Mechanism of Protein Acetylation in Mitochondria

By

Josue Baeza

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN – MADISON

2017

Date of final oral examination: 1/24/2017

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

John M. Denu, Professor, Department of Biomolecular Chemistry

Ronald Raines, Professor, Department of Biochemistry

Richard S. Eisenstein, Professor, Department of Nutritional Sciences

David J. Pagliarini, Associate Professor, Department of Biochemistry

Melissa Harrison, Assistant Professor, Department of Biomolecular Chemistry

Mechanism of Protein Acetylation in Mitochondria

Josue Baeza

Under the supervision of Professor John M. Denu

At the University of Wisconsin-Madison

Reversible N-ε-acetylation has emerged as a widespread and regulatory modification throughout the cell. First identified on N-terminal histone tails over 50 years ago, acetylation has now been identified on over 36,000 lysine sites on organisms ranging from bacteria to mammals, highlighting the strong, conserved role for this modification. Acetylation is controlled by the opposing action of lysine acetyltransferases and deacetylases and recent evidence supports a role for nonenzymatic acetylation. Lysine acetylation influences many biochemical processes including: protein-protein and protein-DNA interactions, stability, cellular localization and enzymatic activity, which is thought to arise from neutralization of the positive charge on lysine residues at neutral pH.

In mitochondria, acetylation generally acts as an inhibitory modification, altering protein interactions, lowered enzymatic activity and mislocalization of mitochondrial proteins. In this case, acetylation acts as a rheostat, modulating the degree of a biochemical process. Therefore, knowing the fraction of the protein that is modified, or stoichiometry, would provide critical information for understanding the regulatory effects of lysine acetylation.

The mechanism of mitochondrial protein acetylation remains unknown. While there is a wealth of data supporting a dynamic and regulatory role for acetylation within mitochondria, there is very little evidence for an enzyme-catalyzed reaction. However, the conditions of the

mitochondrial matrix: elevated acetyl-CoA levels and alkaline pH, raise the possibility that lysine acetylation can occur nonenzymatically.

Throughout my thesis research, I have utilized chemical, biochemical, proteomic, and cellular studies to characterize protein acetylation. Specifically, I developed a mass spectrometry based approach to directly quantify lysine acetylation stoichiometry at the proteome-wide scale. Using this newly developed method, I measured the second order rate constants of nonenzymatic lysine acetylation using mitochondrial and non-mitochondrial proteins, highlighting the wide range of reactivities. The characterization of nonenzymatic rates of native protein demonstrates that the uncatalyzed reaction can account for the levels of acetylation *in vivo*. I also provide evidence that the majority of mitochondrial acetylation occurs within the mitochondria.

Chapter 1 introduces lysine acetylation, the various processes controlling cellular levels including enzymatic and nonenzymatic, and why quantifying acetylation stoichiometry is needed. Chapter 2 discusses the mass spectrometry based method used for quantifying stoichiometry and its application in the model system, *Escherichia coli*. In chapter 3, I quantify the rates of nonenzymatic acetylation using a panel of mitochondrial and nonmitochondrial proteins, which highlights, for the first time, the wide range of lysine reactivities on native proteins. I then utilize an optimized version of the stoichiometry method to query mammalian lysine acetylation stoichiometry, revealing the wide distribution of acetylation across the cell. Conclusions, ongoing work, and future directions are presented in Chapter 5.

Acknowledgements

I want to thank many people who have supported and encouraged me throughout my PhD. I would like to first thank my advisor, Dr. John Denu, for giving me the opportunity to join his lab, and for his support and mentoring throughout my Ph.D. training. I am grateful for the hard work and commitment he gives to his students, which is for our success. I will always remember his dedication as we worked together during those late nights to finish our reactome project. John truly embodies the essence of "leadership by example".

I want to thank my committee members: Dr. Ron Raines, Dr. Rick Eisenstein, Dr. Dave Pagliarini, and Dr. Melissa Harrison for their support and guidance throughout my Ph.D. training. I want to thank Dr. Melissa Harrison for the mentorship and feedback she provided as I prepared for multiple applications including the NSF-GRFP, Postdoctoral application cover letter and Postdoctoral interview preparation. I was blessed to be supported by numerous funding sources including: the UW Science and Medicine Graduate Research Scholars fellowship (SciMed GRS), the NIH Molecular Biosciences Training Grant (MBTG) and the National Science Foundation Graduate Research Scholars Fellowship (NSF-GRFP). I want to thank Dr. Sara Patterson and Abbey Thompson from SciMed GRS, who supported me financially and as I transitioned to graduate school. Thank you to Dr. Jim Keck and Dr. Christina Hull for their support through the Molecular Biosciences Training Grant (MBTG).

I want to thank the Denu lab, past and present, for all their support, friendship, insight, and wonderful times throughout the years. Former Denu lab members include: Sam Oliver, Brittany Albaugh, Elise Wagner, Wei Yu, Melissa Boersma, Kristin Dittenhafer-Reed, Jess Feldman, Zhangli Su, Jing Fan, and Ben Baeten. Thank you to the current members including: James Dowell, Jin-Hee Lee, Rashpal (Rush) Dhillon, Vyacheslav (Slava) Kuznietzov, Wallace Liu, Eric

Armstrong, Kimberly Krautkramer, Anastasia Lindahl, Mark Klein, Spencer Haws, Beatriz Camacho, Alexis Lawton, Yiming (Amy) Qin, and Julie Thelen. Thank you, Michael, Smallegan for your friendship during these years and for teaching me R. We had great times throughout the years and I will never forget the late-night trips to QDoba. I want to especially thank Ann Denu for the work and dedication she provides so the lab runs efficiently. She goes above and beyond to make the lab feel like home and I am grateful for her effort. Finally, I want to thank my two undergraduate students that I had the privilege to work with, Nicole Rademacher and Keighley Reisenauer. You two were a pleasure to work with and I am very proud of the both of you. I wish you all the best in your future goals and look forward to seeing where your lives take you.

My family is the reason why I am here, for their never-ending love and support. For always encouraging me to reach higher, but to never forget who I am or where I came from. Thank you to my parents, Amando and Martha, for teaching me the principle of hard work and patience and for supporting me in my decisions. To my sisters, Tabitha and Jaseileen, for always being there for me as well as their families: Robert Martinez, Aaron Contreras, my beautiful niece, Adriana, and nephews: Cristian, Joaquin, as well as my soon-to-arrive nephew. To my brother-in-law, Misael, thanks for joining us and being a part of our lives here in Wisconsin. Lastly, I want to thank my beautiful wife, Sara Montañez-Sauri, for all she does to make our home feel a little like heaven on earth. You truly exemplify the noble character found in Proverbs 31. “Muchas mujeres han hecho el bien, pero tú las sobrepasas a todas”. We are on this journey together and I am excited for the plans, which the Lord has for us. Te amo, mi Amorcita.

Finally, I want to thank God for directing my path and the opportunities He has given me. “Trust in the Lord with all thine heart; and lean not unto thine own understanding. In all thy ways acknowledge him, and he shall direct thy paths” (Proverbs 3:5-6). “For I know the plans I

have for you," declares the Lord, "plans to prosper you and not to harm you, plans to give you hope and a future" (Jeremiah 29:11). Thank you, God, for the hope and the promise and I look forward to the future you have in store for me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: An Introduction to Reversible Protein Acetylation, Mitochondrial Regulation, and Acetylation Stoichiometry

This chapter is broken up into three sections. The first section will introduce reversible protein acetylation including enzymatic and nonenzymatic mechanisms. The second section will discuss mitochondrial regulation by reversible protein acetylation, specifically by the NAD^+ dependent deacetylase, SIRT3. Finally, the third will introduce the concept of stoichiometry and the importance of quantifying lysine acetylation stoichiometry when trying to understand functional impacts.

Portions of this chapter were published as a co-first author publication in Trends in Biochemical Sciences with the following citation:

Baeza, J., Smallegan, M. J., & Denu, J. M. (2016). Mechanisms and Dynamics of Protein Acetylation in Mitochondria. Trends in Biochemical Sciences, 41(3), 231–244.

J.B. and J.M.S. performed data analysis and generated figures. All authors contributed in drafting and editing the manuscript. We would like to thank members of the Denu laboratory for helpful discussions and thank Q.D. for support throughout the drafting of the manuscript. 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 the National Science Foundation Graduate Research Fellowship Program (NSF-GRFP) DGE-1256259 (J.B.).

1.1 Reversible Lysine Acetylation

1.1.1 Complexity of the proteome

The proteome is the collection of all protein isoforms, or proteoforms, within a cell or organism and is over two orders of magnitude more complex than the genome's predicted protein-coding-genes [1]. Analysis of the human genome estimates 19,000 protein-coding-genes compared to the greater than one million proteoform variants [2]. Each proteoform is a distinct chemical form of a protein that is expressed from a single gene and the chemical diversity can arise through allelic variation, alternative splicing of RNA transcripts, or post-translational modifications (PTM) [3]. Post-translational modification is the modification of a protein following ribosomal translation, which includes proteolytic cleavage of the polypeptide backbone, racemization of amino acid residues, or chemical modification to amino acid functional groups [1]. My doctoral dissertation is specifically focused on chemical modifications which occur on protein lysine residues. Therefore, subsequent mention of post-translational modification refers to the covalent chemical modification of protein residues.

1.1.2 Post-translational modification

Proteins are composed of twenty naturally occurring amino acids. Through post-translational modification of amino acid side chains, nature has expanded the repertoire of functional groups present on proteins. Modifications such as phosphorylation, acetylation, methylation, hydroxylation, deamidation, etc., change the chemical properties of the amino acid side chain, thus altering the protein's structure and function [1]. Even larger protein functional groups such as ubiquitin, Sumo1, and Nedd8 can modify target proteins to form ubiquitination, sumoylation, and neddylation, respectively [4]. Site-specific protein modification regulates diverse biochemical functions, including protein-protein and protein-DNA interactions, cellular

localization, stability, enzyme activity, allostery, etc. Consequently, dysregulation of processes controlling protein post-translational modifications cause deleterious effects to the cell and organism, therefore, understanding the regulatory mechanisms, localizing the site of modification, and quantifying the stoichiometry of modification is critical to elucidate how PTMs function in normal physiology and disease.

1.1.3 Reversible lysine acetylation

Reversible acetylation of the N-ε-amino group of lysine residues is controlled by the opposing activities of lysine acetyltransferases (KAT) and deacetylases (KDAC) (**Figure 1-1)** and has been linked to diverse cellular processes including, gene transcription, metabolism and aging. Lysine acetylation was first discovered over 50 years ago [5], however it has been in the last ten years that mass spectrometry has identified the majority of the acetyl-lysine sites known to date: ~20,000 sites in human cell and tissue samples (**Figure 1-2**), which is in contrast to the greater than ~200,000 phosphorylation sites [6]. In fact, lysine acetylation was discovered 32 years following the initial identification of phosphoserine [7,8], and the field of acetylation has followed steadily behind the phosphorylation field (**Figure 1-3**). In the following sections, I will discuss the major enzyme families involved in reversible lysine acetylation: acetyltransferases and deacetylases as well as the growing evidence for the uncatalyzed, nonenzymatic, acetylation.

1.1.3.1 Lysine acetyltransferases

Lysine acetylation is catalyzed by the distinct action of acetyltransferases utilizing the central carbon metabolite, acetyl Coenzyme A, as a cosubstrate [9]. A total of 37 encoded gene products have been identified to possess lysine acetyltransferase activity with a wide range of activity, function, subcellular localization, and sequence conservation (**Figure 1-4**). Of these, 18 acetyltransferases display robust protein lysine acetylation, representing four major families [10–

12]. Many KAT enzymes are localized to the nucleus and are part of multi-protein complexes involved in chromatin modification of N-terminal histone tails, while others shuttle between the nucleus and cytoplasm. A small number are uniquely located in the cytoplasm, golgi apparatus, endoplasmic reticulum, and mitochondria (**Figure 1-4**).

More than half of the KATs annotated in the literature (19 out of 37) are classified as orphan KATs [10] because little is known of their cellular function. These enzymes are often identified using computational, biochemical, chemical, and cellular methods displaying varying degrees of sequence similarity and enzymatic activity to canonical KATs. However, rigorous biochemical validation and identification of cellular targets are needed to determine if these noncanonical KATs are indeed bona fide protein lysine acetyltransferases. Such is the case for proteins recently reported as mitochondrial localized acetyltransferases, including GCN5L1, ACAT1, and MOF (MYST1 and KAT8) [13–15].

While MOF is classified as a canonical acetyltransferase [16], responsible for deposition of acetyl marks on Histone H4 lysine 16 (H4K16) [17], its mitochondrial subcellular localization has only recently been described [15]. In fact, using genetic, transcriptomic, and biochemical methods, the authors demonstrated regulation of mitochondrial oxidative phosphorylation by controlling gene expression of both nuclear and mitochondrial encoded genes of the OXPHOS system and this regulation was dependent on the catalytic activity of MOF. Despite this evidence, the authors did not identify acetylation sites controlled by MOF. Future studies will be needed to determine which sites are controlled by MOF within mitochondria.

1.1.3.2 Lysine deacetylases (deacylases)

Two different mechanisms have been described for removing acetyl groups from acetyl-lysine residues: metal-dependent acetyl hydrolysis, and acetyl transfer to nicotinamide adenine

dinucleotide (NAD^+). The lysine deacetylase family is divided into four classes: class I, II, and IV are the metal-dependent acetyl hydrolases, and class III are the NAD^+-dependent deacetylases, also called sirtuins [18,19]. There are 18 genes encoded in our genomes which are classified as lysine deacetylases however, the NAD^+-dependent sirtuins have been described to have broader biochemical activity. SIRT1-6 were shown to have general deacylation activity towards varying acyl chain lengths [20]. SIRT4 has also been described to remove lipoyl and biotinyl modifications [21], and SIRT5 displays robust deacetylase activity towards carboxylate containing acyl modifications such as malonylation, succinylation, and glutarylation [22–24]. The presence of these acyl marks as well as the enzymes that catalyze their removal highlight the broad scope of not just lysine acetylation, but general acylation.

1.1.3.3 Nonenzymatic lysine acetylation

Growing lines of evidence suggest that a significant portion of protein acetylation is the result of enzyme-independent or nonenzymatic acetylation. This evidence is exemplified in the mitochondria where widespread acetylation is observed [25–28] with no KAT culpable. Observations made in 1970, prior to the discovery of the first histone acetyltransferase – GCN5, described nonenzymatic acetylation of histone proteins with acetyl CoA [29]. The rate of nonenzymatic acetylation, the authors concluded, was directly proportional to pH and acetyl CoA levels. Additionally, Tanner et al., while performing pH-dependent studies to investigate the function of GCN5 E173 as the general base, quantified the rate of nonenzymatic acetylation on histone proteins. The nonenzymatic rate displayed similar kinetics to the GCN5 E173Q mutant, suggesting that the protonation state of lysine residues determines the propensity for acetylation in the absence of a general base for deprotonation [30].

Interestingly, the factors influencing nonenzymatic acetylation bear resemblance with intramitochondrial conditions as well: alkaline conditions of the mitochondrial matrix due to the OXPHOS proton gradient and acetyl CoA generation through various metabolic pathways (**Figure 1-5**). In 2013, Wagner and Payne demonstrated that alkaline pH and increasing acetyl CoA concentrations was sufficient for increasing acetylation of a purified mitochondrial extract [31]. To rule out the possibility of an unknown acetyltransferase, the authors heat-treated the sample to denature any enzymatic activity and yet, observed a linear increase in acetylation with pH and acetyl CoA levels. Additionally, Weinert et al. identified that acetyl phosphate, another highly reactive acetyl group, could modify proteins in a time and concentration-dependent manner [32]. Acetyl phosphate is only found in prokaryotic organisms and there is no known enzyme to utilize acetyl phosphate as a substrate for lysine acetylation. There is either an as-of-yet unidentified KAT, which utilizes acetyl phosphate as a cosubstrate or, as a general mechanism, highly reactive acetyl groups can modify lysine residues in a nonenzymatic manner.

Figure 1-1: Reversible lysine acetylation

Reversible acetylation of the N-ε-amino group occurs through the opposing actions of protein lysine acetyltransferases and deacetylases. Shown is a diagram representing the acetylation of bovine glutamate dehydrogenase K503 (PDB 3MW9).

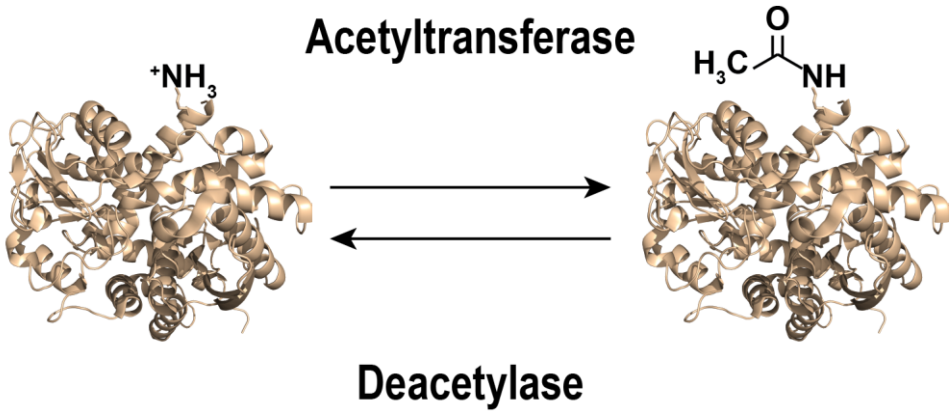


Figure 1-2: Discovery of unique acetylation sites

Bar graph representing the number of unique acetylation sites discovered over time. Notice how the majority of acetylation sites were identified beginning in 2006, corresponding to the first acetylomic MS study [25]. Data was downloaded from the Compendium of Protein Lysine Modifications (http://cplm.biocuckoo.org/) [33,34] Figure design is a courtesy of Michael Smallegan

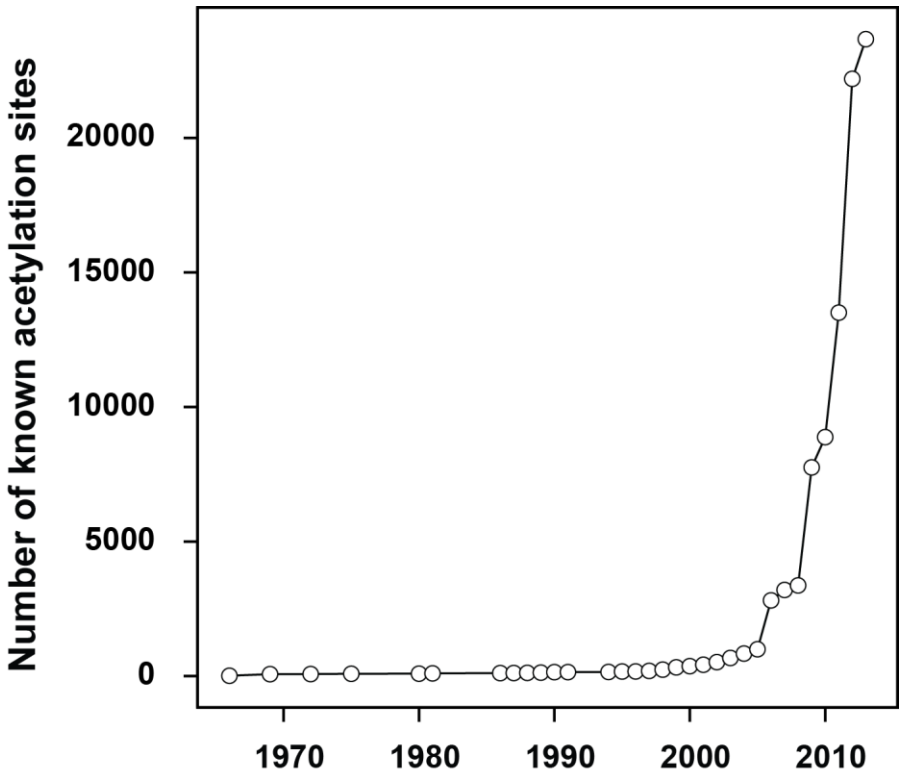


Figure 1-3: Acetylation and phosphorylation literature comparison

Comparison between the number of citations with the keywords: acetylation (red) or phosphorylation (blue) found in Pubmed.

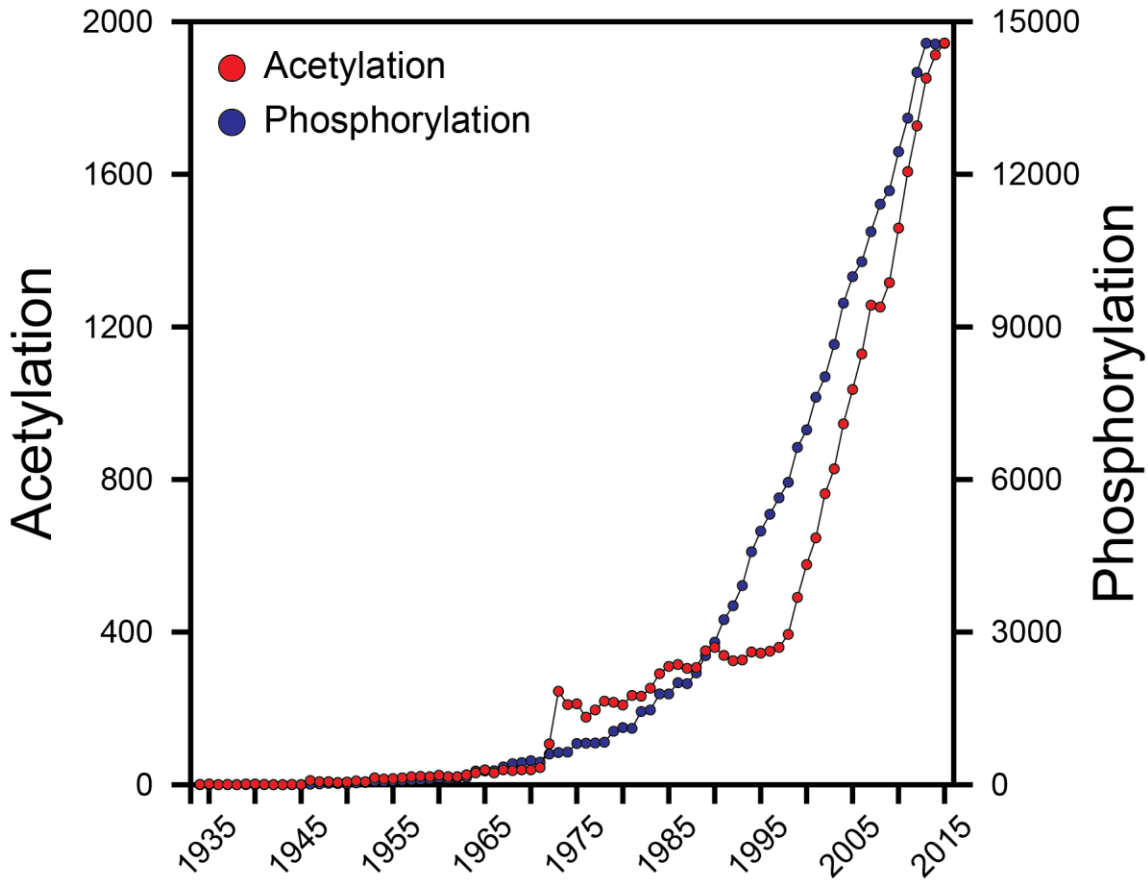


Figure 1-4: Lysine acetyltransferase sequence conservation and subcellular distribution

While the literature commonly reports of four major classes of KATs (GCN5/PCAF, CBP/EP300, MYST, NCOA), phylogenetic analysis reveals broad familial diversity. Figure was downloaded from apps.thesgc.org/resources/phylogenetic_trees/ on November 25, 2016 and the subcellular localization extracted from the Swiss-Prot record of each KAT.

