peptides were resuspended into ~2mL of HPRP buffer A (100 mM Ammonium Formate pH = 10) and injected onto a preequilibrated Phenomenex Gemini® NX-C18 column (5µm, 110 Å, 150 x 2.0mm) with 2% buffer B (10% Buffer A, 90% acetonitrile). Peptides were separated with a Shimadzu LC-20AT HPLC system using a 2% - 40% buffer B linear gradient over 30 minutes at 0.6 mL/min flow rate, collecting 24 fractions throughout the length of the gradient. Fractions

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were dried down using a speedvac and pooled by concatenation into 6 final fractions as described previously (44).

Mass spectrometry

Liquid chromatography

Peptides were separated with a Dionex Ultimate 3000 RSLCnano HPLC using a Waters Atlantis dC18 (100 µm x 150 mm, 3µm) C18 column. The mobile phase consisted of 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). Peptides were eluted with a linear gradient of 2 – 35% B at a flow rate of 800 nL/min over 90 minutes. Peptides were injected by nanoelectrospray ionization (Nanospray Flex™) into the Thermo Fisher Q Exactive™ Hybrid Quadrupole-Orbitrap™ Mass spectrometer.

Data-dependent acquisition mass spectrometry

For data dependent acquisition (DDA), the MS survey scan was performed in positive ion mode with a resolution of 70,000, AGC of 3e6, maximum fill time of 100 ms, and scan range of 400 to 1200 m/z in profile mode. Data dependent MS/MS was performed in profile mode with a resolution of 35,000, AGC of 1e6, maximum fill time of 200 ms, isolation window of 2.0 *m/z*, normalized collision energy of 25, dynamic exclusion was set for 30 seconds, and a loop count of 20.

Data-independent acquisition mass spectrometry

For data-independent acquisition (DIA), the MS survey scan was performed in profile mode with a resolution of 70,000, AGC of 1e6, maximum fill time of 100 ms in the scan range between 400 and 1000 *m/z*. The survey scan was followed 30 DIA scans in profile mode with a resolution of 35,000, AGC 1e6, 20 *m/z* window, and NCE of 25 or 30. For both DDA and DIA methods, the source voltage was set at 2000 V and capillary temperature at 250 °C.

LCMS analysis of single amino acids

The abundances of free lysine, acetyl-lysine, and other amino acids from completely digested protein samples were analyzed using a Thermo Fisher Q Exactive™ Hybrid Quadrupole-Orbitrap™ Mass spectrometer coupled to a Dionex UltiMate 3000 UHPLC system. Samples are separated using a 5 µm polymer 150 2.1 mm SeQuant® ZIC®-pHILIC column, with the following gradient of solvent A (ACN) and solvent B (10 mM ammonium acetate in water, pH 5.5) at a flow rate of 0.3 mL/min: 0-2min, 10% solvent B; 2-14min, linearly increase solvent B to 90%; 14-17min, isocratic 90% solvent B ; 17-20min, equilibration with 10% solvent B. Samples are introduced to the mass spectrometer by heated electrospray ionization. Settings for the ion source are: 10 aux gas flow rate, 35 sheath gas flow rate, 1 sweep gas flow rate, 3.5 kV spray voltage, 320°C capillary temperature, and 300°C heater temperature. Analysis is performed under positive ionization mode, with scan range of 88–500 m/z, resolution of 70 K, maximum injection time of 40 ms, and AGC of 1E6.

To quantify absolute levels of lysine and acetyl-lysine, an external calibration curve was run in the same sequence with the experimental samples. Lysine standard ranges between 10 to 200 µM, and acetyl-lysine standard ranges between 0.5 to 10 µM. Signal from procedural blank was subtracted from samples.