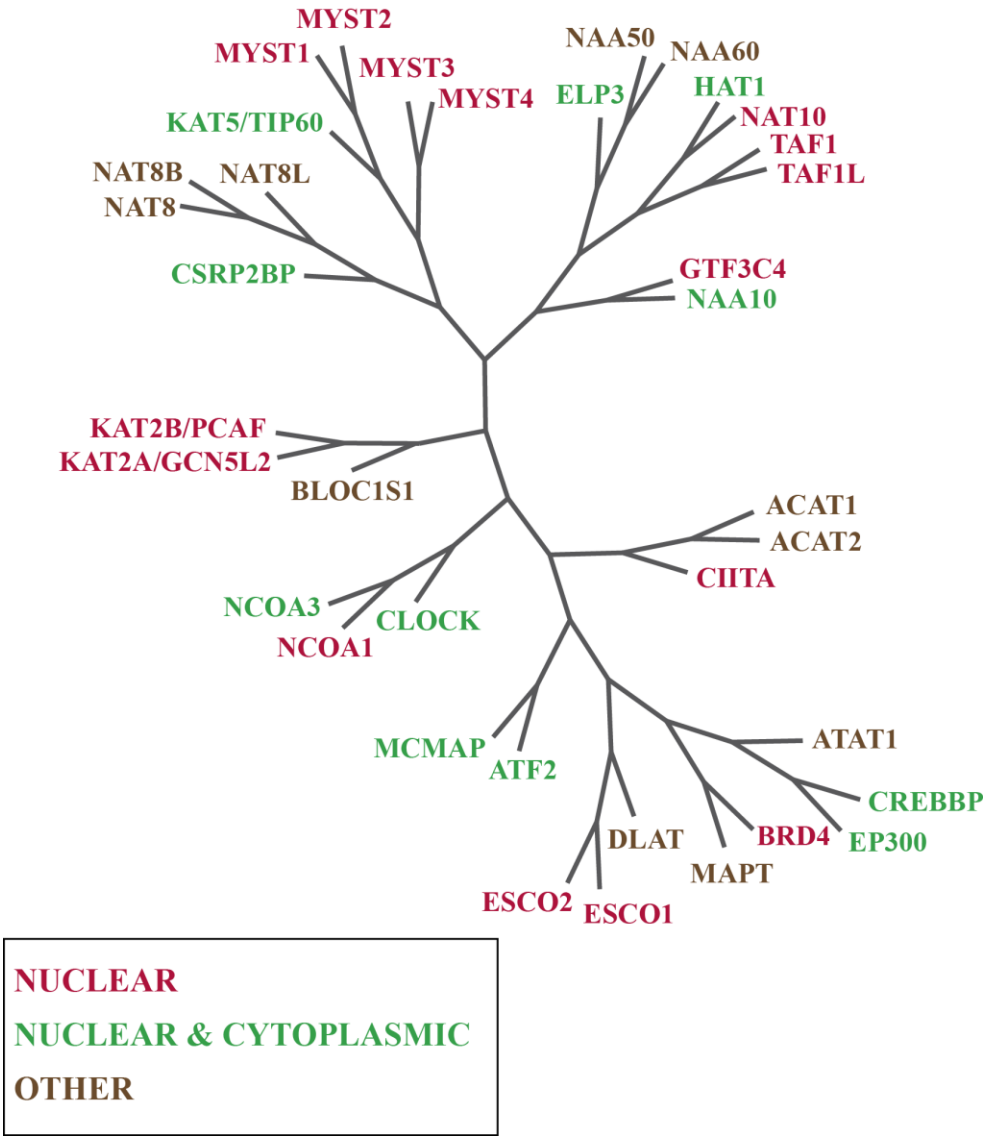
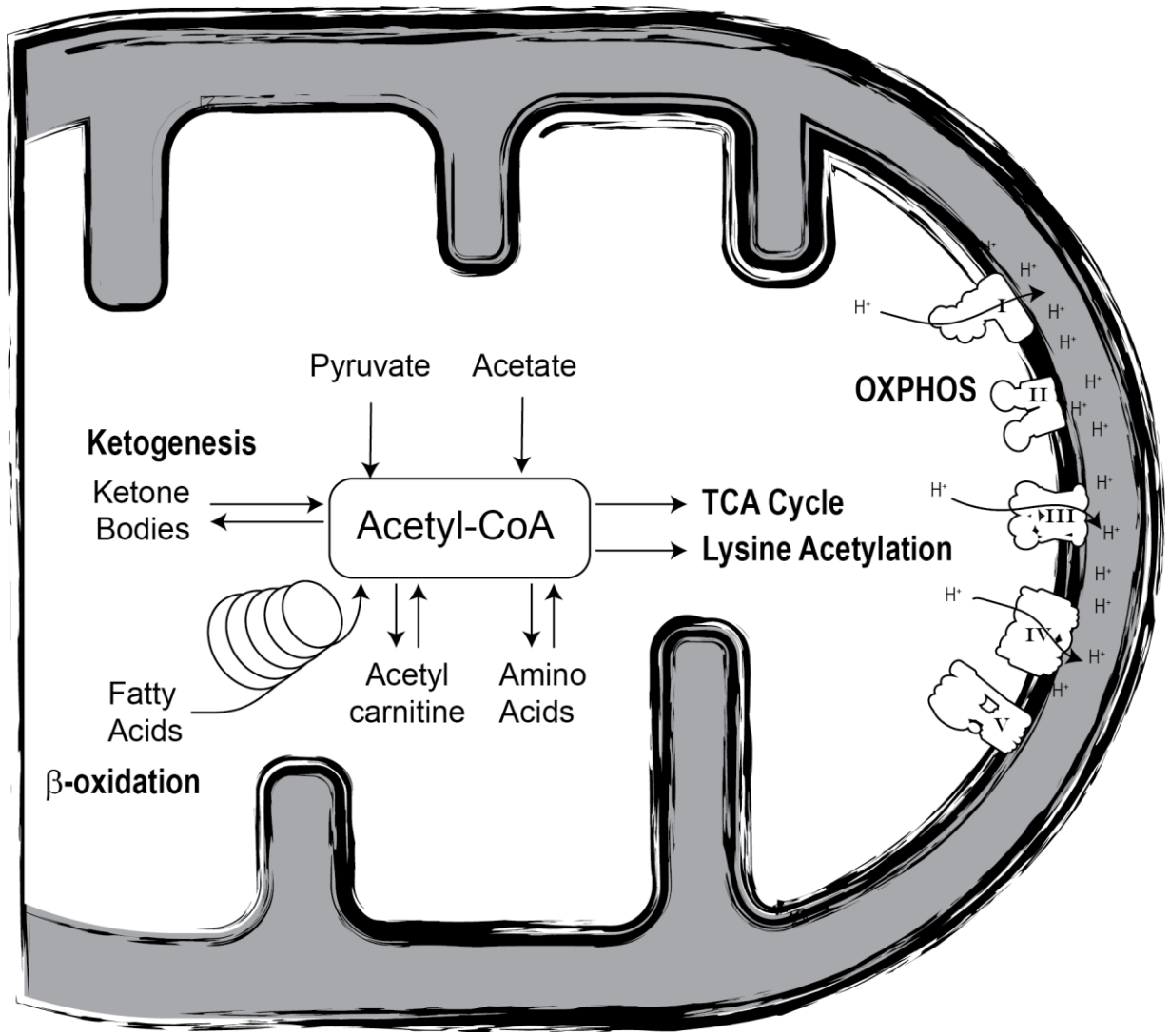


Figure 1-5: Mitochondrial matrix conditions support nonenzymatic acetylation

Through various metabolic pathways, mitochondria produce high levels of acetyl CoA. During oxidative metabolism, the oxidative phosphorylation (OXPHOS) system creates a proton gradient across the inner mitochondrial membrane, elevating the pH within the matrix. Both, high levels of acetyl CoA and alkaline pH, favor nonenzymatic acetylation.



1.2 Lysine Acetylation is a Regulatory Modification

1.2.1 Acetylation regulates gene expression

Acetylation of histone proteins became the first well-established example of biologically functional protein acetylation. Allfrey et al. observed that histones in isolated calf thymus nuclei can be rapidly labeled with radiolabeled acetate, and that these acetylated histones were less inhibitory for RNA polymerase [35]. In the late 1990s the first histone acetyltransferases (HATs) and deacetylases (HDACs) were cloned and linked to the regulation of gene expression on chromatinized templates [36,37]. In this case, acetylation generally correlates with gene expression, acting in part to 'open up' chromatin for appropriate transcriptional machinery to access the DNA template. We now know there are many acetyl-CoA dependent histone acetyltransferases and histone deacetylases that function to regulate all DNA-templated processes, and they are primarily thought to act through the direct reversible acetylation of histone lysine residues [9,38,39]. Current evidence supports the idea that certain site-specific acetylation is sufficient to alter nucleosome dynamics and chromatin folding [40,41]. In addition, acetylated-lysines on histones can function as 'epitopes' for the recruitment of acetyl-lysine binding domains, e.g., bromodomains that are contained within larger protein complexes such as histone acetyltransferases, methyltransferases, transcriptional coactivators, and ATP-dependent chromatin-remodelers [42].

1.2.2 Acetylation regulates non-histone proteins

The acetylation of p53 and α-tubulin were early examples that protein acetylation extends beyond histone proteins [43]. The observation that several deacetylases were localized outside of the nucleus spurred further interest in exploring protein acetylation as a broader phenomenon [5]. Shortly thereafter, the metabolic enzyme, acetyl-CoA synthetase, was found to be regulated by

reversible acetylation in both bacterial and mammalian systems, suggesting that non-histone protein acetylation may be a general mechanism of metabolic regulation. In this case, acetyl-CoA synthetase activity is controlled by acetylation of a single conserved lysine residue in the active site. Acetylation renders the enzyme inactive, while deacetylation restores full activity [44–46]. Collectively, these results demonstrated the existence of functionally-relevant, non-histone targets, which inspired the use of unbiased discovery methods to identify and characterize other acetylation events.

1.2.3 Acetylation regulates mitochondrial enzymes

1.2.3.1 Mass Spectromentry reveals widespread acetylation

Immunoenrichment of acetyl peptides using an anti-acetyllysine antibody coupled to liquid chromatography-mass spectrometry (LC-MS) was the method used to identify protein acetylation. Early acetyl-proteomic studies provided lists of acetylated peptides along with their corresponding proteins. These catalogs were often dominated by metabolic enyzmes and particularly enriched with mitochondrial proteins. Such observations suggested that either the method was biased toward highly-abundant proteins, or there was something unique to metabolic proteins, especially those resident in the mitochondria. Subsequent proteomic studies have confirmed that the modification is widespread, particularly among metabolic and mitochondrial proteins (**Figure 1.6**).

1.2.3.2 Mitochondrial function

Mitochondria function as central mediators of metabolism and energy production [47]. Through the ability to oxidize sugars, fatty acids, and amino acids by reducing molecular oxygen and creating a H^+ gradient across the inner mitochondrial membrane, mitochondria couple this chemiosmotic gradient to the production of ATP. While it is general knowledge that the majority of cellular ATP is generated by these organelles, mitochondria play central roles in many

fundamental cellular processes that include providing precursors for anabolic processes, acting as sentinels of cellular health, and coordinating nucleus–mitochondrion communication. Therefore, revealing previously unknown regulatory networks that operate within mitochondria has broad implications for our understanding of cellular homeostasis and pathology. Identification of N-ε-acetylation in mitochondria is a promising prospect for decoding these regulatory networks within mitochondria.

1.2.3.3 Mitochondrial protein acetylation is dynamic

The accumulated evidence suggests that mitochondrial acetylation is widespread and that the acetylation status of many sites is controlled by the enzymatic activity of the NAD^+-dependent deacetylase sirtuin 3 (SIRT3) [28,48–50]. Other sirtuin family members that reside in the mitochondria include SIRT4 and SIRT5; however, SIRT3 is the only mitochondrial member with robust deacetylation activity [20]. SIRT4 has been shown to possess deacetylation as well as delipoylation and debiotinylation activity, while SIRT5 removes malonyl, succinyl, and glutaryl moieties from lysine residues [21,22,24,51,52]. A definitive mitochondrial acetyltransferase has not been identified; however, nonenzymatic acetylation has been discussed as a possible mechanism to explain the extent of protein acylation [29,31,53,54]. In Chapter 3, I will demonstrate that the chemical reactivities of lysine residues towards acetyl-CoA (as a function of second-order rates) might be sufficient to explain the observed acetylation found in tissues [55].

Although some reports that suggest metabolic proteins can be activated upon acetylation, the majority of well-documented cases indicate that acetylation of mitochondrial enzymes is an inhibitory mark (**Table 1-1**). Overall, the trend suggests that oxidative metabolism is inhibited by higher levels of acetylation among these central metabolic enzymes. Such a regulatory mechanism would serve to sense the overproduction of acetyl-CoA and provide negative feedback to

mitochondrial metabolism. Thus, mitochondria utilize targeted acetylation to decrease the flux through metabolic pathways that operate in oxidative mode. Many of these central enzymes are deacetylated by SIRT3, which reverses the inhibitory effect of acetylation, leading to enhanced oxidative metabolism [28,56,57], and the accompanying stimulation of ROS-mitigating systems, such as isocitrate dehydrogenase (IDH2) and superoxide dismutase 2 (SOD2) [56,58–62]. SIRT3 expression is induced by fasting and chronic caloric restriction [56,63], two conditions that necessitate increased oxidative metabolism of fatty acids and amino acids.

Hebert et al. quantified the fold-change acetylation in SIRT3 wild-type and knockout mice fed either a control or calorie restricted diet [28]. Three classes of acetylation sites emerged from clustering analyses: class 1 are acetylation sites controlled by SIRT3 expression; class 2, acetylation sites primarily affected by a calorie-restricted diet; and class 3, acetylation sites that display minimal or no change under the four conditions [28]. Interestingly, class 3 sites were enriched in loops with a general acidic stretch of amino acids, while class 2 sites were enriched in more hydrophobic sequences. Class 1 sites meet the criteria as targets of SIRT3, with basic amino acids as preferred features around the acetylation site [64]. Collectively, these data suggest that class 3 sites might represent spurious acetylation, whereas class 1 and 2 are regulated sites, with class 1 sites modulated by SIRT3 and class 2 sites being partially buried and inaccessible to SIRT3.

Figure 1-6: Acetylated Proteins in the Context of Global Metabolism

A large proportion of metabolic enzymes are acetylated. Mitochondrial proteins (in blue) and cytoplasmic proteins (in green) that have been found acetylated in M. musculus are overlaid on the KEGG Metabolic Pathways reference pathway. The labeled pathways have had between 8% and 30% of their total lysines acetylated [65,66].

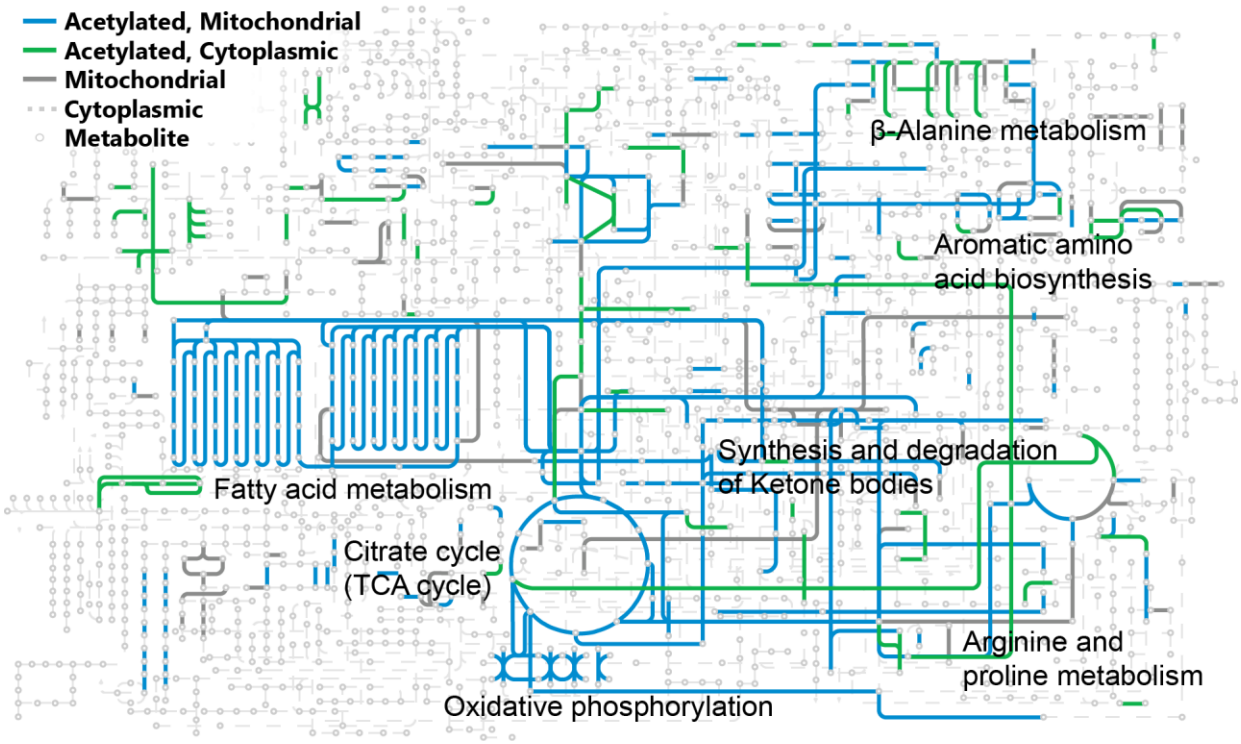


Table 1-1: Mitochondrial Proteins with altered function as a consequence of acetylation.

Gene ID	Gene Name	[1]Organism	Tissue	Site	Effect	Validation	Deacetylase	[2]Deacetylase activity?	Ref.
ACAT1	Acetyl-CoA acetyltransferase 1	Mm	Liver	K260, K265	Inhibitory	AcK.	Sirt3	Yes	[49,67]
AceCS2	Acetyl-CoA synthetase 2	Mm		K635	Inhibitory	AcK (PAT)	Sirt3	Yes	[44]
		Hs		K642	Inhibitory	K→Q	Sirt3	Yes	[45]
ACO2	Aconitase	Mm	Heart	n/a	Stimulatory	Chem AcK	Sirt3	No	[68]
ALDH2	Aldehyde Dehydrogenase 2	Hs		n/a	Stimulatory	n/a	Sirt3	No	[69]
ATP5E	F1F0-ATPase Subunit a	Mm	Liver	n/a	Inhibitory	No	Sirt3	No	[70]
EH/HADH	Enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase	Hs		K165, K171, K346, K584	Stimulatory	4K→Q			[27]
GLUD1	Glutamate dehydrogenase 1	Mm	Liver	n/a	Inhibitory	No	Sirt3	Yes	[50]
		Hs		n/a	Inhibitory	No	Sirt3	Yes	[62]
GOT2	Glutamate oxaloacetate transaminase	Hs		K159, K185, K404	Binding	AcK, K→R, K→Q	Sirt3	No	[71]
HMGCS2	3-Hydroxy-3-Methylglutaryl-CoA Synthase 2, Mitochondrial	Mm	Liver	K310, K447, K473	Inhibitory	K→R	Sirt3	Yes	[72]
IDH2	Isocitrate Dehydrogenase 2	Hs		n/a	Inhibitory	No	Sirt3	Yes	[62]
		Mm	Liver, Inner ear, brain	n/a	Inhibitory	No	Sirt3	Yes	[56]
		Mm	Liver	K413	Inhibitory	AcK	Sirt3	Yes	[61]
LCAD	Long chain acyl-coa dehydrogenase	Hs	Liver	K42	Inhibitory	K→R	Sirt3	Yes	[73]
		Mm	Liver	K318, K322	Inhibitory	Chemical AcK	Sirt3	Yes	[74]
MDH2	Malate Dehydrogenase 2	Hs		K185, K301, K307, K314	Stimulatory	4K→R			[27]
		Mm	Liver	K239	Inhibitory	K→Q	Sirt3	No	[28]
MRPL10	Mitochondrial Ribosomal Protein L10	Mm	Liver	n/a	Inhibitory	No	Sirt3	Yes	[75]
NDUFA9	NADH Dehydrogenase 1 a Subcomplex 9	Mm	Liver	n/a	Inhibitory	No	Sirt3	No	[76]
OGG1	8-Oxoguanine DNA glycosylase 1	Hs		n/a	Instability	No	Sirt3	No	[77]
OPA1	Optic atrophy 1	Mm	Heart	K926, K931	Inhibitory	K→Q, K→R	Sirt3	Yes	[78]
OSCP	Oligomycin sensitivity conferring protein	Hs		n/a	Inhibitory	No	Sirt3	Yes	[79]
OTC	Ornithine Transcarbamoyltransferase	Mm	Liver	K88	Inhibitory	No	Sirt3	Yes	[80]

OXCT	Succinyl CoA:3-ketoacid-CoA transferase	Mm	Brain	K451	Inhibitory	AcK	Sirt3	Yes	[49]
P450scc	P450 Cholesterol side chain cleavage monooxygenase	Hs		K148, K149	Inhibitory	K→A	Sirt3	No	[81]
PDHA	Pyruvate Dehydrogenase E1 alpha subunit	Mm	Muscle	K336	Inhibitory	K→Q, K→R	Sirt3	No	[82]
PPID	Peptidylprolyl cis-trans Isomerase D (Cyclophilin D)	Hs		K145	Binding	K→Q, K→R	Sirt3	No	[83]
		Hs		n/a	Binding	No	Sirt3	No	[84]
SDHA	Succinate Dehydrogenase Complex, Subunit A, Flavoprotein	Mm	Liver	n/a	Inhibitory	No	Sirt3	No	[85]
		Mm	Liver	n/a	Inhibitory	No	Sirt3	Yes	[86]
SKP2	F-box protein S-Phase kinase associated protein 2	Hs		K68, K71	Mislocalization	K→Q, K→R K→L	Sirt3	No	[87]
SOD2	Superoxide Dismutase 2	Mm	Liver	K122	Inhibitory	K→R	Sirt3	Yes	[58]
		Mm	Liver	K53, K89	Inhibitory	K→R	Sirt3	Yes	[60]
		Hs		K68	Inhibitory	K→Q, K→R	Sirt3	No	[59]
VLCAD	Very Long chain acyl-coa dehydrogenase	Hs	Liver	K507	Inhibitory, Mislocalization	Chemical AcK	Sirt3	Yes	[88]

[1]Hs-H. sapiens, Mm-M. musculus

[2]In vitro deacetylase activity with full length protein?

1.3 Quantitative acetyl-proteomics (acetylomics)

1.3.1 Acetylomics by the numbers

Growing interest in the protein acetylation field has fueled concerted efforts to characterize the protein acetylome (or acetyl-proteome). By using various nutritional, genetic, and pharmacological model systems, the roster of acetylated lysine sites has expanded rapidly in the past decade (**Figure 1-7**). In *Mus musculus*, the focus of the majority of acetyl-proteomics studies, 13,186 sites across 4,024 proteins have been identified, and a full 63% of mitochondrially localized proteins contain acetylation sites. A comparable number of sites have been found in human cell lines: ~11,000 sites [33,34,66]. The bulk of these data have been generated in the past five years. In 2011, there were only 552 recorded acetylation sites in *Mus musculus* [34]. It is not solely the raft of new studies that have added to the compendium. Improvements in antibody reagents, chromatography, and MS technology have increased the proportion of the true acetyl-proteome accessible to each study. For acetylome papers published in 2011, the average number of sites reported was 1,357, and in 2015 the average number of sites per paper is 7,784 (**Figure 1-7**). It will be interesting to follow these numbers in the upcoming years and assess whether the total will grow substantially as MS technology makes additional gains in sensitivity.

1.3.2 Identified acetyl sites vs. regulatory acetyl sites

Despite the vast catalog of site-specific mitochondrial acetylation, only a few examples have an established functional consequence. Of the ~700 acetylated mitochondrial proteins, 26 display functional effects when acetylated (**Figure 1-8 and Table 1-1**). This relatively low number is largely due to the technical challenges of acquiring biochemical evidence for altered function. It is even more difficult to determine which biochemical feature is affected: enzyme activity, protein-protein interactions, protein-DNA interactions, stability, localization, allostery, etc. The

most rigorous approach to validate a possible regulatory function is to prepare and characterize a fully unacetylated form and a site-specifically acetylated form that is stoichiometrically modified. Genetic incorporation of an acetyl-lysine onto recombinant proteins using an orthogonal acetyllysl-tRNA synthetase--tRNA pair is one such approach. This system produces homogenous recombinant protein containing an acetyl-lysine at defined sites [41,89]. Another commonly used method to probe acetyl-lysine function is using site-directed mutagenesis. A lysine-to-glutamine (K→Q) substitution is often considered to be an acetyl mimic because of the resemblance of the uncharged functional group. Likewise, a lysine-to-arginine (K→R) substitution preserves the positively charged functional group and is often utilized as an unmodified lysine mimic. However, such mimics are not classical isosteres and accordingly do not always yield the expected results, as illustrated in a recent study on the autoacetylation of a histone acetyltransferase in which both the K→Q and K→R substitutions yielded protein that was ~100-fold less active than the active wild type auto-acetylated species [90]. As an alternative, in vitro acetylation of a target protein can be obtained by using enzymatic or chemical methods [44,74]. The use of a highly specific acetyltransferase can generate a homogenous population of acetylated protein; however, the lack of strong sequence specificity of protein acetyltransferase KAT complexes limits the widespread use of this method. Reactive acetylating agents, such as acetic anhydride and sulfo-N-hydrosuccinimide (NHS)-acetate, are chemical methods to assess the functional impact of protein acetylation. Both strategies suffer from the difficulty of achieving high stoichiometry and from issues of targeting acetylation only to the appropriate lysine residues. These genetic, enzymatic and chemical tools have enabled identification of functionally relevant sites and will continue to be key tools in determining the broader roles of protein acetylation.

1.3.3 Stoichiometry

1.3.3.1 Relative Quantitation vs Stoichiometry

Relative quantitation-based MS has been crucial for understanding acetylation dynamics and the mechanisms that modulate mitochondrial metabolism. However, these methods do not provide direct information on stoichiometry; that is, the fraction of protein that is modified. Stoichiometric information is necessary to provide a more complete picture for the role of acetylation as a regulatory mechanism. Consider the case where two peptides are quantified in two different biological conditions using a relative quantitation-based method, and each peptide displays a five-fold change. Applying a stoichiometry-based method, the same two peptides are measured as 1-5% change and 10-50% change (**Figure 1-9**). In this scenario, the two 'five-fold' changes may have very different effects on the system, given that the level of inhibition (or other regulatory effects) likely scales with the proportion of a protein modified. Whether this scaling is linear or in some way cooperative and nonlinear is an open question. Resolving this question will bring us closer to discerning whether acetyl-peptides in the medium stoichiometry range (5-15%) can have significant biological/regulatory effects.

1.3.3.2 The "Baeza" Method

My graduate thesis work is focused on the development, characterization, and implementation of a method to quantify acetylation stoichiometry at the proteomic scale. This method termed, "The Baeza Method", which was coined and affectionately used by members of the Denu lab, employs a protein chemistry approach to label all unmodified lysine residues with a "heavy" stable isotope acetyl group with the "light" acetyl group originating in vivo through normal cellular processes. Protein digestion generates light and heavy acetyl peptides with matching chromatographic retention times, which can be resolved using high-resolution mass

spectrometry (**Figure 1-10**) [91]. Comparison of the light (endogenous) and isotopic (chemically labeled) acetyl-peptides provides a direct measurement of the stoichiometry. A limitation of this method, however, is the inability to account for other potential lysine modifications, and therefore the ratios that are reported are the ratio of the acetyl-peptide to the acetyl-peptide plus the unmodified corresponding peptide.

1.3.3.3 Mitochondrial acetylation is generally low

Initial stoichiometry measurements suggest that the average level of acetylation across all detected lysine-containing peptides is in the 0-5% range, with most acetylated lysines showing levels below 1% [92]. In *Escherichia coli* and *Saccharomyces cerevisiae*, reported estimated acetylation stoichiometry values range from less than 1% - 98% [32,91,93]. Using an indirect approximation, Weinert et al. estimated that 48 SIRT3-regulated sites would exhibit stoichiometry >1% in SIRT3 knockout mice, with one site predicted to be up to 87% acetylated. Together, these observations suggest that most lysine residues detected in MS acetyl-proteomic analyses harbor relatively low acetylation stoichiometry. This raises many questions about the roles played by site-specific acetylation events. One could argue that such low overall stoichiometry means that most acetylations are spurious and of no biological consequence. This argument has merit and is supported by the existence of class 3 acetylation sites that were not altered by SIRT3 expression or dietary restriction. However, such a blanket argument does not account for the existing biochemical and biological data supporting functionally relevant acetylation. In addition, the current data are limited to a few biological conditions, and the possibility that cellular stresses could greatly induce high stoichiometry remains to be evaluated. Nevertheless, a logical conclusion is that spurious acetylation coexists with targeted (functional) acetylation, and the

overall steady-state levels are a function of the rates of acetylation, deacetylation and protein turnover, which are subject to metabolic flux, enzyme expression, activity and protein stability.

Figure 1-7. Timeline of Detected Acetylated Peptides per Publication, 2006-2015

The number of acetylated peptides reported per acetyl-proteome study, grouped by year. Included experiments used liquid chromatography coupled mass spectrometry (LC-MS) to query the acetylation on the proteome level with the data type indicated by fill color. "Identification" studies report a list of peptides, "quantification" studies report the change in relative abundance in an acetyl-peptide between two experimental conditions, and "stoichiometry" studies report the ratio of acetyl-peptide to the amount of acetyl-peptide plus the unmodified corresponding peptide. The papers include experiments performed in the following organisms: M. musculus, H. sapiens, S. enterica, D. melanogaster, S. cerevisiae, R. norvegicus, and E. Coli. [25–28,32,48,49,57,67,91–109]

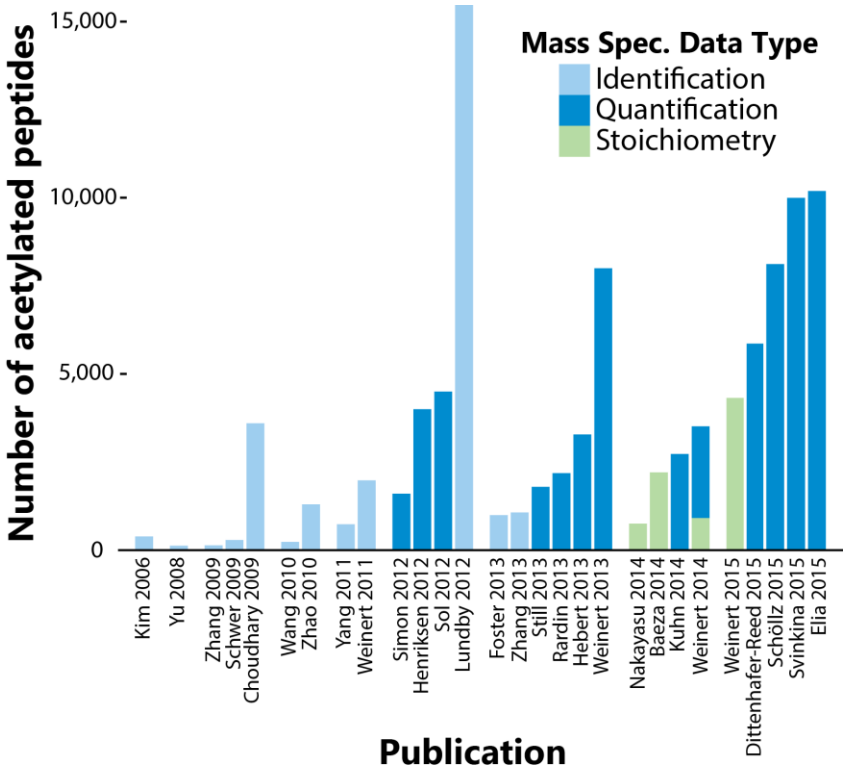


Figure 1-8: Overview of proteins in mitochondrial metabolism with functional consequence caused by reversible acetylation

Diagram visualizing key metabolic pathways in mitochondrial metabolism. Each enzyme or subunit of a multi-enzyme complex contains acetylation site(s) that cause a functional consequence on enzymatic activity.

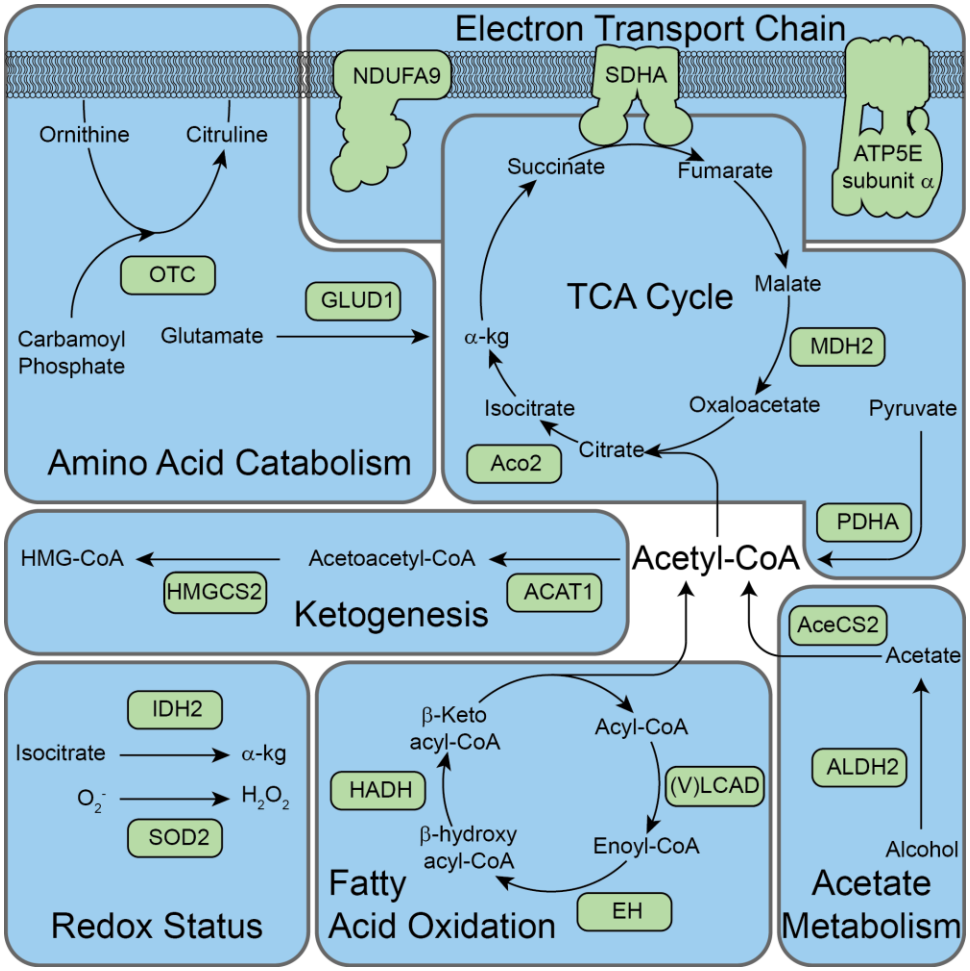


Figure 1-9: Relative quantitation vs. Stoichiometry

Diagram depicting the key difference between quantifying relative changes with stoichiometry. Relative quantitation would consider both scenarios identical, 5-fold increase, while stoichiometry would determine 1-5% and 10-50% change.

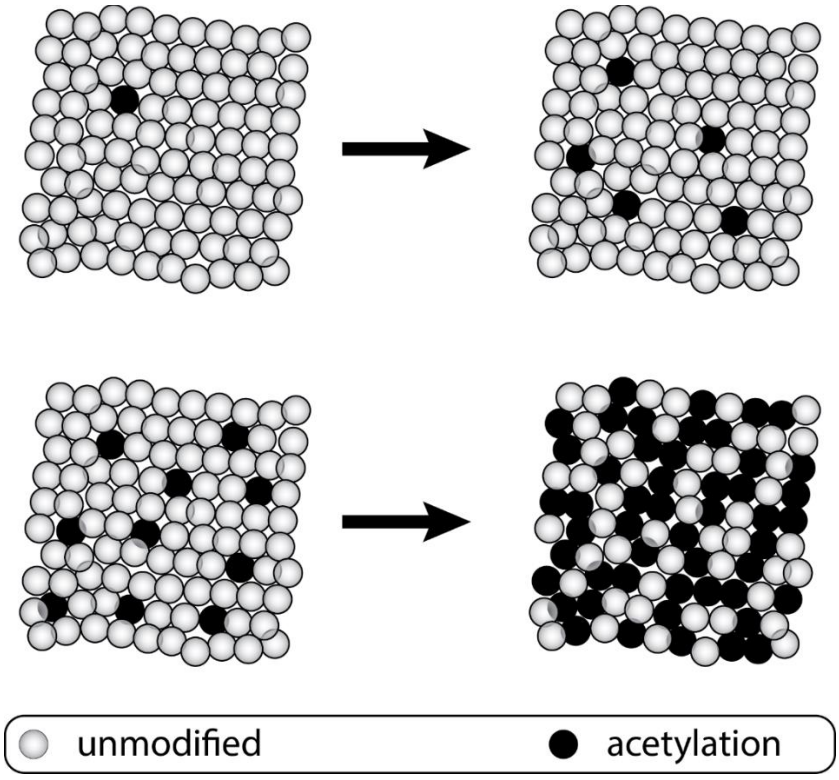


Figure 1-10: General scheme for determining acetylation stoichiometry

A protein sample with varying degree of acetylation is denatured and chemically acetylated using isotopic acetic anhydride followed by trypsin digestion. The heavy and light peptides are resolved using high resolution mass spectrometry and stoichiometry is determined using the peak area of the light and heavy peptides with the following equation: L / (L+H).

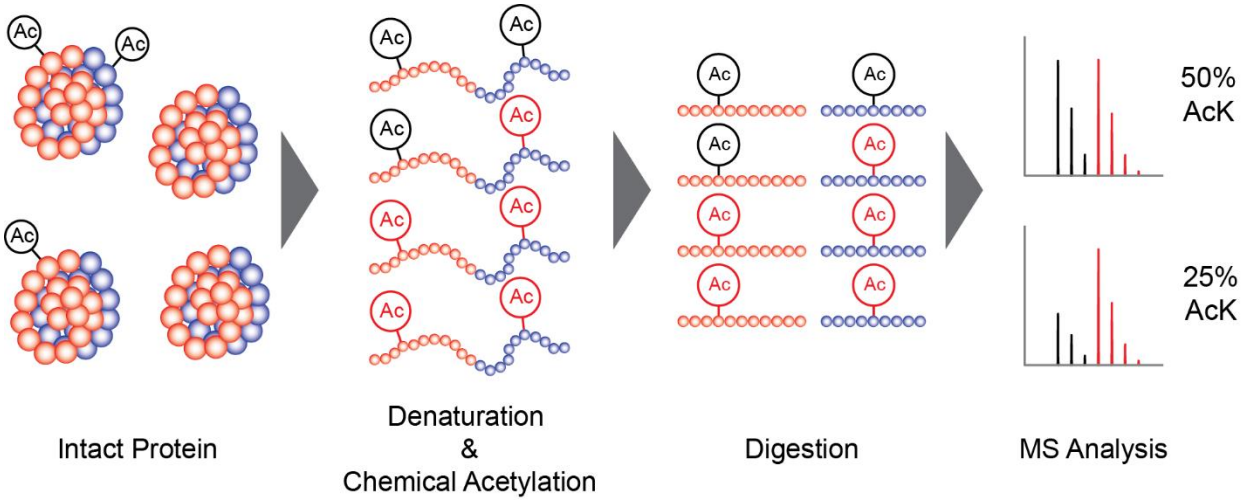
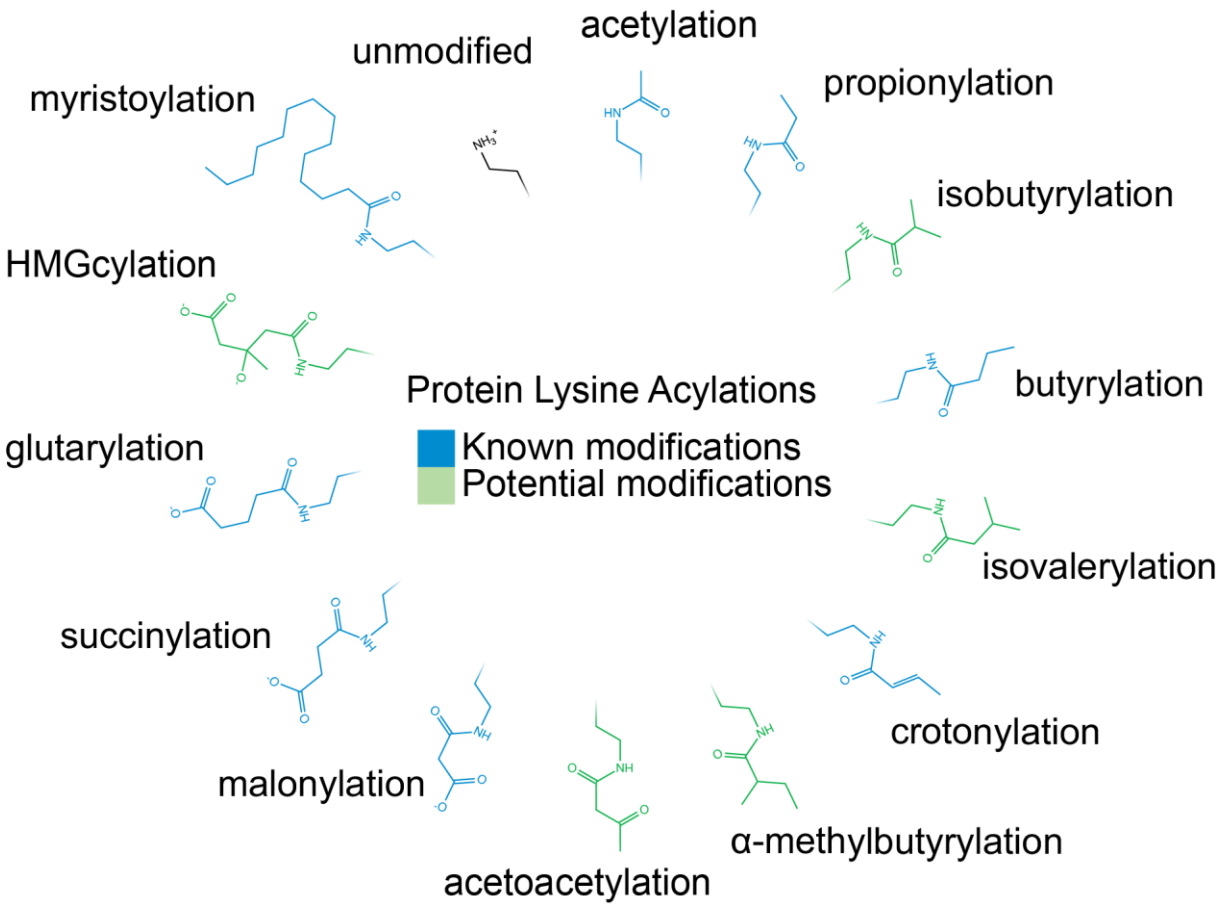


Figure 1-11: Described and Potential Lysine Acylations

Lysine has been shown to be modified by a number of acyl-groups (in blue), which are often donated from the corresponding acyl-CoA. Acetylation, succinylation [110], propionylation [111], butyrylation [112,113], malonylation [22], glutrarylation [24], myristoylation [114], and crotonylation [115] are among lysine modifications previously discovered. Other acyl-CoAs that are present in the mitochondria are likely to also modify lysines (in green), particularly if the acylation process is primarily nonenzymatically driven.