Data Processing

Generating ^{12}C-AcK and D_3-AcK Spectral Library

The spectral library consists of a catalogue of high quality MS/MS fragmentation spectra resulting from data-dependent acquisition (DDA) MS runs. For our workflow, we performed DDA runs on an MCF7 lysate which was chemically acetylated with ^{12}C-acetic anhydride, digested with trypsin and gluC, followed by HPRP prefractionation (see above). Prior to MS analysis, iRT peptides (Biognosys) were spiked into each sample following manufacturer's guidelines. Database search was performed using MaxQuant version 1.5.4.1 using lysine acetylation and methionine oxidation as variable modifications

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and cysteine carbamidomethylation as fixed modification. The MaxQuant search results were imported into Spectronaut to build the 12C-AcK library. The 12C-AcK spectral library was then exported as a spreadsheat from Spectronaut and imported into a custom spectral library modifier, which completes the spectral library for all combinations of light and heavy acetylated peptides. With this *in silico* approach to inflate the 12C-AcK library, every acetylated peptide precursor will be represented by 2^n versions differing in the number and position of heavy/light acetylated lysine, where n is the number of acetylation sites in the peptide. The spectral library was completed with the corresponding precursor m/z values and fragment m/z values. The most intense fragment ions selected from the initial MS2 spectrum were cloned to the other peptide precursor versions. All peptide precursor versions will have identical retention time and hence iRT was also cloned.

DIA MS data analysis

Data from DIA-MS was analyzed using Spectronaut 10. Thermo raw files were converted to HTRMS files with the Spectronaut Raw to HTRMS converter using the default settings and input into Spectronaut. The Spectronaut default settings for quantitation were used with slight modification: Identification-Qvalue score was set to 0.1 and Workflow-Unify peptide peaks was selected. This will cause Spectronaut to use the same integration boundaries for all light/heavy versions of one acetylated peptide within one LC-MS run. This change in workflow will instruct Spectronaut to select for a given acetylated peptide precursor the best signal (by q-value) of the 2^n versions in the spectral library (see above). With this workflow Spectronaut will then transfer the integration boundaries of the best scoring peptide precursor to the other peptide precursors. Because all of the 2^n peptide precursor versions only differ by the number of heavy instead of light acetylated lysine the retention time is expected to be identical. The spectral libraries which were completed as described above for all the light/heavy peptide precursor versions were used with this workflow. A Spectronaut output file containing all the fragment ion peak areas along with the corresponding peptide and protein identification was exported and used to compute the lysine site stoichiometry. A list containing all the data categories used for downstream stoichiometry analysis is found in supplemental information.

Stoichiometry data processing

Data processing was performed in R v3.5.0 (http://www.r-project.org/) using an in-house made R script, which is available in the supplementary information. The stoichiometry preprocessing pipeline consists of two major steps: quantifying fragment ion stoichiometry and natural abundance isotopic correction.

Quantifying site-specific stoichiometry

DIA MS measures multiple peptide fragment ion abundances so our approach allows for quantitation of multiple lysines within a peptide. Acetylation stoichiometry of unique lysine sites are quantified by matching light and heavy fragment ion pairs and using the equation:

$$\frac{XIC_L}{XIC_L + XIC_H} \qquad \text{Equation 1}$$

where XIC_L is the peak area of the light fragment ion and XIC_H is the peak area of the heavy fragment ion.

Isotopic purity correction

The mass shift of the light and heavy AcK peptides is 3 Da. This causes the M+0 peak of the heavy AcK peptide to overlap with the M+3 peak of the light AcK peptide. Therefore, we are correcting for the isotopic distribution overlap between the peptide pairs. This is done using an in-house R script as well as the R package, BRAIN v1.16.0 (Baffling Recursive Algorithm for Isotopic DistributioN calculations), available from Bioconductor, the open source, software project (http://www.Bioconductor.org/) (45). To

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

correct for natural abundance of ^{13}C isotope, the M+0 and M+I, where I represents the isotopic mass shift +1 or +3, were used to calculate the correction coefficient.

$$\text{Correction coefficient} = \frac{1}{\left(\frac{M+0}{M+I}\right)} \qquad \text{Equation 2}$$

The correction coefficient is used to calculate the correction value:

$$\text{Correction value} = XIC_L * \text{correction coefficient} \qquad \text{Equation 3}$$

where the XIC_L is the peak area of the light fragment ion. Finally, the corrected heavy peak area ($^{Corr}XIC_H$) is calculated:

$$^{Corr}XIC_H = XIC_H - \text{Correction value} \qquad \text{Equation 4}$$

where the XIC_H is the peak area of the heavy fragment ion. The corrected stoichiometry is quantified using equation 1, substituting with $^{Corr}XIC_H$.