1.3.4 Regulatory mechanisms that feature low stoichiometry

While it might appear reasonable to discount very low stoichiometry as spurious, and to assign functional meaning to only very high stoichiometry, we must explore the possibility that many regulated acetylation sites with significant but low stoichiometry (e.g. 5-15%) are sufficient to induce clear phenotypes. In the following discussion, I will explain potential models where low but significant acetylation can impact function.

1.3.4.1 Stoichiometry of acylation

The large-scale MS methods mapping acetyl, malonyl, succinyl, and glutaryl moieties on lysine residues have treated each modification as a unique entity. Each modification alters the charge state from z = +1 (unmodified) to z = 0 (acetyl) or z = -1 (malonyl, succinyl, glutaryl). In addition, butyrylation and propionylation have also been documented in liver mitochondria [113,116]. All such acylations are most likely derived from their CoA derivatives, possibly through a nonenzymatic mechanism [29,54,55,93,116–118]. Based on these observations, there is potential that other acylations on lysine residues await discovery (**Figure 1-11**). Specific acylations may be enriched on metabolic enzymes belonging to the pathway that generates the respective acyl-CoA. For example, 3-hydroxy-3-methylglutarylation and acetoacetylation may be found on enzymes involved in ketogenesis, such as acetyl-CoA acetyltransferase 1 (ACAT1), 3-Hydroxy-3-Methylglutaryl-CoA Synthase 2 (HMGCS2) and Hydroxymethylglutaryl-CoA lyase (HMGCL), while isobutarylation (a valine derivative), isovalerylation (a leucine derivative), and α-methylbutarylation (an isoleucine derivative) may be found on enzymes involved in branched chain amino acid catabolism. In a biochemical pathway, sets of different lysine modifications are likely to exist among a protein population, and might have similar regulatory outcomes. Therefore, we need to account for the full distribution of a peptide across all possible modification states

(including but not limited to the above acylations). The current stoichiometry methodology only accounts for the ratio of the unmodified peptide to the corresponding acetylated peptide. Tools to query the full range of possible acylations at the proteome scale are needed.

1.3.4.2 Shared Functional Outcome of Multi-Site Acetylation on Individual Proteins

Some documented cases show that a single acetylation site directly controls enzymatic activity, such as for Acetyl-CoA Synthetase 2 (AceCS2) [44,45]. In other cases, multiple sites affect activity, such as with ACAT1 (sites K260, K265), HMGCS2 (sites K310, K447, K473) and SOD2 (sites K122, K53, K89, K68) (**Table 1-1**). On most proteins, the impact of acetylation is observed consistently across many studies, but the functional impact of multi-site acetylation has not been experimentally determined. For example, carbamoyl-phosphate synthase (CPS1) has 59 documented acetylation sites, but no single regulatory acetyl-site in particular is known to affect activity, even though $Sirt3^{-/-}$ mice exhibit CPS1 hyperacetylation and an altered urea cycle, where CPS1 plays a vital role in removing ammonia from the cell [28]. Given that proteins are digested into peptides before MS analysis, one cannot discern whether multi-site modifications exist on the same protein molecule or on different protein molecules. There are two mechanisms by which low stoichiometry across multiple sites can have additive effects on activity: (i) acetyl-sites co-occuring on the same molecule could ratchet activity down (or up) in accordance with how many sites are occupied, or (ii) acetyl-sites on different molecules with similar functional effects could in each contribute to dampening (or enhancing) activity.

1.3.4.3 Cumulative effect of low-stoichiometry acetylation across a pathway

While a low level of acetylation on one particular enzyme may have a limited effect on its catalytic capacity, the low-stoichiometry acetylation of multiple enzymes in a metabolic pathway, each causing a small decrease (or increase) in activity, could cumulatively result in physiologically

meaningful alteration in metabolite flux. Hebert et al. report nine multi-enzyme mitochondrial pathways that have perturbed metabolite concentrations in Sirt3$^{-/-}$ mice. β-Oxidation, which was found to have altered metabolites, contains five enzymes targeted by SIRT3 [28]. Similarly, Dittenhafer-Reed et al. demonstrated dysregulated ketone body utilization in the brain of mice lacking SIRT3, a pathway which in brain tissue was found to have 13 predicted SIRT3 targeted acetyl-sites distributed over five proteins involved in ketone body utilization [49]. In light of this, it is very possible that no single enzyme in particular is culpable, but instead the observed metabolite concentration changes are the additive effects of low-stoichiometry modulation of several enzymes in the same pathway.

1.3.4.4 Acetylation and cooperativity in multimeric enzyme complexes

Metabolic enzymes often function as higher-ordered complexes including hetero- and homo-oligomerization of polypeptide chains. Glutamate dehydrogenase (GDH) is a homohexamer which displays cooperativity effects upon ligand binding [119,120]. GDH contains an ‘antenna’ domain that protrudes from the structure and was proposed to be an intersubunit communication link during cooperativity [121]. Acylation of two lysine sites (K477 and K480 on bovine GDH) at the apex of the antennae has been reported, and, interestingly, these sites displayed the highest reactivity towards acetyl-CoA [28,55]. How would acetylation affect the cooperativity of the enzyme complex? What level of stoichiometry would be needed to disrupt protein-protein interactions between multimeric enzymes? Would the acetylated monomer weaken the interactions among the rest of the complex? It is tempting to suggest that acetylation could disrupt complex formation or the communication between monomers, in which case having a low stoichiometry would have a non-linear effect.

1.3.4.5 Acetylation measurements are an aggregate of many levels of biological hierarchies

When acetylation stoichiometry is discussed in a particular biological condition, it is important to note that MS measurements collapse many levels of biological hierarchies into one number. In a cell culture study, these data produce one value that represents the average stoichiometry over the whole population of cells in the sample, over the whole population of mitochondria within each cell, and over sub-compartmental, local concentrations of proteoforms, all of which could contain significant heterogeneity. This issue is compounded in studies of bulk tissue, which additionally contain a heterogenous population of cell types and functionality. What appears to be very low stoichiometry in the aggregate could be a combination of highly acetylated populations and populations with no significant modifications, or any intermediate between these extremes. For instance, mitochondria within the same cell or between cells in a population might display very diverse protein-acetylation levels but, with the current approaches, these are averaged during analysis. It is intriguing to suggest that mitochondrial dysfunction caused by high levels of protein acetylation might drastically impair functionality organelle-wide and initiate mitophagy. Such a scenario might predict that highly acetylated mitochondria are rapidly turned over which would contribute to low observable stoichiometry under steady-state conditions. Such catastrophic events could be induced by supra-physiological levels of mitochondrial acetyl-CoA, perhaps combined with loss of SIRT3 function. It would not be unprecedented for acetyl-CoA levels to induce the degradation machinery; cytosolic depletion of acetyl-CoA has recently been shown to stimulate autophagy [122].

1.4 Concluding Remarks

A significant part of my Ph.D. research was aimed at quantification of acetylation stoichiometry and elucidation of the mechanism of mitochondrial protein acetylation. In chapter

2, I apply a novel method to quantify acetylation stoichiometry in *Escherichia coli* in order to understand changes in acetylation stoichiometry dependent on the bacterial NAD^+-dependent deacetylase, CobB. In chapter 3, I quantify the second order rate constants of lysine acetylation for a panel of mitochondrial and non-mitochondrial proteins. This study is the first to characterize nonenzymatic acetylation using quantitative techniques and supports the idea of a nonenzymatic mechanism responsible for the widespread acetylation in vivo. Chapter 4 quantifies the cellular distribution of acetylation stoichiometry. This global analysis of stoichiometry highlights the regulatory dichotomy of protein acetylation of nuclear and mitochondria proteomes. For example, the general trends suggest that nuclear protein acetylation is enzyme-catalyzed, usually leading to up-regulated processes such as transcriptional activation. In mitochondria, current evidence supports nonenzymatic mechanisms for acetylation, with most acetylated proteins exhibiting a loss of function. Both organelles utilize deacetylases to remove these modifications. High levels of nuclear acetyl-CoA represent pro-growth signal. In the mitochondria, acetyl-CoA levels in excess of demand constitute a signal to slow flux through oxidative energy production. By inducing acetyl-CoA dependent protein acetylation, a substrate-level braking system is applied. When energy demands require more oxidative metabolism, induced SIRT3 expression relieves the brake, allowing the cell to push the accelerator for more energy production.

1.5 References

1 Walsh, C.T. et al. (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. Angew. Chem. Int. Ed Engl. 44, 7342–7372

2 Ezkurdia, I. et al. (2014) Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. Hum. Mol. Genet. 23, 5866–5878

3 Smith, L.M. et al. (2013) Proteoform: a single term describing protein complexity. Nat. Methods 10, 186–187

4 Swatek, K.N. and Komander, D. (2016) Ubiquitin modifications. Cell Res. 26, 399–422

5 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

6 Hornbeck, P.V. et al. (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. 43, D512–20

7 Phillips, D.M. (1963) The presence of acetyl groups of histones. Biochem. J 87, 258–263

8 Lipmann, F.A. and Levene, P.A. (1932) SERINEPHOSPHORIC ACID OBTAINED ON HYDROLYSIS OF VITELLINIC ACID. J. Biol. Chem. 98, 109–114

9 Roth, S.Y. et al. (2001) Histone acetyltransferases. Annu. Rev. Biochem. 70, 81–120

10 Montgomery, D.C. et al. (2014) Chemoproteomic profiling of lysine acetyltransferases highlights an expanded landscape of catalytic acetylation. J. Am. Chem. Soc. 136, 8669–8676

11 Arrowsmith, C.H. et al. (2012) Epigenetic protein families: a new frontier for drug discovery. Nat. Rev. Drug Discov. 11, 384–400

12 Liu, L. et al. (2012) ChromoHub: a data hub for navigators of chromatin-mediated signalling. Bioinformatics 28, 2205–2206

13 Scott, I. et al. (2012) Identification of a molecular component of the mitochondrial acetyltransferase programme: a novel role for GCN5L1. Biochem. J 443, 655–661

14 Fan, J. et al. (2014) Tyr phosphorylation of PDP1 toggles recruitment between ACAT1 and SIRT3 to regulate the pyruvate dehydrogenase complex. Mol. Cell 53, 534–548

15 Chatterjee, A. et al. (2016) MOF Acetyl Transferase Regulates Transcription and Respiration in Mitochondria. Cell 167, 722–738.e23

16 Allis, C.D. et al. (2007) New nomenclature for chromatin-modifying enzymes. Cell 131, 633–636

17 Taipale, M. et al. (2005) hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. Mol. Cell. Biol. 25, 6798–6810

18 Lombardi, P.M. et al. (2011) Structure, mechanism, and inhibition of histone deacetylases and related metalloenzymes. Curr. Opin. Struct. Biol. 21, 735–743

19 Feldman, J.L. et al. (2012) Sirtuin catalysis and regulation. J. Biol. Chem. 287, 42419–42427

20 Feldman, J.L. et al. (2013) Activation of the protein deacetylase SIRT6 by long-chain fatty acids and widespread deacylation by mammalian sirtuins. J. Biol. Chem. 288, 31350–31356

21 Mathias, R.A. et al. (2014) Sirtuin 4 is a lipoamidase regulating pyruvate dehydrogenase complex activity. Cell 159, 1615–1625

22 Peng, C. et al. (2011) The first identification of lysine malonylation substrates and its regulatory enzyme. Mol. Cell. Proteomics 10, M111.012658

23 Rardin, M.J. et al. (2013) SIRT5 regulates the mitochondrial lysine succinylome and metabolic networks. Cell Metab. 18, 920–933

24 Tan, M. et al. (2014) Lysine glutarylation is a protein posttranslational modification

regulated by SIRT5. Cell Metab. 19, 605–617

25 Kim, S.C. et al. (2006) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Mol. Cell 23, 607–618

26 Choudhary, C. et al. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 325, 834–840

27 Zhao, S. et al. (2010) Regulation of cellular metabolism by protein lysine acetylation. Science 327, 1000–1004

28 Hebert, A.S. et al. (2013) Calorie restriction and SIRT3 trigger global reprogramming of the mitochondrial protein acetylome. Mol. Cell 49, 186–199

29 Paik, W.K. et al. (1970) Nonenzymatic acetylation of histones with acetyl-CoA. Biochim. Biophys. Acta 213, 513–522

30 Tanner, K.G. et al. (1999) Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. J. Biol. Chem. 274, 18157–18160

31 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

32 Weinert, B.T. et al. (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in E. coli. Mol. Cell 51, 265–272

33 Liu, Z. et al. (2014) CPLM: a database of protein lysine modifications. Nucleic Acids Res. 42, D531–6

34 Liu, Z. et al. (2011) CPLA 1.0: an integrated database of protein lysine acetylation. Nucleic Acids Res. 39, D1029–34

35 Allfrey, V.G. et al. (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

36 Brownell, J.E. et al. (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell 84, 843–851

37 Taunton, J. et al. (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science 272, 408–411

38 Guarente, L. (2011) Sirtuins, Aging, and Medicine. N. Engl. J. Med. 364, 2235–2244

39 Haberland, M. et al. (2009) The many roles of histone deacetylases in development and physiology: implications for disease and therapy. Nat. Rev. Genet. 10, 32–42

40 Shogren-Knaak, M. et al. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science 311, 844–847

41 Neumann, H. et al. (2009) A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. Mol. Cell 36, 153–163

42 Filippakopoulos, P. et al. (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. Cell 149, 214–231

43 Piperno, G. et al. (1987) Microtubules containing acetylated alpha-tubulin in mammalian cells in culture. J. Cell Biol. 104, 289–302

44 Hallows, W.C. et al. (2006) Sirtuins deacetylate and activate mammalian acetyl-CoA synthetases. Proc. Natl. Acad. Sci. U. S. A. 103, 10230–10235

45 Schwer, B. et al. (2006) Reversible lysine acetylation controls the activity of the mitochondrial enzyme acetyl-CoA synthetase 2. Proc. Natl. Acad. Sci. U. S. A. 103, 10224–10229

46 Starai, V.J. et al. (2002) Sir2-dependent activation of acetyl-CoA synthetase by

deacetylation of active lysine. Science 298, 2390–2392

47 Wallace, D.C. (2005) A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. Annu. Rev. Genet. 39, 359–407

48 Rardin, M.J. et al. (2013) Label-free quantitative proteomics of the lysine acetylome in mitochondria identifies substrates of SIRT3 in metabolic pathways. Proc. Natl. Acad. Sci. U. S. A. 110, 6601–6606

49 Dittenhafer-Reed, K.E. et al. (2015) SIRT3 mediates multi-tissue coupling for metabolic fuel switching. Cell Metab. 21, 637–646

50 Lombard, D.B. et al. (2007) Mammalian Sir2 homolog SIRT3 regulates global mitochondrial lysine acetylation. Mol. Cell. Biol. 27, 8807–8814

51 Park, J. et al. (2013) SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways. Mol. Cell 50, 919–930

52 Laurent, G. et al. (2013) SIRT4 coordinates the balance between lipid synthesis and catabolism by repressing malonyl CoA decarboxylase. Mol. Cell 50, 686–698

53 Ghanta, S. et al. (2013) Mitochondrial protein acetylation as a cell-intrinsic, evolutionary driver of fat storage: chemical and metabolic logic of acetyl-lysine modifications. Crit. Rev. Biochem. Mol. Biol. 48, 561–574

54 Wagner, G.R. and Hirschey, M.D. (2014) Nonenzymatic protein acylation as a carbon stress regulated by sirtuin deacylases. Mol. Cell 54, 5–16

55 Baeza, J. et al. (2015) Site-specific reactivity of nonenzymatic lysine acetylation. ACS Chem. Biol. 10, 122–128

56 Someya, S. et al. (2010) Sirt3 mediates reduction of oxidative damage and prevention of age-related hearing loss under caloric restriction. Cell 143, 802–812

57 Sol, E.M. et al. (2012) Proteomic investigations of lysine acetylation identify diverse substrates of mitochondrial deacetylase sirt3. PLoS One 7, e50545

58 Tao, R. et al. (2010) Sirt3-mediated deacetylation of evolutionarily conserved lysine 122 regulates MnSOD activity in response to stress. Mol. Cell 40, 893–904

59 Chen, Y. et al. (2011) Tumour suppressor SIRT3 deacetylates and activates manganese superoxide dismutase to scavenge ROS. EMBO Rep. 12, 534–541

60 Qiu, X. et al. (2010) Calorie restriction reduces oxidative stress by SIRT3-mediated SOD2 activation. Cell Metab. 12, 662–667

61 Yu, W. et al. (2012) SIRT3 protein deacetylates isocitrate dehydrogenase 2 (IDH2) and regulates mitochondrial redox status. J. Biol. Chem. 287, 14078–14086

62 Schlicker, C. et al. (2008) Substrates and regulation mechanisms for the human mitochondrial sirtuins Sirt3 and Sirt5. J. Mol. Biol. 382, 790–801

63 Hirschey, M.D. et al. (2010) SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation. Nature 464, 121–125

64 Smith, B.C. et al. (2011) SIRT3 substrate specificity determined by peptide arrays and machine learning. ACS Chem. Biol. 6, 146–157

65 Yamada, T. et al. (2011) iPath2.0: interactive pathway explorer. Nucleic Acids Res. 39, W412–5

66 Calvo, S.E. et al. (2015) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res. DOI: 10.1093/nar/gkv1003

67 Still, A.J. et al. (2013) Quantification of Mitochondrial Acetylation Dynamics Highlights Prominent Sites of Metabolic Regulation. J. Biol. Chem. 288, 26209–26219

68 Fernandes, J. et al. (2015) Lysine Acetylation Activates Mitochondrial Aconitase in the

Heart. Biochemistry 54, 4008–4018

69 Xue, L. et al. (2012) Acetylation-dependent regulation of mitochondrial ALDH2 activation by SIRT3 mediates acute ethanol-induced eNOS activation. FEBS Lett. 586, 137–142

70 Bao, J. et al. (2010) SIRT3 is regulated by nutrient excess and modulates hepatic susceptibility to lipotoxicity. Free Radic. Biol. Med. 49, 1230–1237

71 Yang, H. et al. (2015) SIRT3-dependent GOT2 acetylation status affects the malate-aspartate NADH shuttle activity and pancreatic tumor growth. EMBO J. 34, 1110–1125

72 Shimazu, T. et al. (2010) SIRT3 deacetylates mitochondrial 3-hydroxy-3-methylglutaryl CoA synthase 2 and regulates ketone body production. Cell Metab. 12, 654–661

73 Hirschey, M.D. et al. (2010) SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation. Nature 464, 121–125

74 Bharathi, S.S. et al. (2013) Sirtuin 3 (SIRT3) protein regulates long-chain acyl-CoA dehydrogenase by deacetylating conserved lysines near the active site. J. Biol. Chem. 288, 33837–33847

75 Yang, Y. et al. (2010) NAD+-dependent deacetylase SIRT3 regulates mitochondrial protein synthesis by deacetylation of the ribosomal protein MRPL10. J. Biol. Chem. 285, 7417–7429

76 Ahn, B.-H. et al. (2008) A role for the mitochondrial deacetylase Sirt3 in regulating energy homeostasis. Proc. Natl. Acad. Sci. U. S. A. 105, 14447–14452

77 Cheng, Y. et al. (2013) Interaction of Sirt3 with OGG1 contributes to repair of mitochondrial DNA and protects from apoptotic cell death under oxidative stress. Cell Death Dis. 4, e731

78 Samant, S.A. et al. (2014) SIRT3 deacetylates and activates OPA1 to regulate mitochondrial

dynamics during stress. Mol. Cell. Biol. 34, 807–819

79 Wu, Y.-T. et al. (2013/1) Regulation of mitochondrial FoF1ATPase activity by Sirt3-catalyzed deacetylation and its deficiency in human cells harboring 4977 bp deletion of mitochondrial DNA. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 1832, 216–227

80 Hallows, W.C. et al. (2011) Sirt3 promotes the urea cycle and fatty acid oxidation during dietary restriction. Mol. Cell 41, 139–149

81 Li, D. et al. (2012) Resveratrol stimulates cortisol biosynthesis by activating SIRT-dependent deacetylation of P450scc. Endocrinology 153, 3258–3268

82 Jing, E. et al. (2013) Sirt3 regulates metabolic flexibility of skeletal muscle through reversible enzymatic deacetylation. Diabetes 62, 3404–3417

83 Shulga, N. et al. (2010) Sirtuin-3 deacetylation of cyclophilin D induces dissociation of hexokinase II from the mitochondria. J. Cell Sci. 123, 894–902

84 Wei, L. et al. (2013) Oroxylin A induces dissociation of hexokinase II from the mitochondria and inhibits glycolysis by SIRT3-mediated deacetylation of cyclophilin D in breast carcinoma. Cell Death Dis. 4, e601

85 Cimen, H. et al. (2010) Regulation of succinate dehydrogenase activity by SIRT3 in mammalian mitochondria. Biochemistry 49, 304–311

86 Finley, L.W.S. et al. (2011) Succinate dehydrogenase is a direct target of sirtuin 3 deacetylase activity. PLoS One 6, e23295

87 Inuzuka, H. et al. (2012) Acetylation-dependent regulation of Skp2 function. Cell 150, 179–193

88 Zhang, Y. et al. (2015) SIRT3 and SIRT5 regulate the enzyme activity and cardiolipin

binding of very long-chain acyl-CoA dehydrogenase. PLoS One 10, e0122297

89 Neumann, H. et al. (2008) Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. Nat. Chem. Biol. 4, 232–234

90 Albaugh, B.N. et al. (2011) Autoacetylation of the histone acetyltransferase Rtt109. J. Biol. Chem. 286, 24694–24701

91 Baeza, J. et al. (2014) Stoichiometry of site-specific lysine acetylation in an entire proteome. J. Biol. Chem. 289, 21326–21338

92 Weinert, B.T. et al. (2015) Analysis of acetylation stoichiometry suggests that SIRT3 repairs nonenzymatic acetylation lesions. EMBO J. DOI: 10.15252/embj.201591271

93 Weinert, B.T. et al. (2014) Acetylation dynamics and stoichiometry in Saccharomyces cerevisiae. Mol. Syst. Biol. 10, 716

94 Yu, B.J. et al. (2008) The diversity of lysine-acetylated proteins in Escherichia coli. J. Microbiol. Biotechnol. 18, 1529–1536

95 Zhang, J. et al. (2009) Lysine acetylation is a highly abundant and evolutionarily conserved modification in Escherichia coli. Mol. Cell. Proteomics 8, 215–225

96 Schwer, B. et al. (2009) Calorie restriction alters mitochondrial protein acetylation. Aging Cell 8, 604–606

97 Wang, Q. et al. (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. Science 327, 1004–1007

98 Yang, L. et al. (2011) The fasted/fed mouse metabolic acetylome: N6-acetylation differences suggest acetylation coordinates organ-specific fuel switching. J. Proteome Res. 10, 4134–4149

99 Weinert, B.T. et al. (2011) Proteome-wide mapping of the Drosophila acetylome

demonstrates a high degree of conservation of lysine acetylation. Sci. Signal. 4, ra48

100 Simon, G.M. et al. (2012) Quantitative assessment of the impact of the gut microbiota on lysine epsilon-acetylation of host proteins using gnotobiotic mice. Proc. Natl. Acad. Sci. U. S. A. 109, 11133–11138

101 Henriksen, P. et al. (2012) Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in Saccharomyces cerevisiae. Mol. Cell. Proteomics 11, 1510–1522

102 Lundby, A. et al. (2012) Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. Cell Rep. 2, 419–431

103 Foster, D.B. et al. (2013) The cardiac acetyl-lysine proteome. PLoS One 8, e67513

104 Zhang, Q.-F. et al. (2013) Reversibly acetylated lysine residues play important roles in the enzymatic activity of Escherichia coli N-hydroxyarylamine O-acetyltransferase. FEBS J. 280, 1966–1979

105 Nakayasu, E.S. et al. (2014) A method to determine lysine acetylation stoichiometries. Int. J. Proteomics 2014, 730725

106 Kuhn, M.L. et al. (2014) Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation. PLoS One 9, e94816

107 Schölz, C. et al. (2015) Acetylation site specificities of lysine deacetylase inhibitors in human cells. Nat. Biotechnol. 33, 415–423

108 Svinkina, T. et al. (2015) Deep, Quantitative Coverage of the Lysine Acetylome Using Novel Anti-acetyl-lysine Antibodies and an Optimized Proteomic Workflow. Mol. Cell. Proteomics 14, 2429–2440

109 Elia, A.E.H. et al. (2015) Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. Mol. Cell 59, 867–881

110 Zhang, Z. et al. (2011) Identification of lysine succinylation as a new post-translational modification. Nat. Chem. Biol. 7, 58–63

111 Garrity, J. et al. (2007) N-lysine propionylation controls the activity of propionyl-CoA synthetase. J. Biol. Chem. 282, 30239–30245

112 Chen, Y. et al. (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. Mol. Cell. Proteomics 6, 812–819

113 Fritz, K.S. et al. (2013) Ethanol metabolism modifies hepatic protein acylation in mice. PLoS One 8, e75868

114 Jiang, H. et al. (2013) SIRT6 regulates TNF-α secretion through hydrolysis of long-chain fatty acyl lysine. Nature 496, 110–113

115 Tan, M. et al. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell 146, 1016–1028

116 Pougovkina, O. et al. (2014) Aberrant protein acylation is a common observation in inborn errors of acyl-CoA metabolism. J. Inherit. Metab. Dis. 37, 709–714

117 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

118 Weinert, B.T. et al. (2013) Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. Cell Rep. 4, 842–851

119 Wacker, S.A. et al. (2010) Ligand-induced changes in the conformational stability and flexibility of glutamate dehydrogenase and their role in catalysis and regulation. Protein Sci. 19, 1820–1829

120 Alex, S. and Bell, J.E. (1980) Dual nucleotide specificity of bovine glutamate

dehydrogenase. The role of negative co-operativity. Biochem. J 191, 299–304

121 Peterson, P.E. and Smith, T.J. (1999) The structure of bovine glutamate dehydrogenase provides insights into the mechanism of allostery. Structure 7, 769–782

122 Mariño, G. et al. (2014) Regulation of autophagy by cytosolic acetyl-coenzyme A. Mol. Cell 53, 710–725

Chapter 2: Stoichiometry of site-specific lysine acetylation in an entire proteome

This chapter was published as a first author publication in The Journal of Biological Chemistry with the following citation:

Baeza, J., Dowell, J. A., Smallegan, M. J., Fan, J., Amador-Noguez, D., Khan, Z., & Denu, J. M. (2014). Stoichiometry of site-specific lysine acetylation in an entire proteome. The Journal of Biological Chemistry, 289(31), 21326–21338.

Author Contributions: J.B., J.A.D., and J.M.D. conceived the idea for quantifying acetylation stoichiometry. J.B. and J.A.D. executed MS experiments. Z.K. developed software to analyze data. J.B. M.J.S. Z.K. J.M.D. analyzed results. J.F. D.A.N. performed metabolomics experiments. J.B., J.A.D. J.M.D. drafted the manuscript. All authors edited final draft. J.M.D. is the corresponding author. We thank Jorge Escalante for providing the BL21 (DE3) ΔCobB E. coli strain and the Sushmita Roy lab for helpful discussions on performing the network analysis. 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 the National Science Foundation Graduate Research Fellowship Program (NSF-GRFP) DGE-1256259 (J.B.).

2.1 Abstract

Acetylation of lysine ε-amino groups influences many cellular processes and has been mapped to thousands of sites across many organisms. Stoichiometric information of acetylation is essential to accurately interpret biological significance. Here we developed and employed a novel method for directly quantifying stoichiometry of site-specific acetylation in the entire proteome of E. coli. By coupling isotopic labeling and a novel pairing algorithm, our approach performs an in silico enrichment of acetyl-peptides, circumventing the need for immunoenrichment. We investigated the function of the sole NAD^+-dependent protein deacetylase, CobB, on both site-specific and global acetylation. We quantified 2206 peptides from 899 proteins and observed a wide distribution of acetyl-stoichiometry, ranging from less than 1% up to 98%. Bioinformatic analysis revealed that metabolic enzymes, which either utilize or generate acetyl-CoA, and proteins involved in transcriptional and translational processes displayed the highest degree of acetylation. Loss of CobB led to increased global acetylation at low stoichiometry sites, induced site-specific changes at high stoichiometry sites, and biochemical analysis revealed altered acetyl-CoA metabolism. Thus, this study demonstrates that sirtuin deacetylase deficiency leads to both site-specific and global changes in protein acetylation stoichiometry, affecting central metabolism.

2.2 Introduction

Acetylation of the ε-amino group of lysine (K) residues is now considered a major regulatory protein modification that influences many cellular processes including protein-protein interactions, protein-DNA interactions, stability, cellular localization and enzymatic activity. (1, 2) Acetylation occurs by enzymatic and non-enzymatic mechanisms. (3) Non-enzymatic acetylation occurs through nucleophilic attack of the lysine side chain on esters of the metabolic intermediates, acetyl-CoA and acetyl-phosphate. (4, 5) In prokaryotes, acetyl-phosphate was reported to serve as the major acetyl donor. (4) Removal of acetyl-modifications requires protein deacetylases. (6)

Mass spectrometry-based proteomic studies have mapped thousands of acetylated sites in a wide range of organisms. (7–11) Recently, acetyl-proteome measurements have utilized labeling strategies, such as tandem mass tags or stable isotope labeling by amino acids in cell culture (SILAC), followed by immunoenrichment of acetylated peptides and mass spectrometry. (7–9, 12, 13) While these approaches are valuable for comparing relative changes across conditions, they do not provide direct information on stoichiometry at individual sites. (1, 13) Stoichiometric information is essential to accurately interpret the biological significance of these acetylation sites. For example, if only relative acetylation can be determined, as is the case with current acetylation methods, then a 1 to 5% and a 20 to 100% change would appear identical, i.e., a 5-fold change. The biological significance of these changes are entirely different, especially in cases in which acetylation is inhibitory.

Mass spectrometry has been used to examine phosphorylation stoichiometry in large-scale studies. (13–15) One central issue that is addressed by these approaches is that post-translationally modified peptides and their corresponding unmodified counterparts have varying ionization

efficiencies, and thus, they cannot be directly compared. The issue of differing ionization efficiency can be addressed by making indirect measurements that incorporate correlated changes in phosphorylated and non-phosphorylated versions of peptides that are corrected for overall protein level changes across experiments with very different amounts of phosphorylation. (16) Very recently this indirect strategy was applied to estimate acetylation stoichiometry in log phase yeast compared to growth arrested yeast. (17) Yet, this strategy requires quantifying modified and unmodified versions of the same peptide, which are difficult to detect even with enrichment. It is biased in that it can only measure stoichiometry for peptides where PTM levels are dramatically different across conditions. Furthermore, it is unclear how the estimates behave when the changes in PTM status across conditions are small. In the context of phosphorylation, these challenges can be addressed by treating an identical stable isotope labeled sample with a phosphatase and measuring and comparing the increased abundance of the unphosphorylated peptide. (15) The percent increase provides a measurement of the stoichiometry of a phosphorylated site that circumvents the need for enrichment and the requirement for a change in the level of a PTM site. Yet no such approach exists for lysine acetylation.

Here, we develop and utilized the first unbiased and direct method for quantifying the stoichiometry of site-specific acetylation at the proteome-wide scale, without enrichment. Our approach uses a stable isotope chemical labeling step that acetylates all unmodified lysines resulting in a “heavy” and “light” acetyl-lysine pair across the entire proteome. This pair is then analyzed using high-resolution, high-accuracy mass spectrometry to yield proteome-wide, site-specific stoichiometry. Using this approach, we interrogated the acetyl-proteome of E. coli and investigated the functional consequences of genetically removing the sole NAD^+-dependent protein deacetylase, CobB, on both site-specific and global acetylation. We quantified 2206

peptides from 899 proteins in the range of less than 1% up to 98% acetylation. Loss of CobB resulted in a slight increase of global acetylation at low stoichiometry sites, as well as site-specific changes at high stoichiometry sites. Proteins with the highest stoichiometry included those involved in central metabolism, transcription and translation.

2.3 Experimental Procedures

2.3.1 Sample Preparation

2.3.1.1 Cell culture and sample preparation

E. coli BL21 (DE3) Wild type and ΔCobB (with empty pQE80) were grown in 2XYT media under ampicillin selection overnight. Cells were harvested, flash frozen and stored in -80 °C. Frozen cells were resuspended in ~10 volumes 8M urea solution (8 M urea, 5 mM DTT, 100 mM ammonium bicarbonate pH 8) and lysed by sonication. After centrifugation, the protein concentration was determined using Bradford reagent (Bio-Rad).

2.3.1.2 AceCS2 expression

AceCS2 pQE80 was transformed into E. coli BL21 (DE3) and protein expression was performed as previously described. (18)

2.3.2 Mass Spectrometry

2.3.2.1 Protein chemical acetylation and digestion

Equal amount of protein (100 µg) was diluted using 8M urea solution (2 µg/uL) and incubated at 60 °C, 1000 RPM, for 30 minutes on the Thermomixer (Eppendorf). Cysteine alkylation with iodoacetamide was performed for 30 minutes in the dark. Chemical acetylation was performed as previously described, with slight modification. (19) Equal volume of ammonium bicarbonate solution (1M NH_4OAc, 8M urea, 100 mM ammonium bicarbonate pH 8) was added to each sample. ~20 µmol of acetic anhydride [acetic anhydride, acetic anhydride-d_6 or acetic anhydride-$^{13}C_4$,d_6 (Sigma Aldrich)] was added to each sample and incubated on Thermomixer at 4 °C, 1000 RPM, for 20 minutes. After incubation, the pH was raised to ~8 using ammonium hydroxide and checked using litmus paper. Chemical acetylation was repeated twice more and buffer exchange occurred with 50 mM ammonium bicarbonate (pH 8), 10% methanol using 10K

MWCO spin filters (Millipore). Trypsin was added at a 1:100 ratio and incubated overnight, 37 °C, 300 RPM on Thermomixer.

2.3.2.2 LC-MS/MS

Peptides were separated with a Dionex Ultimate 3000 RSLCnano HPLC using a Waters Atlantis dC18 (100 µm x 150 mm) reverse phase column. The mobile phase consisted of (A) water with 0.1% formic acid and (B) acetonitrile with 0.1% formic acid. Peptides were eluted with a linear gradient of 2 – 40% B at a flow rate of 0.7 µl/minute over 120 minutes and introduced into a hybrid quadrupole-orbitrap mass spectrometer (Thermo Q Exactive) by nanoelectrospray ionization (Thermo Nanospray Flex). The MS survey scan was performed in positive ion mode with a resolution of 70,000, AGC of 1E6, maximum fill time of 100 ms, and scan range of 400 to 2000 m/z. Data dependent MS/MS was performed with a resolution of 17,500, AGC of 1E5, maximum fill time of 64 ms, isolation window of 1.5 Da, underfill ratio of 0.1%, normalized collision energy of 28, dynamic exclusion of 20 seconds, and a loop count of 20. The source voltage was set at 2000 V and capillary temperature at 250 °C.

2.3.2.3 MS data analysis

Data was analyzed using a modified version of the December 25th, 2013 release of open-source, quantitative mass spectrometry analysis tool PVIEW downloaded from (http://compbio.cs.princeton.edu/pview). (20, 21) We generated a peptide database with cleavage only at arginine, as trypsin did not cleave at acetylated lysines. Up to 2 missed cleavages were allowed. MS1 precursor tolerance was set to ±10 ppm and MS2 fragment mass tolerance was set to ±15 ppm. Peptide identifications were obtained at a false discovery rate (FDR) of 1%. Estimation of false discovery rate (FDR) was carried out using only peptide spectrum matches from paired XICs.

2.3.2.4 Acetylation stoichiometry calculations

The equation used to calculate stoichiometry of single lysine containing peptides:

$$\frac{XIC_L}{XIC_L + XIC_H}$$

and double lysine containing peptides:

$$\frac{XIC_L + \frac{1}{2}XIC_M}{XIC_L + XIC_M + XIC_H}$$

where XIC_L is the extracted ion chromatogram of the light acetyl-peptide, XIC_H is the extracted ion chromatogram of the heavy acetyl-peptide (acetic anhydride-d_6 or acetic anhydride-$^{13}C_4$,d_6) and XIC_M, when present, is the extracted ion chromatogram for the double lysine containing peptide with a light and heavy acetyl-lysine.

Due to the isotopic purity of the commercially available, acetic anhydride-$^{13}C_4$,d_6 labeling reagent (97 atom % D, 99 atom % ^{13}C), a global correction factor was applied to the stoichiometry of the Δ5-acetic anhydride labeled sample.

$$\frac{(XIC_L + \frac{1}{2} XIC_M) - (\frac{0.04}{0.96}) \cdot (\frac{1}{2} XIC_M + XIC_H)}{XIC_L + XIC_M + XIC_H}$$

2.3.3 Determination of chemical labeling efficiency

Equal amount of protein (BSA or E. coli lysate with expressed mouse AceCS2) was prepared as described above. An unlabeled sample was used as a control. Protein digestion, LC-MS/MS, and database search was performed as described above.

Unmodified tryptic peptides of the labeled and control sample were monitored and compared to determine labeling efficiency. Briefly, the peak area for a lysine containing peptide in the control sample is measured. In the labeled sample, the same lysine site would be acetylated to some degree, generating a larger acetyl-peptide and also the unmodified peptide proportionate to the degree of chemical labeling. As a result, the unmodified peptide in both the labeled and control samples can be compared to determine labeling efficiency. For normalization across samples, a peptide lacking lysine residues is used. Peak areas for each peptide were determined using Xcalibur 2.2 SP1 (Thermo Scientific).

2.3.4 Stoichiometry curve determination

Equal amount of BSA was chemically acetylated with acetic anhydride or acetic anhydride-$^{13}C_4$,d_6 and digested as described above. Light and heavy BSA peptides were resuspended to equal concentrations and then mixed at varying ratios corresponding to 1, 5, 10, 20, 25, 40, 50, 60, 75, 80, 90, 95, 99% acetylation. The BSA samples were then spiked into a trypsin digested E. coli proteome at equal concentrations, analyzed by LC-MS/MS and data processed as described above.

2.3.5 Bioinformatics

Gene ontology (GO) and pathway analysis was performed using DAVID v6.7. (22, 23) For the enrichment analysis, an unlabeled, trypsin-digested, E. coli proteome was used as the background. Network analysis was performed using STRING v9.1 with proteins having an

acetylation stoichiometry greater than 7% as the input. (24, 25) The network was visualized using Cytoscape v3.1.0. (26)

2.3.6 Metabolite analysis

In order to measure intracellular metabolite levels, cells were grown in liquid culture with 2XYT media to early stationary phase. Metabolites were extracted by filtering 5ml culture on nylon filters (Millipore) and quenching the cells in cold extraction solvent (40:40:20 methanol: acetonitrile: water). Cell extracts were centrifuged followed by one round of re-extraction. Supernatants from two rounds of extraction were combined, dried under N_2, and resuspended in HPLC water for analysis. The metabolites were analyzed by LC-MS as previously described. (27) Briefly, LC separation on a Synergy Fusion-RP column (100 mm × 2 mm, 2.5 μm particle size, Phenomenex, Torrance, CA) using a gradient of solvent A (97:3 H2O/MeOH with 10 mM tributylamine and 5 mM ammonium bicarbonate), and solvent B (100% MeOH) was coupled by negative mode electrospray ionization to a stand-alone Q Exactive orbitrap mass spectrometer (Thermo Scientific).

2.4 Results

2.4.1 Method development and validation

To measure acetylation stoichiometry on the proteome scale, we developed a robust workflow using chemical acetylation with isotopic acetic anhydride followed by trypsin digestion and high-resolution mass spectrometry (**Figure 2-1**). The general strategy involves preparing denatured protein extracts from cells, chemically acetylating free lysine residues with isotopic acetic anhydride, followed by trypsin cleavage and MS analysis. Using this process, every lysine bears an acetyl group, either the light version derived from endogenous acetylation or the heavy version originating from in vitro chemical acetylation. The trypsin cleavage step produces chemically identical peptides, which have matching retention times. The corresponding pairs of peptides that differ in mass are resolved using mass spectrometry and direct stoichiometry is determined by dividing the light peak area over the sum of the light and heavy peak areas.

To ensure accurate determination of stoichiometry, we first evaluated acetic anhydride labeling efficiency within a simple and complex protein sample. An E. coli whole-cell lysate was chemically modified with acetic anhydride and qualitatively assessed for the degree of acetylation. Western blot analysis revealed that acetic anhydride chemically acetylated E. coli proteins across the full range of protein sizes (**Figure 2-2A**). To quantitatively assess the degree of chemical acetylation in a simple protein sample, equal amounts of labeled and unlabeled bovine serum albumin (BSA) were digested and analyzed by mass spectrometry. Unmodified BSA tryptic peptides in both samples were monitored and compared to determine the degree of acetylation. The labeling efficiency across eight out of nine BSA peptides was greater than 98% (**Figure 2-2B**). To determine the labeling efficiency in a complex sample, the mouse AceCS2 protein was recombinantly expressed in BL21 (DE3) bacteria. The cells were lysed, protein was labeled with

increasing amounts of acetic anhydride and the labeling efficiency was calculated for peptides of AceCS2 protein. Across the 8 sites in AceCS2, the labeling efficiency was >98% (**Table 2-1**). These results indicate that proteins in complex mixtures can be chemically acetylated with high efficiency.

We developed a computational method for stoichiometric determination of acetylation that avoids the need for immunoenrichment. Existing proteome-wide acetylation studies require an enrichment step in order to increase sensitivity in the PTM analysis. (13) Sensitivity is achieved by the increased abundance and subsequent identification of acetylated peptides. In this study, we achieve high sensitivity by coupling a novel computational algorithm, implemented in a peptide quantification software tool called PVIEW, to the chemical acetylation strategy. This algorithm identifies acetyl-peptides (heavy or light) using one of two databases (**Figure 2-3A**) and alternates between iterations of heavy-to-light and light-to-heavy XIC pairing (**Figure 2-3B, D**), using the high-abundant species as a proxy to identify the corresponding, lower abundant, peptide pair. In essence, the algorithm identifies the best acetyl-peptide and based on charge and isotopic shift, the corresponding peptide pair is identified, matched, and quantified. In this manner, our algorithm performs an in silico enrichment of acetylated peptides.

We demonstrated the capability of measuring acetylation stoichiometry across a broad range. BSA protein was chemically acetylated with either acetic anhydride (light) or isotopic acetic anhydride (heavy) followed by trypsin digestion. Peptides were then mixed at distinct ratios corresponding to 1 to 99% acetylation and added to a trypsin digested E. coli proteome. This experiment served to mimic the full range of endogenous acetylation that might be encountered from diverse biological conditions. Samples were analyzed by LC-MS/MS and the stoichiometries of five BSA acetyl-peptides were quantified using our novel quantification algorithm. As a

representative, the MS/MS spectrum for the acetyl-peptide, ALK(Ac)AWSVAR, is shown in Figure 2-4A. MS1 spectra corresponding to 10, 50, and 90% stoichiometry are shown in Figure 2-4B, D, respectively. The stoichiometry curve encompassing 1 to 99% stoichiometry is shown in Figure 2-4E. Linear regression analysis of the stoichiometry curves from all five acetyl-BSA peptides show a high correlation with R^2 values ranging between 0.970 – 0.995 (**Figure 2-4F**). These results establish the feasibility of accurately determining stoichiometry across a wide range of values in a complex mixture.

Figure 2-1: Diagram of method for determining direct acetylation stoichiometry

Extracted protein was denatured, chemically acetylated using isotopic acetic anhydride followed by trypsin digestion. Peptides were analyzed by LC-MS/MS and quantified to determine site-specific stoichiometry.