MSstats

Protein abundance summarization was performed using MSstats v3.12.0 with the output of Spectronaut as the input. The function "SpectronauttoMSstatsFormat" was used with the following arguments: intensity set to "PeakArea", filter_w_Qvalue set to TRUE, qvalue_cutoff set to 0.01, useUniquePeptide set to TRUE, fewMeasurements set to "remove", removeProtein_with1Feature set to FALSE, and summaryforMultipleRows set to "max". The dataProcesses function was then performed using the default arguments.

NCE Optimization

To quantify site-specific acetylation stoichiometry from peptides containing multiple lysines, the fragmentation spectra of precursor ions must contain a high b- and y-ion coverage. To this end, we compared and optimized the number of peptide spectral matches (PSMs) as well as b- and y-ion coverage of MCF7 peptides (chemically acetylated with ^{12}C-acetic anhydride followed by trypsin and gluC digestion) with a Q-Exactive MS using varying NCE settings (15, 20, 25, 30, 35, 40, 45, 50). For all NCE conditions, precursors between 400 - 1200 *m/z* were selected for fragmentation. MS1 resolution was set to 70,000, 3e6 target AGC, and 100 ms max IT in profile mode. MS2 resolution was set to 35,000, 1e6 target AGC, 200 ms max IT in profile mode with 15 sec dynamic exclusion. Database search was performed using MaxQuant version 1.5.4.1 followed by data analysis in R.

Stoichiometry curve

We determined the accuracy and precision of our stoichiometry method by generating an 11-point stoichiometry curve using a complex sample. For this, we used a HEK293 lysate that was grown using standard culture conditions and harvested by centrifugation. The packed cell volume was resuspended using urea buffer (6-8M urea, 100mM ammonium bicarbonate pH = 8.0) and lysed by sonication. Protein concentration was measured using Bradford reagent (Bio-Rad).

To quantify stoichiometry ranging between 1-99%, we varied the amount of starting material to be chemically acetylated with ^{12}C-acetic anhydride or D_6-acetic anhydride using a total of 200 µg of protein for each stoichiometry point. For example, to measure a sample as 10% acetylated, we labeled 20 µg of HEK293 lysate with ^{12}C-acetic anhydride and 180 µg of HEK293 lysate with D_6-acetic anhydride. The starting protein amounts were varied to generate stoichiometries of: 1, 5, 10, 20, 40, 50, 60, 80, 90, 95, and 99% acetylation. Upon chemical acetylation, the sample was pooled together, digested using trypsin and we performed an offline HPRP prefractionation as outlined above.

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bioinformatics

Functional annotation

Functional annotation of enriched gene ontology (GO) terms were assessed using DAVID v6.8 (46, 47). For enrichment analysis, the background was set to all the proteins identified in the DDA spectral library totaling 2400 unique protein IDs. The list of high-confidence acetylation stoichiometries, which consist of lysine sites with at least 2 fragment ion quantitation values ± 5% stoichiometry, were ordered from lowest to highest stoichiometry and split into 5 equal quantiles: (A < 0.0095, n = 97), (B 0.0096 - 0.0246, n = 97), (C 0.0247 - 0.0464, n = 96), (D 0.0465 - 0.2741, n = 96), (E 0.2747 - 0.9704, n = 96). Each quantile was analyzed using DAVID for the following GO terms: GOTERM_BP_DIRECT, GOTERM_CC_DIRECT, and GOTERM_MF_DIRECT. A Benjamini Hochberg-corrected pvalue < 0.05 was used corresponding to < 1% FDR for all GO terms. To generate the GO Term scatter plot, we used the -Log10 transformed pvalue, fold enrichment, and coverage from the DAVID output. The scatter plot was generated in R with the ggplot2 v2.1.0 package.

Subcellular localization assignment

To assign protein subcellular localization, we used the MitoCarta (34, 35) and Uniprot (http://www.uniprot.org/) databases. For "Mitochondrial" assignment of proteins, we used the Mitocarta database. Additionally, we used "Subcellular location" or "GO - Cellular component" from the Uniprot database to assign "Mitochondrial", "Nuclear", and "Cytoplasmic" pools. Other subcellular locations, such as endoplasmic reticulum, golgi apparatus, cell membrane, etc., were assigned to the "Nuclear" fraction due to the likelihood that these cellular compartments, during differential centrifugation, would sediment in the "Nuclear" spin, which occurs at 1000 xg.

Quantitative Site set functional Score Analysis (QSSA)

The intersection of the KEGG pathway map (48) and proteins in the spectral library detected with < 1% FDR was used for the gene set background. Acetylation coverage for each (p) pathway was calculated as the number of acetyl sites identified (n_{ack}) over the total number of lysines in the pathway (n_k), counted using protein sequences from Uniprot. The extent of acetylation was taken into account by summing the acetylation stoichiometry (s) across all conditions and all sites in each pathway. To allow for combining acetylation coverage and stoichiometry, the standard score of each quantity was taken. The overall pathway score was then calculated as the sum of the individual z-scores.

$$QSSA_p = z\left(\frac{n_{acK}}{n_K}\right) + z\left(\sum_{i \in p} s_i\right)$$

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Author contributions: J.B., J.F., and J.M.D. designed the research study. J.B. and A.L. performed acetylation stoichiometry mass spectrometry experiments and analyzed the data. J.F. performed single amino acid mass spectrometry experiments and analyzed data. I.L., T.G., O.M.B., and L.R. developed software to analyze DIA MS data as well as for custom modification of the library. J.B., A.L. and M.J.S. performed bioinformatic analysis. J.B. and A.L. designed the figures and drafted the manuscript. J.B., A.L., J.F., M.J.S., L.R. and J.M.D edited manuscript. J.M.D. is the corresponding author.

Acknowledgements: We would like to thank Greg Barrett-Wilt and Greg Sabat at the University of Wisconsin-Madison Biotechnology center for use of the Mascot database server. This work was supported, in whole or in part, by National Institutes of Health (NIH) Grant GM065386 (J.M.D.), NIH National Research Service Award T32 GM007215 (J.B. and A.L.) and the National Science Foundation Graduate Research Fellowship Program (NSF-GRFP) DGE-1256259 (J.B.).

Additional Information

The authors I.L., T.G., O.M.B. and L.R. are employees of Biognosys AG (Zurich, Switzerland).
Spectronaut is a trademark of Biognosys AG.

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Kouzarides, T. (2000) Acetylation: a regulatory modification to rival phosphorylation? *EMBO J.* 19, 1176–1179

2. Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964) ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proc. Natl. Acad. Sci. U. S. A.* 51, 786–794

3. Verdin, E., and Ott, M. (2015) 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat. Rev. Mol. Cell Biol.* 16, 258–264

4. Close, P., Creppe, C., Gillard, M., Ladang, A., Chapelle, J.-P., Nguyen, L., and Chariot, A. (2010) The emerging role of lysine acetylation of non-nuclear proteins. *Cell. Mol. Life Sci.* 67, 1255–1264

5. Anderson, K. A., and Hirschey, M. D. (2012) Mitochondrial protein acetylation regulates metabolism. *Essays Biochem.* 52, 23–35

6. Pehar, M., and Puglielli, L. (2013) Lysine acetylation in the lumen of the ER: a novel and essential function under the control of the UPR. *Biochim. Biophys. Acta* 1833, 686–697

7. Pougovkina, O., te Brinke, H., Ofman, R., van Cruchten, A. G., Kulik, W., Wanders, R. J. A., Houten, S. M., and de Boer, V. C. J. (2014) Mitochondrial protein acetylation is driven by acetyl-CoA from fatty acid oxidation. *Hum. Mol. Genet.* 23, 3513–3522

8. Filippakopoulos, P., and Knapp, S. (2014) Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat. Rev. Drug Discov.* 13, 337–356

9. Josling, G. A., Selvarajah, S. A., Petter, M., and Duffy, M. F. (2012) The role of bromodomain proteins in regulating gene expression. *Genes* 3, 320–343

10. Baeza, J., Smallegan, M. J., and Denu, J. M. (2016) Mechanisms and Dynamics of Protein Acetylation in Mitochondria. *Trends Biochem. Sci.* 41, 231–244

11. Olia, A. S., Barker, K., McCullough, C. E., Tang, H.-Y., Speicher, D. W., Qiu, J., LaBaer, J., and Marmorstein, R. (2015) Nonenzymatic Protein Acetylation Detected by NAPPA Protein Arrays. *ACS Chem. Biol.* 10, 2034–2047

12. Wagner, G. R., and Payne, R. M. (2013) Widespread and enzyme-independent Nε-acetylation and Nε-succinylation of proteins in the chemical conditions of the mitochondrial matrix. *J. Biol. Chem.* 288, 29036–29045