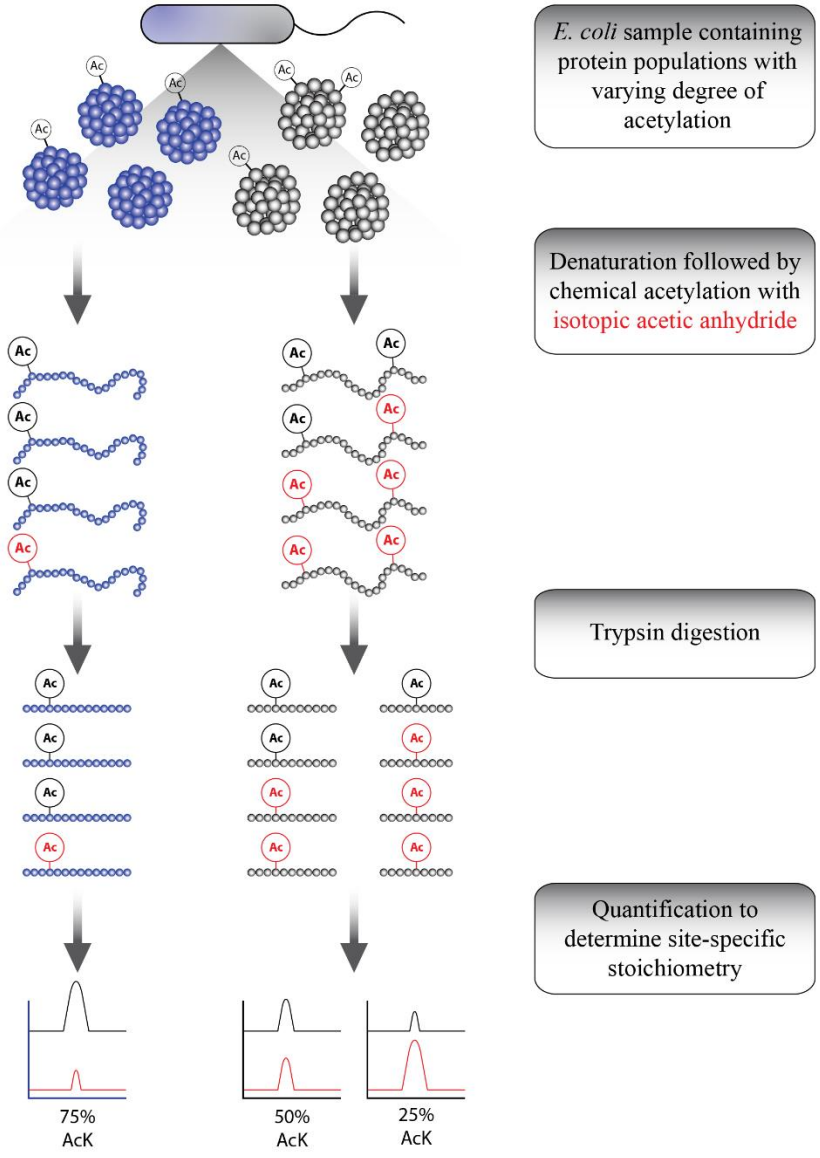


Figure 2-2: Acetic anhydride comprehensively modifies lysine residues on a wide range of protein sizes and to a high degree

Chemical acetylation by acetic anhydride was assessed qualitatively and quantitatively. (A) Western blot analysis with α-AcK antibody with an unmodified and modified E. coli sample shows chemical labeling across a full range of proteins in a complex mixture. (B) The degree of acetylation was determined using bovine serum albumin (BSA). The indicated peptides were quantified in the unmodified and modified sample. The abundance of each peptide correlates to the degree of chemical acetylation by acetic anhydride.

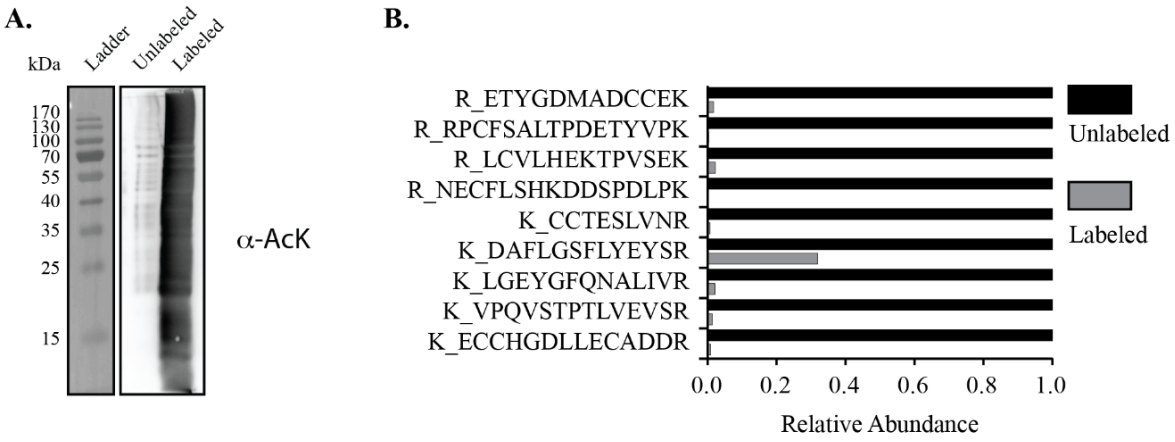


Figure 2-3: Diagram of LC-MS/MS data analysis algorithm used for identifying isotopic pairs for stoichiometry determination

The pairing algorithm was carried out using extracted ion chromatograms (XICs). The peaks here are for illustration purposes only. (A) We created two separate MS2 peptide search databases: (1) A “light” database with native acetylation as a fixed modification for each lysine. (2) A “heavy” database where every lysine is modified by the isotopic acetic anhydride. The algorithm used two iterations: (B) light-to-heavy pairing when the native acetyl-peptide was the most highly abundant species, and (C) a heavy-to-light pairing when chemically labeled peptide was the most highly abundant species. In each respective iteration for an XIC with an associated MS2 spectrum, we determined the best matching peptide based on MS2 search score using the light or heavy database. We used the expected isotopic shift to find the corresponding MS1 peak. (D) For low stoichiometry sites, the algorithm is able to perform in silico enrichment. The algorithm uses the MS2 spectrum and MS1 signal of the heavy acetyl-peptide as a proxy for the native acetyl-peptide. Using this proxy information, the algorithm identifies low-level signal from the native acetyl-peptide.

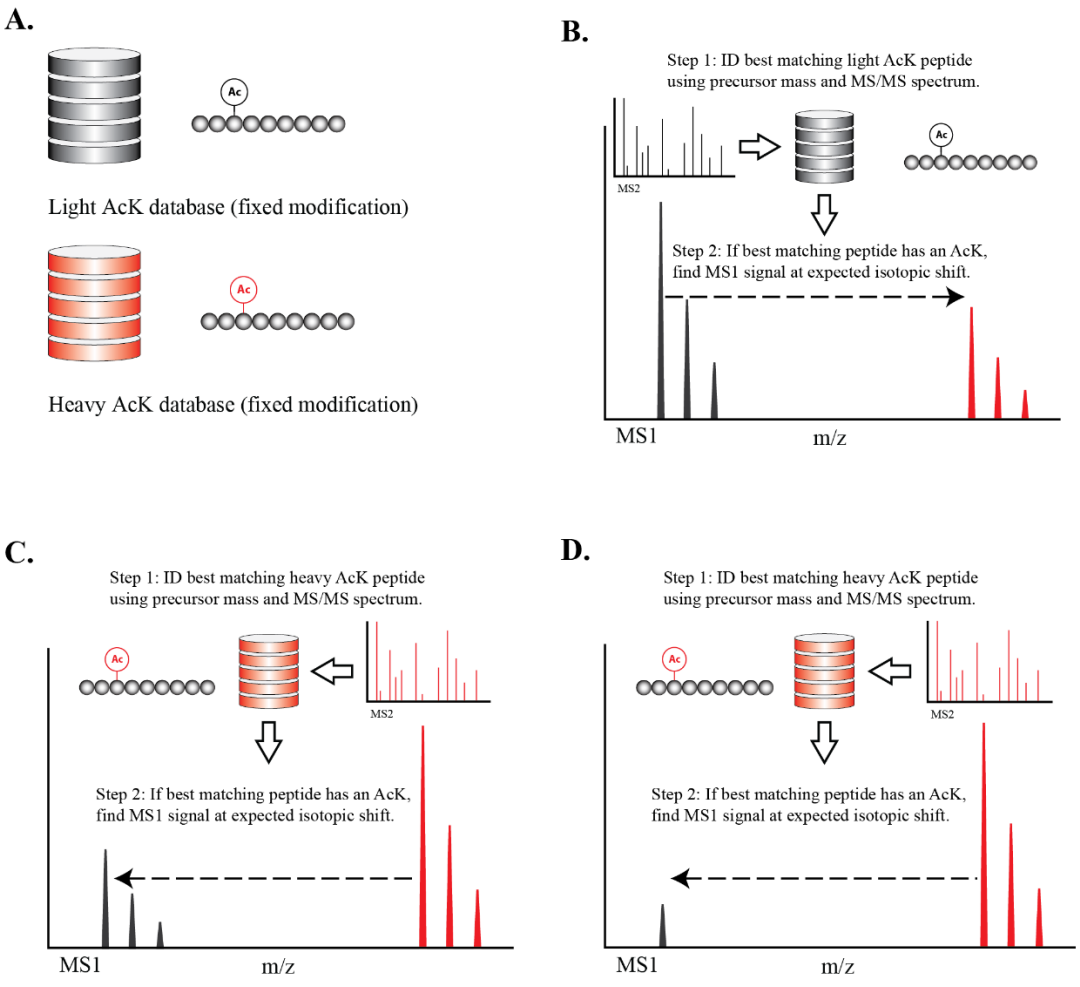
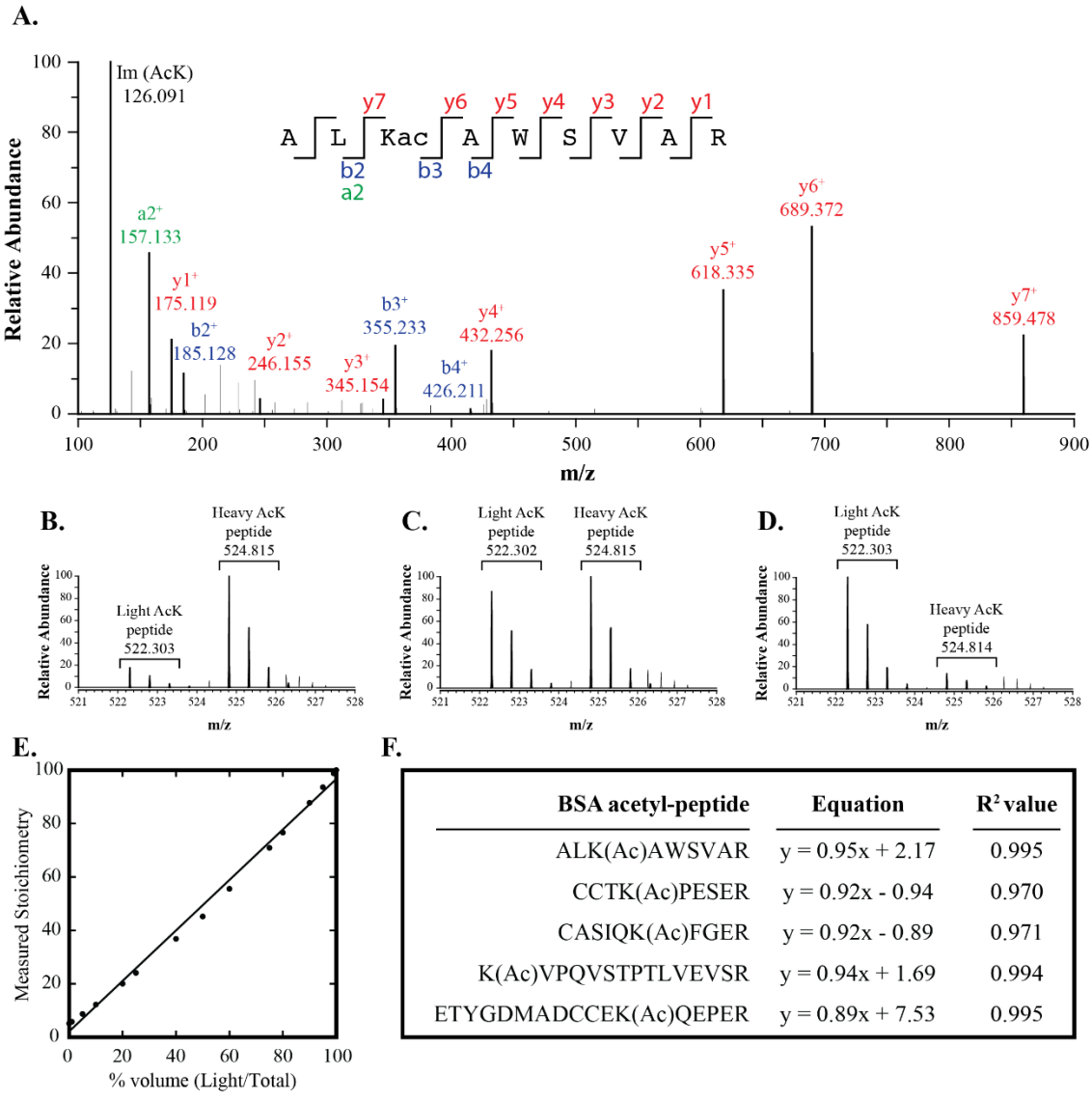

A.
Ac
Light AcK database (fixed modification)
Ac
Heavy AcK database (fixed modification)
B.
Step 1: ID best matching light AcK peptide
using precursor mass and MS/MS spectrum.
MS2
Ac
Step 2: If best matching peptide has an AcK,
find MS1 signal at expected isotopic shift.
MS1
m/z
C.
Step 1: ID best matching heavy AcK peptide
using precursor mass and MS/MS spectrum.
Ac
MS2
Step 2: If best matching peptide has an AcK,
find MS1 signal at expected isotopic shift.
MS1
m/z
D.
Step 1: ID best matching heavy AcK peptide
using precursor mass and MS/MS spectrum.
Ac
MS2
Step 2: If best matching peptide has an AcK,
find MS1 signal at expected isotopic shift.
MS1
m/z

Figure 2-4: Acetylation stoichiometry is detected and quantified across a full range of values

Light and Δ5-acetic anhydride labeled BSA protein digests were mixed at different ratios corresponding to 1 to 99% acetylation followed by MS analysis and quantitation. (A) MS/MS spectrum for the acetylated BSA peptide, ALK(Ac)AWSVAR, is shown. (Product ions shown are charge 1^+). (B – D) MS1 spectrum showing light and heavy forms of the peptide in A (charge 2^+) corresponding to 10, 50, 90% acetylation, respectively with a mass shift of ~2.5 m/z. (E) Scatter plot comparing percent volume input (Light/Total) to the corrected measured stoichiometry of the BSA peptide, ALK(Ac)AWSVAR. Stoichiometry values were corrected due to the isotopic purity of ~96%. (F) Linear regression analysis of all the acetyl-peptides quantified shows the best-fit line and high R^2 value.



BSA acetyl-peptide	Equation	R^2 value
ALK(Ac)AWSVAR	y = 0.95x + 2.17	0.995
CCTK(Ac)PESER	y = 0.92x - 0.94	0.970
CASIQK(Ac)FGER	y = 0.92x - 0.89	0.971
K(Ac)VPQVSTPTLVEVSR	y = 0.94x + 1.69	0.994
ETYGDMADCCEK(Ac)QEPER	y = 0.89x + 7.53	0.995

Table 2-1: Degree of chemical acetylation with increasing amount of acetic anhydride in a complex mixture

An E. coli lysate with expressed AceCS2 protein (Uniprot ID Q99NB1) was used to determine acetylation efficiency by acetic anhydride. Non-acetylated tryptic peptides of the labeled and control sample were monitored and compared to determine labeling efficiency. The XIC for a lysine containing peptide in the control sample is quantified. In the labeled sample, the same lysine site would be acetylated to some degree, generating two peptides after trypsin cleavage: 1). a larger acetyl-peptide and 2). the unmodified peptide, proportionate to the degree of chemical labeling. Therefore, the unmodified peptide from the labeled and control samples can be compared to determine labeling efficiency. For normalization across samples, a peptide lacking lysine residues is used.

Peptide	nmol acetic anhydride / ug protein		
	~200	~300	~400
K_AVITFNQGLR	0.998	0.998	0.998
K_GNVLEGGDVSGALCISQAWPGMAR	0.924	0.931	0.967
K_INQFYGAPTAVR	0.980	0.989	0.990
K_SCPTVQHVLVAHR	0.991	0.999	0.997
K_SPETIALIWER	0.997	0.998	0.997
R_GQDLGDTTTLEDPSVITEILSAFQK	0.969	0.992	0.994
R_LGTAEIEDAMADHPAVPETAVIGYPHDIK	0.990	0.994	0.997
R_TLGSVGEPINHEAWEWLHK	0.972	0.989	0.992

2.4.2 Acetylation stoichiometry values from proteomes of WT and ΔCobB bacteria

Many bacteria, like E. coli, contain a single NAD^+-dependent protein deacetylase that is the ortholog to CobB from Salmonella enterica. In S. enterica, CobB regulates metabolic pathways that permit growth on alternate carbon sources such as short chain fatty acids (SCFA) (28, 29) and citrate. (30) Here, we applied our new method to investigate the role of CobB on both site-specific and global acetylation stoichiometries.

We determined protein acetylation of BL21 (DE3) wild type and ΔCobBstrains, analyzing three biological replicates for each strain. Cells were grown in rich media, harvested during stationary phase, and processed as outlined in Figure 2-1, using acetic anhydride-d_6 (Δ3-acetic anhydride) or acetic anhydride-$^{13}C_4$, d_6 (Δ5-acetic anhydride) for the isotopic chemical labeling. We identified a total of 4025 and 2659 acetyl-lysine sites corresponding to 852 and 576 proteins for the Δ3 and Δ5 acetic anhydride labeling, respectively (**Table 2-2**). As a consequence of all lysines bearing an acetyl moiety, trypsin cleavage occurs only at arginine residues. (31, 32) Using Δ3 and Δ5 acetic anhydride labeling strategies, we quantified stoichiometry for peptides bearing one and two lysine residues, corresponding to 3123 and 2099 peptides, respectively (**Table 2-2**).

We observed a wide distribution of acetyl-stoichiometry, ranging from less than 1% up to 98% (**Figure 2-5A**). Among the peptides quantified, ~2500 displayed stoichiometry between 0-10%, ~500 were acetylated between 10-20%, ~60 between 20-30%, and 70 displayed stoichiometry greater than 30%. Thus, the largest fraction of peptides exhibiting stoichiometry greater than 20% AcK is 4% of the quantified proteome (**Figure 2-4A inset**).

To assess reproducibility and reveal global differences between strains, we compared peptide level acetylation stoichiometry across all Δ5-acetic anhydride prepared biological samples. Stoichiometry values of acetylated peptides within the same biological condition show the highest

correlation with a Spearman's Correlation coefficient of 0.811 – 0.849 for wild type and 0.813 – 0.845 for ΔCobB (**Figure 2-5B**). However, when stoichiometries are compared across biological conditions, a decrease in correlation is observed (0.563 – 0.621), suggesting that the ΔCobBstrain differs substantially in the acetylation profile compared to the WT strain. To assess global alterations in acetylation, we plotted a histogram of peptide level ΔCobB/WT Log_2 fold change. The ΔCobBstrain displayed a slight global increase in acetylation when compared to WT, suggesting that CobB alters the E. coli acetyl proteome under stationary phase growth (**Figure 2-5C**).

We next determined site-specific, acetyl-stoichiometry changes upon CobB deletion by generating a stoichiometry line plot that compares the acetylation of WT and ΔCobBpeptides and quantifies the magnitude change between the two conditions, revealing peptides that have higher stoichiometry in either ΔCobB (red lines) or WT (black lines) strains **(Figure 2-6**). This analysis also revealed three groups of peptide stoichiometry profiles: 1.) peptides with low acetylation stoichiometry and low magnitude change, 2.) peptides with higher acetylation stoichiometry and low magnitude change, and 3.) peptides with different stoichiometries and a high magnitude change between conditions. Group 1 represents the majority of peptides, which is reflected by the low average global acetylation stoichiometry (**Figure 2-6**). Interestingly, group 2 peptides had a higher stoichiometry regardless of CobB status, which suggests these AcK sites are not targets of CobB. Finally, group 3 peptides had the highest magnitude change between WT and ΔCobB, and included acetyl-peptides that increase in the CobB deletion. These include tryptophan-tRNA ligase (TrpS), 30S ribosomal protein S16 (RpsP), peptide deformylase (Def), xanthine phospho-ribosyltransferase (Gpt), and acetyl-CoA synthetase (Acs).

Table 2-2: Acetylation stoichiometry values

The number of acetyl-peptides was quantified across the experimental conditions and strains. This includes total number of acetylation sites quantified, 1K and 2K peptides, and total proteins identified.

	Acetic anhydride labeling				
	Δ3	Δ5	ΔCobB	WT	Total
lysine sites detected	4025	2659	2573	1945	2206
single lysine peptides with stoichiometry	2221	1539	1451	1103	1602
double lysine peptides with stoichiometry	902	560	561	421	603
proteins identified	852	576	827	680	899

Figure 2-5: Site-specific and global acetylation stoichiometry in BL21 (DE3) WT and ΔCobB strains

(A) Distribution of acetylation stoichiometry obtained in WT and ΔCobBstrains and (inset) a pie chart showing the proportion of stoichiometry across the entire E. coli proteome. (B) Acetylated peptides of all biological conditions from the Δ5-acetic anhydride labeling were compared. Left of diagonal shows the scatter plot of acetyl-peptides from each biological condition. Right of diagonal is the Spearman's correlation coefficient. Correlation within biological replicates is high (0.811 – 0.849 for WT and 0.813 – 0.845 for ΔCobB), while correlation decreased across biological conditions (0.5632 – 0.6208). Number in parenthesis represents the total number of peptides analyzed in scatter plot. (C) Histogram of Log_2 fold change between ΔCobBand WT acetyl-peptides. The plot shows a normal distribution and slight shift towards the right.

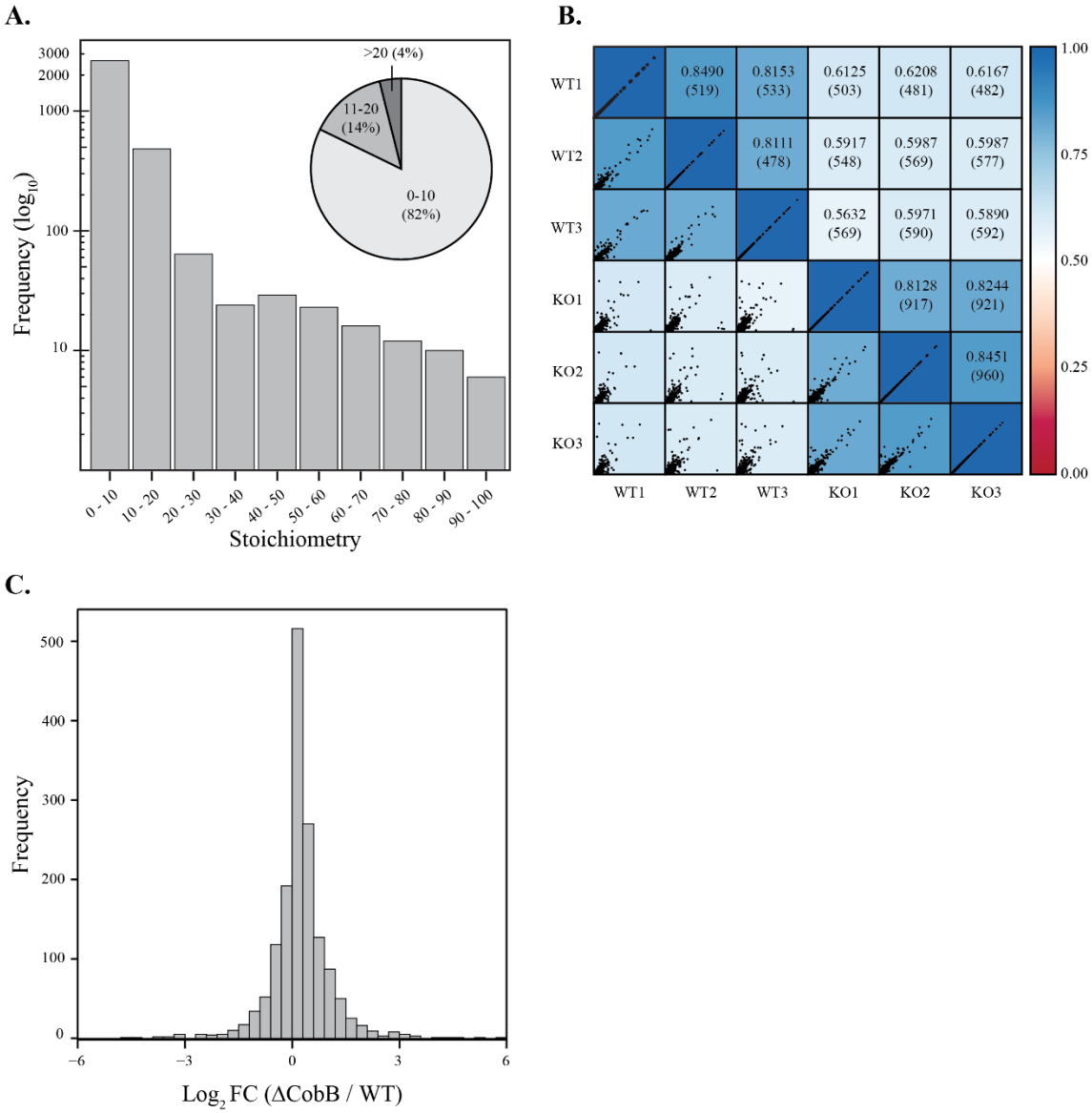
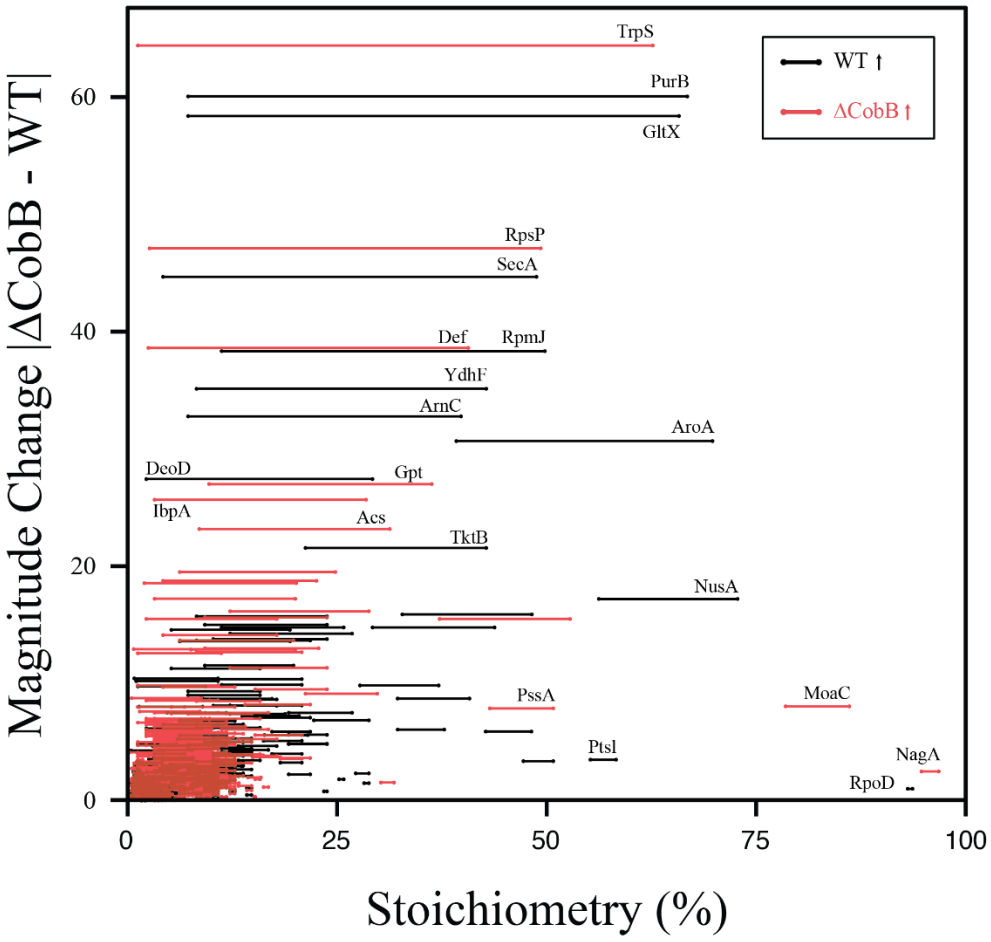

A.
Frequency (log10)
3000
2000
1000
100
10
>20 (4%)
11-20
(14%)
0-10
(82%)
0 - 10
10 - 20
20 - 30
30 - 40
40 - 50
50 - 60
60 - 70
70 - 80
80 - 90
90 - 100
Stoichiometry
B.
WT1
WT2
WT3
KO1
KO2
KO3
0.8490
(519)
0.8153
(533)
0.6125
(503)
0.6208
(481)
0.6167
(482)
0.8111
(478)
0.5917
(548)
0.5987
(569)
0.5987
(577)
0.5632
(569)
0.5971
(590)
0.5890
(592)
0.8128
(917)
0.8244
(921)
0.8451
(960)
1.00
0.75
0.50
0.25
0.00
C.
Frequency
500
400
300
200
100
0
−6
−3
0
3
6
Log2 FC (ΔCobB / WT)

Figure 2-6: Dynamics of acetylation stoichiometry in WT and ΔCobB strains

(A) Stoichiometry line plot demonstrates peptide level stoichiometry in WT and ΔCobBand the magnitude change between strains. Peptides are shown that have a higher stoichiometry in WT (black lines) and ΔCobB (red lines).



2.4.3 Bioinformatic and network analysis of acetylation stoichiometry

To examine whether proteins that display substantial stoichiometry are enriched in specific cellular processes, we performed an enrichment analysis of our unique stoichiometry results using the database for annotation, visualization, and integrated discovery (DAVID). (22, 23) Certain metabolic pathways and ribosomal proteins were significantly enriched above background (**Figure 2-7A)**. Gene Ontology (GO) analysis also revealed that acetyl-CoA metabolic processes, transcription, translation, and tRNA aminoacylation were significantly enriched (**Figure 2-7B**). A network analysis of the identified pathways was performed using the STRING Database. (24, 25) Acetylation stoichiometries along with the number sites were superimposed across the protein interaction network and visualized using Cytoscape. (26) Acetylation sites per protein varied across the proteome (**Figure 2-7C**), therefore we used the highest stoichiometry quantified for each protein in the network. The pentose phosphate pathway, glycolysis, pyruvate metabolism and the TCA cycle revealed proteins with sites of higher stoichiometry (**Figure 2-7D**). The proteins in these pathways were also enriched in total acetylation, as many proteins in these pathways had more than 7 AcK sites (**Figure 2-7D, F**). Proteins involved in transcription and translation displayed many sites of acetylation and high stoichiometries (**Figure 2-7E, F**). This unique network analysis revealed the range of stoichiometries in major cellular pathways and processes.

Total acetylation site number and high stoichiometry enrichment in central metabolic pathways suggests that CobB deletion affects these pathways. To visualize these results, we plotted the TCA cycle and pyruvate metabolism pathways, with acetyl-CoA at the crossroads of these two pathways (**Figure 2-8**). As evident from this analysis, enzymes that either utilize or generate acetyl-CoA exhibited peptides with higher stoichiometry, such as acetyl-CoA carboxylase subunit alpha (AccA), biotin carboxylase (AccC), formate acetyltransferase (PflB), phosphate

acetyltransferase (Pta), acetyl-CoA synthetase (Acs), citrate synthase (GltA), and aldehyde-alcohol dehydrogenase (AdhE) (**Figure 2-8**). Also, proteins in these pathways displayed a higher number of quantifiable peptides with above average stoichiometry. For example, pyruvate kinase II (PykA), phosphoenolpyruvate synthase (PpsA), Pta, aconitate hydratase 1 and 2 (AcnA, AcnB) α-ketoglutarate dehydrogenase E1 component (SucA) and AdhE had at least six peptides with a stoichiometry higher than 7%.

Acetyl-CoA is at the junction of central metabolism. It is striking that most of the metabolic enzymes one reaction away from generating or utilizing acetyl-CoA have a higher acetylation stoichiometry in the ΔCobBstrain during stationary phase, suggesting that CobB is involved in many facets of acetyl-CoA utilization.

2.4.4 Metabolite analysis reveals increased levels of acetyl-phosphate in the CobB mutant

NAD^+-dependent protein deacetylases (sirtuins) regulate many metabolic enzymes in prokaryotes and eukaryotes. (33) In prokaryotes, CobB regulates acetyl-CoA synthetase, which is required for cell growth on acetate. (28, 29) Activation of Acs by reversible acetylation is thought to be essential for "the acetate switch." (34) This transition occurs when bacterial cells have depleted carbon sources such as glucose and begin to utilize environmental acetate for growth. If flux through Acs is inhibited, for example by acetylation, then acetate can be diverted through the acetate kinase (AckA), phosphate acetyltransferase (Pta) pathway to generate acetyl-CoA (**Figure 2-8**). The higher stoichiometry as well as multiple acetylation sites on Pta suggests an inhibitory effect on protein function. Because acetate is rerouted through the AckA-Pta pathway due to Acs inhibition, we speculated that there would be a buildup of acetyl-phosphate if Pta function were inhibited by acetylation. To test this, we measured acetyl-phosphate in the WT and ΔCobBstrains. Indeed, there was a significant increase of acetyl-phosphate in the CobB mutant (**Figure 2-8B**).

This suggests that strains lacking CobB have an inability to efficiently utilize acetyl-phosphate in the ΔCobBstrain, consistent with inhibition of Pta activity.

Figure 2-7: Bioinformatic and network analysis of acetylation stoichiometry

(A) KEGG pathway analysis and (B) Gene ontology of biological process for acetylated proteins in both WT and ΔCobBusing an acetylation stoichiometry greater than 7% (average stoichiometry). Pathways and biological processes are significantly enriched ($p < 0.05$) above an unmodified, trypsin-digested, E. coli proteome (background proteome used for DAVID analysis). (C) Pie chart of acetylation sites detected per protein. Network analysis of (D) metabolic pathways and (E) transcriptional and translational biological processes. (F) Legend for D and E. Shapes represent the number of acetylation sites detected per protein. The size of the shape is proportional to the highest acetyl-peptide stoichiometry of a protein in either WT or ΔCobB. Proteins are colored according to pathway or biological process.

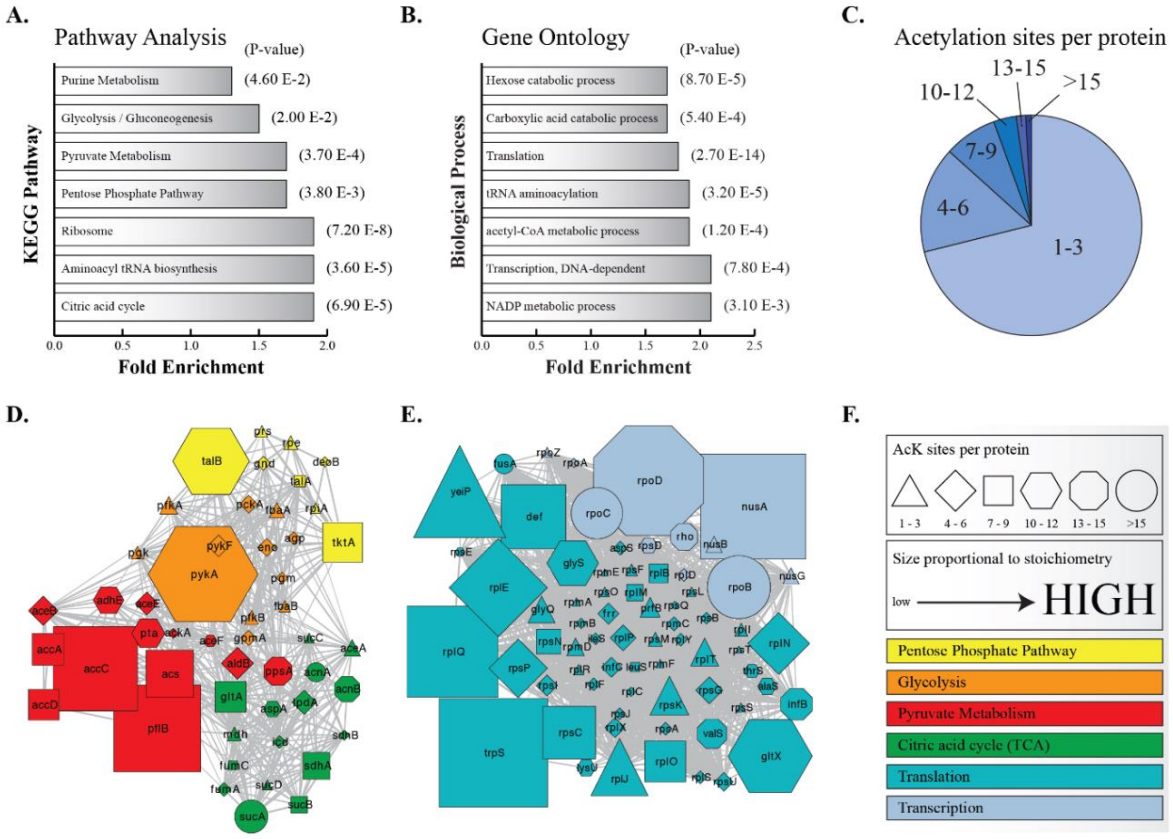
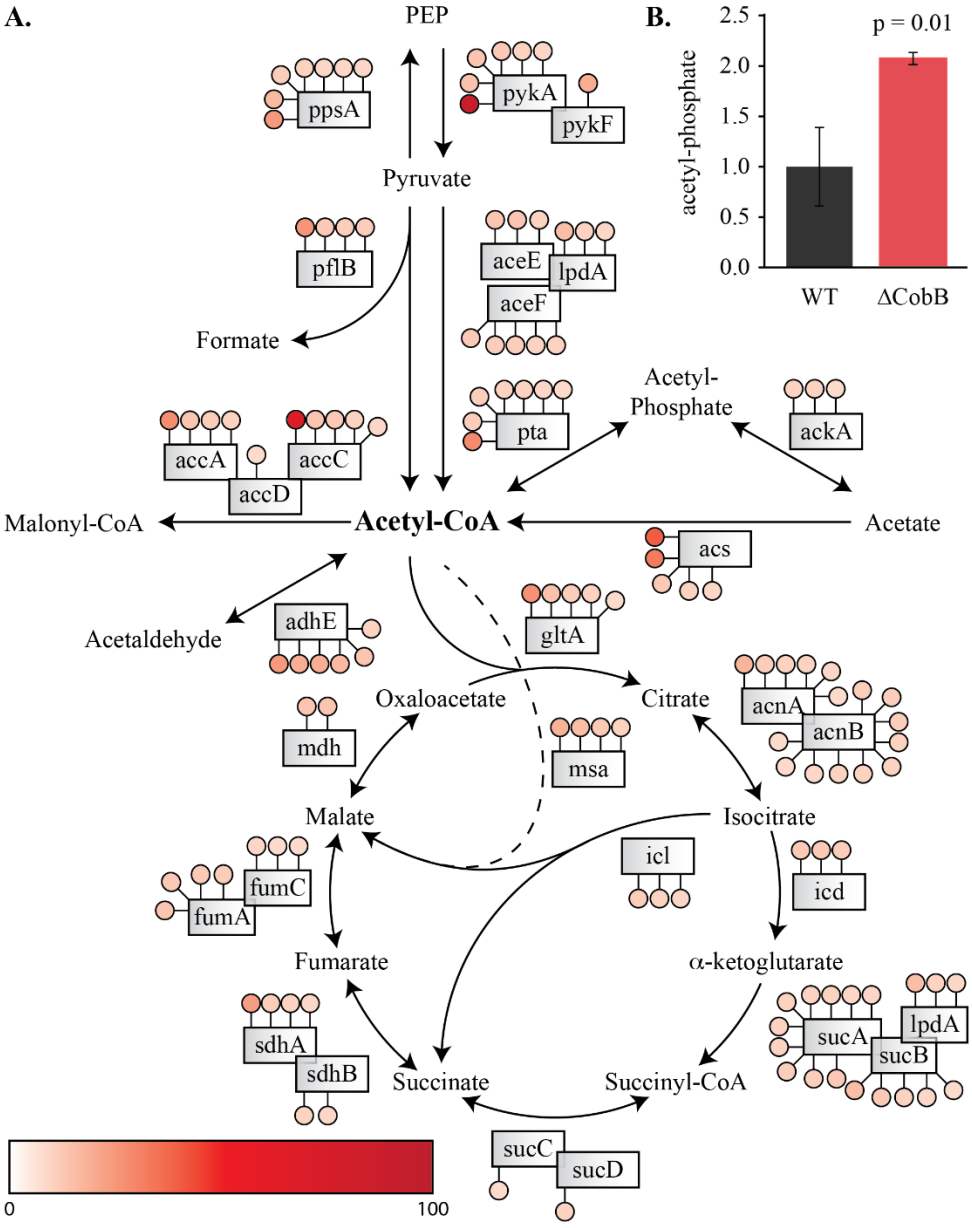

A.
Pathway Analysis
(P-value)
Purine Metabolism
(4.60 E-2)
Glycolysis / Gluconeogenesis
(2.00 E-2)
Pyruvate Metabolism
(3.70 E-4)
Pentose Phosphate Pathway
(3.80 E-3)
Ribosome
(7.20 E-8)
Aminoacyl tRNA biosynthesis
(3.60 E-5)
Citric acid cycle
(6.90 E-5)
KEGG Pathway
0.0 0.5 1.0 1.5 2.0
Fold Enrichment
B.
Gene Ontology
(P-value)
Hexose catabolic process
(8.70 E-5)
Carboxylic acid catabolic process
(5.40 E-4)
Translation
(2.70 E-14)
tRNA aminoacylation
(3.20 E-5)
acetyl-CoA metabolic process
(1.20 E-4)
Transcription, DNA-dependent
(7.80 E-4)
NADP metabolic process
(3.10 E-3)
Biological Process
0.0 0.5 1.0 1.5 2.0 2.5
Fold Enrichment
C.
Acetylation sites per protein
13-15
>15
10-12
7-9
4-6
1-3
D.
E.
F.
AcK sites per protein
1 - 3
4 - 6
7 - 9
10 - 12
13 - 15
>15
Size proportional to stoichiometry
low
HIGH
Pentose Phosphate Pathway
Glycolysis
Pyruvate Metabolism
Citric acid cycle (TCA)
Translation
Transcription

Figure 2-8: Acetylation of enzymes in metabolic pathways and metabolite analysis

(A) Enzymes of pyruvate metabolism and the TCA cycle are shown. Filled circles represent a peptide identified with an acetyl-stoichiometry greater than 7% (including single and double lysine containing peptides). The color of each circle represents the degree of acetylation according to heat map legend. Enzymes with the highest degree of acetylation include pykA, pta, accA, accC, acs, gltA, adhE, and msa. Enzymes with the highest number of AcK peptides include ppsA, pta, adhE, acnA, acnB, and sucA. (B) Acetyl-phosphate is significantly increased in the ΔCobB strain ($p = 0.01$).


A.
PEP
ppsA
pykA
pykF
Pyruvate
pflB
aceE
lpdA
aceF
Formate
Acetyl-
Phosphate
pta
ackA
accA
accC
accD
Malonyl-CoA
Acetyl-CoA
Acetate
acs
Acetaldehyde
adhE
gltA
Oxaloacetate
Citrate
acnA
acnB
mdh
msa
Malate
Isocitrate
icl
icd
fumC
fumA
Fumarate
α-ketoglutarate
sdhA
sdhB
Succinate
Succinyl-CoA
sucA
lpdA
sucB
sucC
sucD
0
100
B.
acetyl-phosphate
2.5
2.0
1.5
1.0
0.5
0.0
p = 0.01
WT
ΔCobB

2.5 Discussion

In this work, we developed a novel method to quantify site-specific acetylation stoichiometry across an entire proteome. This method utilizes isotopic acetic anhydride to chemically acetylate all unmodified lysines in a protein sample. Coupling this protein chemistry to high-resolution mass spectrometry and novel bioinformatic data processing, acetylation stoichiometry was quantified. Applying this method to bacterial cells, we describe the effects of the bacterial sirtuin member, CobB, on site-specific and global acetylation.

2.5.1 Site-specific acetylation stoichiometry

One role of the NAD^+-dependent sirtuins is the regulation of metabolism by deacetylation of key transcription factors and metabolic enzymes themselves. (6) CobB deacetylates acetyl-CoA synthetase at a conserved lysine residue, thereby activating Acs. (28, 29) Acs forms acetyl-CoA from acetate and ATP, thus allowing bacterial cells to utilize alternate carbon sources for energy when glucose is depleted. The active site lysine was not identified in this study, likely due to the small size of the peptide (<350 m/z) upon trypsin cleavage. We did however identify and quantify another acetyl site on Acs, K348, which displayed 32.5% acetylation in ΔCobBand 8.3% in WT. Further investigation is needed to determine if this site is critical for Acs activity as it lies at the dimeric interface.

Metabolic enzymes of the TCA cycle, pyruvate metabolism, and glycolysis were also highly enriched in acetylation, most notably in the number of AcK sites. Of these enriched metabolic pathways, enzymes one reaction away from acetyl-CoA displayed higher stoichiometry, including Pta, AccA, AccC, GltA, AdhE, and Msa (**Figure 2-8**). It is noteworthy that the enzymes, which utilize or generate acetyl-CoA, could also be regulated by reversible acetylation. This observation coupled with the known inhibition of Acs, led us to examine and demonstrate

increased levels of acetyl-phosphate in the CobB deficient strain, suggesting an inhibitory effect of acetylation on Pta activity.