13. James, A. M., Hoogewijs, K., Logan, A., Hall, A. R., Ding, S., Fearnley, I. M., and Murphy, M. P. (2017) Non-enzymatic N-acetylation of Lysine Residues by AcetylCoA Often Occurs via a Proximal S-acetylated Thiol Intermediate Sensitive to Glyoxalase II. *Cell Rep.* 18, 2105–2112

14. Baeza, J., Smallegan, M. J., and Denu, J. M. (2015) Site-specific reactivity of nonenzymatic lysine acetylation. *ACS Chem. Biol.* 10, 122–128

15. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E. (2015)

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* 43, D512–20

16. Latorre-Muro, P., Baeza, J., Armstrong, E. A., Hurtado-Guerrero, R., Corzana, F., Wu, L. E., Sinclair, D. A., López-Buesa, P., Carrodeguas, J. A., and Denu, J. M. (2018) Dynamic Acetylation of Phosphoenolpyruvate Carboxykinase Toggles Enzyme Activity between Gluconeogenic and Anaplerotic Reactions. *Mol. Cell* 71, 718–732.e9

17. Xu, Y. S., Liang, J. J., Wang, Y., Zhao, X.-Z. J., Xu, L., Xu, Y.-Y., Zou, Q. C., Zhang, J. M., Tu, C.-E., Cui, Y.-G., Sun, W.-H., Huang, C., Yang, J.-H., and Chin, Y. E. (2016) STAT3 Undergoes Acetylation-dependent Mitochondrial Translocation to Regulate Pyruvate Metabolism. *Sci. Rep.* 6, 39517

18. Baeza, J., Dowell, J. A., Smallegan, M. J., Fan, J., Amador-Noguez, D., Khan, Z., and Denu, J. M. (2014) Stoichiometry of site-specific lysine acetylation in an entire proteome. *J. Biol. Chem.* 289, 21326–21338

19. Nakayasu, E. S., Wu, S., Sydor, M. A., Shukla, A. K., Weitz, K. K., Moore, R. J., Hixson, K. K., Kim, J.-S., Petyuk, V. A., Monroe, M. E., Pasa-Tolic, L., Qian, W.-J., Smith, R. D., Adkins, J. N., and Ansong, C. (2014) A method to determine lysine acetylation stoichiometries. *Int. J. Proteomics* 2014, 730725

20. Zhou, T., Chung, Y.-H., Chen, J., and Chen, Y. (2016) Site-Specific Identification of Lysine Acetylation Stoichiometries in Mammalian Cells. *J. Proteome Res.* 15, 1103–1113

21. Meyer, J. G., D'Souza, A. K., Sorensen, D. J., Rardin, M. J., Wolfe, A. J., Gibson, B. W., and Schilling, B. (2016) Quantification of Lysine Acetylation and Succinylation Stoichiometry in Proteins Using Mass Spectrometric Data-Independent Acquisitions (SWATH). *J. Am. Soc. Mass Spectrom.* 27, 1758–1771

22. Bruderer, R., Bernhardt, O. M., Gandhi, T., Miladinović, S. M., Cheng, L.-Y., Messner, S., Ehrenberger, T., Zanotelli, V., Butscheid, Y., Escher, C., Vitek, O., Rinner, O., and Reiter, L. (2015) Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell. Proteomics* 14, 1400–1410

23. Escher, C., Reiter, L., MacLean, B., Ossola, R., Herzog, F., Chilton, J., MacCoss, M. J., and Rinner, O. (2012) Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* 12, 1111–1121

24. Bruderer, R., Bernhardt, O., Gandhi, T., and Reiter, L. (2016) High precision iRT retention time prediction in the targeted analysis of data-independent acquisition and its impact on identification and quantitation. *Proteomics*,

25. Bruderer, R., Bernhardt, O. M., Gandhi, T., Xuan, Y., Sondermann, J., Schmidt, M., Gomez-Varela, D., and Reiter, L. (2017) Optimization of Experimental Parameters in Data-Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results. *Mol. Cell. Proteomics* 16, 2296–2309

26. Greer, T., Lietz, C. B., Xiang, F., and Li, L. (2015) Novel isotopic N,N-dimethyl leucine (iDiLeu)

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

reagents enable absolute quantification of peptides and proteins using a standard curve approach. *J. Am. Soc. Mass Spectrom.* 26, 107–119

27. Weinert, B. T., Moustafa, T., Iesmantavicius, V., Zechner, R., and Choudhary, C. (2015) Analysis of acetylation stoichiometry suggests that SIRT3 repairs nonenzymatic acetylation lesions. *EMBO J.* 34, 2620–2632

28. Dittenhafer-Reed, K. E., Richards, A. L., Fan, J., Smallegan, M. J., Fotuhi Siahpirani, A., Kemmerer, Z. A., Prolla, T. A., Roy, S., Coon, J. J., and Denu, J. M. (2015) SIRT3 mediates multi-tissue coupling for metabolic fuel switching. *Cell Metab.* 21, 637–646

29. Choudhary, C., Weinert, B. T., Nishida, Y., Verdin, E., and Mann, M. (2014) The growing landscape of lysine acetylation links metabolism and cell signalling. *Nat. Rev. Mol. Cell Biol.* 15, 536–550

30. Cobbold, S. A., Santos, J. M., Ochoa, A., Perlman, D. H., and Llinás, M. (2016) Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. *Sci. Rep.* 6, 19722

31. Carabetta, V. J., Greco, T. M., Tanner, A. W., Cristea, I. M., and Dubnau, D. (2016) Temporal Regulation of the Bacillus subtilis Acetylome and Evidence for a Role of MreB Acetylation in Cell Wall Growth. *mSystems* 1,

32. Henriksen, P., Wagner, S. A., Weinert, B. T., Sharma, S., Bacinskaja, G., Rehman, M., Juffer, A. H., Walther, T. C., Lisby, M., and Choudhary, C. (2012) Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in Saccharomyces cerevisiae. *Mol. Cell. Proteomics* 11, 1510–1522

33. Fan, J., Baeza, J., and Denu, J. M. (2016) Investigating Histone Acetylation Stoichiometry and Turnover Rate. *Methods Enzymol.* 574, 125–148

34. Pagliarini, D. J., Calvo, S. E., Chang, B., Sheth, S. A., Vafai, S. B., Ong, S.-E., Walford, G. A., Sugiana, C., Boneh, A., Chen, W. K., Hill, D. E., Vidal, M., Evans, J. G., Thorburn, D. R., Carr, S. A., and Mootha, V. K. (2008) A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134, 112–123

35. Calvo, S. E., Clauser, K. R., and Mootha, V. K. (2016) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res.* 44, D1251–7

36. Shi, L., and Tu, B. P. (2015) Acetyl-CoA and the regulation of metabolism: mechanisms and consequences. *Curr. Opin. Cell Biol.* 33, 125–131

37. Kim, S. C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L., Grishin, N. V., White, M., Yang, X.-J., and Zhao, Y. (2006) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell* 23, 607–618

38. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834–840

bioRxiv preprint doi: https://doi.org/10.1101/472530; this version posted November 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

39. Magnuson, B., Ekim, B., and Fingar, D. C. (2012) Regulation and function of ribosomal protein S6 kinase (S6K) within mTOR signalling networks. *Biochem. J* 441, 1–21

40. Rosner, M., Schipany, K., and Hengstschläger, M. (2013) Merging high-quality biochemical fractionation with a refined flow cytometry approach to monitor nucleocytoplasmic protein expression throughout the unperturbed mammalian cell cycle. *Nat. Protoc.* 8, 602–626

41. Schwämmle, V., and Jensen, O. N. (2010) A simple and fast method to determine the parameters for fuzzy c-means cluster analysis. *Bioinformatics* 26, 2841–2848

42. Choi, M., Chang, C.-Y., Clough, T., Broudy, D., Killeen, T., MacLean, B., and Vitek, O. (2014) MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* 30, 2524–2526

43. Peng, Y., and Puglielli, L. (2016) Nϵ-lysine acetylation in the lumen of the endoplasmic reticulum: A way to regulate autophagy and maintain protein homeostasis in the secretory pathway. *Autophagy* 12, 1051–1052

44. Dowell, J. A., Frost, D. C., Zhang, J., and Li, L. (2008) Comparison of two-dimensional fractionation techniques for shotgun proteomics. *Anal. Chem.* 80, 6715–6723

45. Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oleś, A. K., Pagès, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L., and Morgan, M. (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* 12, 115–121

46. Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13

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

48. Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G. D. (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5, e13984