Of the group 3 peptides (those showing substantially increased acetylation in the ΔCobB), RpsP, TrpS, Def, and Gpt had acetylation sites with the largest magnitude change in the CobB mutant. RpsP had a magnitude difference of 47% acetylation between the WT and ΔCobBstrains. As a component of the 30S ribosome, RpsP interacts with the RNA component of the ribosome. The ε-amino group of K46 appears to make a critical contact with the 2'-OH of a guanine nucleotide of the RNA. (35) Upon acetylation, this contact could destabilize the interaction between RpsP and RNA. TrpS generates tryptophanyl-tRNA that is used for protein synthesis and K312 has a magnitude difference of 64% acetylation between WT and ΔCobB. There is not a solved crystal structure for the E. coli form, however modeling with the Yersinia pestis variant shows that the analogous K site, K324, is 2.7 Å away from the backbone oxygen of P86, which likely forms H-bonding contact between the two residues. (36, 37) Def removes the formyl group from the N-terminal methionine and the peptide containing K158/161 has a magnitude change of 39%. These sites are on the surface of the protein and it is not known whether they are critical for protein-protein interaction. Gpt is involved in purine metabolism. The acetyl-site, K30, has a magnitude change of 27% between WT and ΔCobBstrains. The crystal structure reveals this site is also surface exposed and it is unknown if it plays a role in interactions. (38)

2.5.2 Global acetylation stoichiometry

The ability to measure acetyl-lysine stoichiometry allowed us to estimate the amount of total acetate attached to protein. The average global acetyl-stoichiometry was ~7% and we determined stoichiometry for ~30% of the quantifiable lysines in the E. coli proteome. Using these values and known, calculated values of E. coli for cell volume (BNID 100004), protein

concentration (BNID 108263), and the frequency of lysines present in E. coli (BNID 100636), (39) we estimated the amount of acetate on protein to be between 1.7 and 2.8 mM. This is striking since the concentration of acetyl-CoA is estimated to be less than 500 µM in E. coli. (40) This analysis reveals the high amount of acetate stored on proteins, though future experiments will be needed to assess the turnover rate with respect to acetyl-CoA and acetyl-phosphate consumption.

2.5.3 Stoichiometry vs. relative quantitation

Knowledge of stoichiometry is a key piece of information to accurately interpret the functional roles of acetylation. Current acetyl-proteomic methods rely on relative changes of peptides across conditions. In most cases, researchers have used an arbitrary 2-fold change cutoff as meaningful, (4, 10) which can greatly skew interpretation. For example, in our study, the acetyl-peptide from glutathione reductase (Gor) containing K105/K112 has a ΔCobB/WT fold change of 9.3. When the stoichiometry is considered, this peptide is 0.5% acetylated in WT and 4.3% in ΔCobB. This results in a magnitude change of 3.8% acetylation between the two conditions, which might only reflect biological variation. For the converse example, the acetyl-peptide from 3-phosphoshikimate 1-carboxyvinyltransferase (AroA) containing K354 has a ΔCobB/WT fold change of 0.6. However, the stoichiometry of this peptide is 38.8 and 69.5% for ΔCobBand WT, respectively (**Figure 2-6**). This is a magnitude difference of 30.7% acetylation and is likely more impactful on overall protein function than the former example.

2.5.4 Unique features of methodology

Major advantages/improvements of the presented method include: 1.) Quantitation is independent of protein level since stoichiometry depends on the total protein in a given sample. This eliminates the requirement to normalize to total protein amount. 2.) Stoichiometry as well as relative changes can be determined using this method. 3.) There is no need for immunoenrichment

of acetyl-peptides. This avoids any issues toward antibody enrichment bias introduced in the analysis. This issue is circumvented with the pairing algorithm of PVIEW by performing an in silico enrichment of acetyl-peptides. 4.) Low amount of starting material is used. Because of the minimal sample processing, less protein can be used, facilitating the analysis of limited, difficult to obtain samples. 5.) This method has broad applicability, from an isolated purified protein to cell culture or tissue sample.

2.6 References

1. Xiong, Y., and Guan, K.-L. (2012) Mechanistic insights into the regulation of metabolic enzymes by acetylation. J. Cell Biol. 198, 155–164
2. Yang, X.-J., and Seto, E. (2008) Lysine acetylation: codified crosstalk with other posttranslational modifications. Mol. Cell. 31, 449–461
3. Tanner, K. G., Trievel, R. C., Kuo, M. H., Howard, R. M., Berger, S. L., Allis, C. D., Marmorstein, R., and Denu, J. M. (1999) Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. J. Biol. Chem. 274, 18157–18160
4. Weinert, B. T., Iesmantavicius, V., Wagner, S. A., Schölz, C., Gummesson, B., Beli, P., Nyström, T., and Choudhary, C. (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in E. coli. Mol. Cell. 51, 265–272
5. 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
6. Feldman, J. L., Dittenhafer-Reed, K. E., and Denu, J. M. (2012) Sirtuin catalysis and regulation. J. Biol. Chem. 287, 42419–42427
7. 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
8. 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

9. Hebert, A. S., Dittenhafer-Reed, K. E., Yu, W., Bailey, D. J., Selen, E. S., Boersma, M. D., Carson, J. J., Tonelli, M., Balloon, A. J., Higbee, A. J., Westphall, M. S., Pagliarini, D. J., Prolla, T. A., Assadi-Porter, F., Roy, S., Denu, J. M., and Coon, J. J. (2013) Calorie restriction and SIRT3 trigger global reprogramming of the mitochondrial protein acetylome. Mol. Cell. 49, 186–199
10. Rardin, M. J., Newman, J. C., Held, J. M., Cusack, M. P., Sorensen, D. J., Li, B., Schilling, B., Mooney, S. D., Kahn, C. R., Verdin, E., and Gibson, B. W. (2013) Label-free quantitative proteomics of the lysine acetylome in mitochondria identifies substrates of SIRT3 in metabolic pathways. Proc. Natl. Acad. Sci. U. S. A. 110, 6601–6606
11. Lundby, A., Lage, K., Weinert, B. T., Bekker-Jensen, D. B., Secher, A., Skovgaard, T., Kelstrup, C. D., Dmytriyev, A., Choudhary, C., Lundby, C., and Olsen, J. V. (2012) Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. Cell Rep. 2, 419–431
12. Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol. Cell. Proteomics. 1, 376–386
13. Olsen, J. V., and Mann, M. (2013) Status of large-scale analysis of post-translational modifications by mass spectrometry. Mol. Cell. Proteomics. 12, 3444–3452
14. Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., Gnad, F., Cox, J., Jensen, T. S., Nigg, E. A., Brunak, S., and Mann, M. (2010) Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Sci. Signal. 3, ra3
15. Wu, R., Haas, W., Dephoure, N., Huttlin, E. L., Zhai, B., Sowa, M. E., and Gygi, S. P. (2011)

A large-scale method to measure absolute protein phosphorylation stoichiometries. Nat. Methods. 8, 677–683

16. Steen, H., Jebanathirajah, J. A., Springer, M., and Kirschner, M. W. (2005) Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS. Proc. Natl. Acad. Sci. U. S. A. 102, 3948–3953
17. Weinert, B. T., Iesmantavicius, V., Moustafa, T., Schölz, C., Wagner, S. A., Magnes, C., Zechner, R., and Choudhary, C. (2015) Acetylation dynamics and stoichiometry in Saccharomyces cerevisiae. Mol. Syst. Biol. 11, 833
18. Hallows, W. C., Lee, S., and Denu, J. M. (2006) Sirtuins deacetylate and activate mammalian acetyl-CoA synthetases. Proc. Natl. Acad. Sci. U. S. A. 103, 10230–10235
19. Riordan, J. F., and Vallee, B. L. (1972) [41] Acetylation. in Methods in Enzymology, pp. 494–499, Academic Press, Volume 25, 494–499
20. Khan, Z., Amini, S., Bloom, J. S., Ruse, C., Caudy, A. A., Kruglyak, L., Singh, M., Perlman, D. H., and Tavazoie, S. (2011) Accurate proteome-wide protein quantification from high-resolution 15N mass spectra. Genome Biol. 12, R122
21. Khan, Z., Bloom, J. S., Garcia, B. A., Singh, M., and Kruglyak, L. (2009) Protein quantification across hundreds of experimental conditions. Proc. Natl. Acad. Sci. U. S. A. 106, 15544–15548
22. 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
23. 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

24. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 41, D808–15
25. Snel, B., Lehmann, G., Bork, P., and Huynen, M. A. (2000) STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. Nucleic Acids Res. 28, 3442–3444
26. Saito, R., Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., Lotia, S., Pico, A. R., Bader, G. D., and Ideker, T. (2012) A travel guide to Cytoscape plugins. Nat. Methods. 9, 1069–1076
27. Lu, W., Clasquin, M. F., Melamud, E., Amador-Noguez, D., Caudy, A. A., and Rabinowitz, J. D. (2010) Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. Anal. Chem. 82, 3212–3221
28. Starai, V. J., Celic, I., Cole, R. N., Boeke, J. D., and Escalante-Semerena, J. C. (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science. 298, 2390–2392
29. Starai, V. J., Takahashi, H., Boeke, J. D., and Escalante-Semerena, J. C. (2003) Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in Salmonella enterica and Saccharomyces cerevisiae. Genetics. 163, 545–555
30. Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z.-B., Zeng, R., Xiong, Y., Guan, K.-L., Zhao, S., and Zhao, G.-P. (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. Science. 327, 1004–1007
31. Olsen, J. V., Ong, S.-E., and Mann, M. (2004) Trypsin cleaves exclusively C-terminal to

arginine and lysine residues. Mol. Cell. Proteomics. 3, 608–614

32. Beck, H. C., Nielsen, E. C., Matthiesen, R., Jensen, L. H., Sehested, M., Finn, P., Grauslund, M., Hansen, A. M., and Jensen, O. N. (2006) Quantitative proteomic analysis of post-translational modifications of human histones. Mol. Cell. Proteomics. 5, 1314–1325
33. Schwer, B., and Verdin, E. (2008) Conserved metabolic regulatory functions of sirtuins. Cell Metab. 7, 104–112
34. Wolfe, A. J. (2005) The acetate switch. Microbiol. Mol. Biol. Rev. 69, 12–50
35. Pulk, A., and Cate, J. H. D. (2013) Control of Ribosomal Subunit Rotation by Elongation Factor G. Science. 340, 1235970
36. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources. Nucleic Acids Res. 37, D387–92
37. Kopp, J., and Schwede, T. (2004) The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. Nucleic Acids Res. 32, D230–4
38. Vos, S., Parry, R. J., Burns, M. R., de Jersey, J., and Martin, J. L. (1998) Structures of free and complexed forms of Escherichia coli xanthine-guanine phosphoribosyltransferase. J. Mol. Biol. 282, 875–889
39. Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010) BioNumbers--the database of key numbers in molecular and cell biology. Nucleic Acids Res. 38, D750–3
40. Danchin, A., Dondon, L., and Daniel, J. (1984) Metabolic alterations mediated by 2-ketobutyrate in Escherichia coli K12. Mol. Gen. Genet. 193, 473–478

Chapter 3: Site-specific reactivity of non-enzymatic lysine acetylation

This chapter was published as a co-first author publication in ACS Chemical Biology with the following citation:

Baeza, J., Smallegan, M. J., & Denu, J. M. (2015). Site-specific reactivity of nonenzymatic lysine acetylation. ACS Chemical Biology, 10(1), 122–128.

Author contributions: J.B. and J.M.D. designed research study. J.B. performed experiments. J.B., M.J.S., and J.M.D analyzed results. M.J.S. performed bioinformatic analysis. J.B. and M.J.S. designed figures. All authors drafted and edited the manuscript. J.B. and M.J.S. are co-first authors. J.M.D. is corresponding author. We would like to thank K. Dittenhafer-Reed for purification of TFAM, J. Lee for purification of Histone H3, J. Feldman for the synthesis of Histone H3 and H4 peptide, V. Kuzniestsov for the assistance with the graphical abstract and Q. Doba for assistance in preparing the manuscript. 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 the National Science Foundation Graduate Research Fellowship Program (NSF-GRFP) DGE-1256259 (J.B.).

3.1 Abstract

Protein acetylation of lysine ε-amino groups is abundant in cells, particularly within mitochondria. The contribution of enzyme-catalyzed and non-enzymatic acetylation in mitochondria remains unresolved. Here, we utilize a newly developed approach to measure site-specific, non-enzymatic acetylation rates for 90 sites in eight native purified proteins. Lysine reactivity (as second-order rate constants) with acetyl-phosphate and acetyl-CoA ranged over three orders of magnitude, and higher chemical reactivity tracked with likelihood of dynamic modification in vivo, providing evidence that enzyme-catalyzed acylation might not be necessary to explain the prevalence of acetylation in mitochondria. Structural analysis revealed that many highly reactive sites exist within clusters of basic residues, whereas lysines that show low reactivity are engaged in strong attractive electrostatic interactions with acidic residues. Lysine clusters are predicted to be high-affinity substrates of mitochondrial deacetylase SIRT3 both in vitro and in vivo. Our analysis describing rate determination of lysine acetylation is directly applicable to investigate targeted and proteome-wide acetylation, whether or not the reaction is enzyme catalyzed.

3.2 Introduction

Protein acetylation is a post-translational modification affecting diverse cellular processes.(1–3) Although regulation of transcription by reversible acetylation of histone proteins is well-known, mass spectrometry based proteomic studies have catalogued thousands of acetylation sites in the mitochondria.(4–8) Unlike nuclear acetylation, which is catalyzed by several families of lysine acetyltransferases (KATs), direct evidence of acetyltransferases in mitochondria is lacking. However, mitochondria do harbor the bona fide deacetylase, SIRT3, a member of the NAD^+-dependent deacylases.(9, 10) SIRT3 has been shown to deacetylate and stimulate the activity of a number of metabolic enzymes, leading to enhanced oxidative metabolism, urea cycle, and ROS detoxification.(11, 12)

Lack of evidence for mitochondrial KATs raises the possibility that lysine acetylation is largely an uncatalyzed reaction, whereby the unprotonated lysine side-chain reacts with the thioester of acetyl-CoA.(13) Non-enzymatic acetylation has been observed on histone proteins.(14–16) Due to higher pH and increased acetyl-CoA levels in mitochondria, non-enzymatic acetylation appears to be a viable possibility.(13, 15, 17–19) Consistent with this idea, mitochondrial acetyl-CoA metabolism altered by dietary regimens or genetic manipulations affect lysine acetylation.(4, 5, 10, 13, 20, 21) Additionally, in vitro incubation of acetyl-CoA with non-histone proteins leads to a time dependent increase in acetylation.(17) Acetyl-phosphate, a prokaryotic metabolic intermediate can modify lysine residues in vivo and in vitro, further illustrating the general plausibility of non-enzymatic acetylation.(18, 19) We have previously demonstrated that KAT families, GCN5 and MYST, utilize general base catalysis to remove a proton from the ε-amino group, permitting nucleophilic attack on the bound acetyl-CoA.(15, 22) In this regard, the pK_a of the ε-amino group has little effect on the enzyme-catalyzed reaction,

while the uncatalyzed reaction is directly dependent on the amount of unprotonated lysine.(15) Regardless of whether enzymatic or nonenzymatic acetylation is the primary mode, knowledge of lysine reactivity toward these acetylating agents would provide crucial insight in the general mechanism of protein acetylation.

To establish an understanding of lysine reactivity for acetylation, the chemical kinetics at the site-specific level and among a broad range of proteins must be quantified. To date, there is only one report that quantifies the nonenzymatic second order rate for acetylation, which was reported for lys-36 of histone H3.(16) Here, we utilize a newly developed mass spectrometry method to investigate the chemical reactivities (second order rate constants) of 90 lysine sites within eight purified proteins from mitochondrial and nonmitochondrial sources. The analysis revealed acetylation reactivities ranging over three orders of magnitude. Structural and bioinformatic analyses explore the molecular basis for high and low reactivity sites on two mitochondrial proteins known to be regulated by acetylation.

3.3 Experimental Procedures

3.3.1 Protein samples

Purified bovine GDH and porcine αKGDH were purchased from Sigma Aldrich. Recombinant Mouse ACAT1 (amino acids 31-424) with C-terminal $(His)_6$ tag was purified according to previous methods.(23) Recombinant Histone H3.2 and H4 (Xenopus laevis) were expressed and purified according to previous methods.(24) Recombinant mouse TFAM pQE80-L (amino acids 43-246) was expressed by transforming plasmid into BL21 DE3 PlysS E. coli and grown in 2XYT media. Cells were grown until mid-log phase (OD_{600} 0.6) at 37°C with shaking and induced with 420 μM IPTG and grown overnight at 18°C. Cells were harvested and resuspended in 30 mM sodium phosphate, 100 mM NaCl, 1mM beta-mercaptoethanol, 10 mM imidazole, 1 mM PMSF, 10 μg/mL aprotinin and 10 μg/mL leupeptin. Cells were lysed by sonication and clarified by centrifugation at 16,000 RPM for 30 minutes. Protein was purified nickel chelating chromatography (Wash buffer: 30 mM sodium phosphate, 100 mM NaCl, 1 mM BME, 10 mM imidazole; Elution buffer: 30 mM sodium phosphate, 100 mM NaCl, 1 mM BME, 250 mM imidazole) by FPLC (AKTA, GE Healthcare Life Sciences ©) using a linear gradient. A 10% step using elution buffer was performed. Protein containing fractions were pooled, concentrated and dialyzed overnight. Protein concentration was determined using Bradford reagent (Bio-Rad).

3.3.2 Acetylation kinetics

3.3.2.1 Time-dependent acetylation

Purified protein samples were incubated at a 1 $mg \cdot mL^{-1}$ concentration in 100 mM ammonium bicarbonate buffer (pH 8.4) with acetyl-CoA or acetyl-phosphate and incubated at 37°C on Thermomixer C at 1000 RPM (Eppendorf) in a 120 uL total volume. Time points were

taken at 0, 5, 10, 30, 60 minutes and prepared for mass spectrometry analysis as described below. BSA was treated with 4.5 mM AcCoA and 8.5 mM AcP.

3.3.2.2 Concentration-dependent acetylation

Purfied protein samples were incubated at 1 $mg \cdot mL^{-1}$ concentration in 100 mM ammonium bicarbonate buffer (pH = 8.4) with varying concentration of AcCoA or AcP 0.5, 1, 2, 4, and 8 mM in 40 uL total volume. Reaction was incubated at 37°C on Thermomixer C at 1000 RPM for 60 minutes and prepared for mass spectrometry analysis as described below.

3.3.3 Sample preparation, chemical acetylation and digestion

Due to the nature of non-enzymatic reactions, quenching of the reaction requires the removal of the reactive chemical species. To do this, a 20uL aliquot of the reaction volume was transferred to 10 kDa MWCO filter unit (EMD Millipore) that was prefilled with 450 µL urea buffer (8M urea, 100 mM ammonium bicarbonate pH 7.4, 5 mM DTT), washed twice and transferred to a new 1.5mL microcentrifuge tube according to manufacturer's guidelines. This, in essence, lowered the concentration of both protein and acetylating reagent.

Each sample was heat denatured and reduced at 60°C on the Thermomixer C at 1000 RPM for 30 minutes followed by cysteine alkylation using 40 mM iodoacetamide for 30 minutes in the dark. The pH of the reaction volume was increased to ~8 by adding NH_4OH and protein acetylation was performed by adding ~50 µmol acetic anhydride-d6 (99.8% isotopic purity, Cambridge Isotopes) and incubating at 60°C for 30 minutes. Buffer exchange occurred with 50 mM ammonium bicarbonate (pH 8.4) using 10K MWCO filter units. For the reactions using acetyl-CoA, protein was digested with a 1:80 GluC-to-protein ratio for 4 hours at 37°C while shaking at 500 RPM on Thermomixer C. Sample was dried down using speedvac and resuspended using 50 mM ammonium bicarbonate (pH 8.4) and digested with trypsin at a 1:80 ratio. For the reactions

using acetyl-phosphate, protein was digested with GluC as above followed by chemical acetylation with acetic anhydride-d6 labeling N-terminus of the GluC peptides. GluC digested peptides were dried down and resuspended in 50 μL of 50 mM ammonium bicarbonate (pH 8.4) and digested with trypsin as above. Peptides were then acidified with 1% TFA, dried down and resuspended in 2% acetonitrile, 0.1% formic acid and analyzed by mass spectrometry as described below.

3.3.4 Determination of second order rate constant for purified peptide

The purified peptide, KQTARKSTGGKAPKWW, with N^{α}-acetylation was used to determine the rate constant of a primary amine. 1 μmol of purified peptide was incubated in 100 mM ammonium bicarbonate buffer (pH 8.5) with varying concentration of AcCoA (0.5, 1, 2, 4, 8 mM) in 50 uL total volume for 60 minutes at 37°C. To quench the reaction, 50 μmol of acetic anhydride-d6 was added to the sample and incubated at 37°C for 20 min. Sample was acidified with 1% TFA, desalted with OMIX C18 Tips (Agilent Technologies), resuspended in 2% acetonitrile, 0.1% formic acid and analyzed by mass spectrometry as described below.

3.3.5 Mass Spectrometry

3.3.5.1 LC-MS/MS

Peptides were separated using a Dionex Ultimate 3000 RSLCnano HPLC using a Waters Atlantis reverse phase column (100 μm x 150 mm). Mobile phase consisted of (A) 0.1% formic acid in HPLC grade H_2O and (B) 0.1% formic acid in HPLC grade acetonitrile. Peptides were eluted in a linear gradient of 2 – 40% B at 700 nL/min over 60 minutes with a column temperature at 60 °C and introduced into a Thermo Q-Exactive hybrid quadrupole Orbitrap mass spectrometer by nanoelectrospray ionization. Survey scan was performed in positive ion mode with a 140,000 resolution, AGC of 1E6, max fill time 250 ms, and a scan range of 350 to 2000 m/z. Data dependent MS/MS was performed with a resolution of 17,500, AGC of 1E5, max fill time 100 ms, isolation

window of 1.5 m/z, and a loop count of 10. The source voltage was set at 2000 V and capillary temperature at 250 °C.

3.3.5.2 Database search and data analysis

MS data was searched using high-resolution database search algorithm, Morpheus. Precursor mass tolerance was set to 10 ppm and product mass tolerance to 15 ppm with a false discovery rate of 1%. Up to 2 missed cleavages were allowed for a tandem GluC and Trypsin digest. Carbamidomethylation of cysteines was set as a fixed modification while methionine oxidation and lysine acetylation-d3 (unimod accession #: 56) was set as a variable modification. For Histone H3 and H4 only, we used the database search engine, Mascot, for peptide identification using the same settings as above, except Trypsin was used as the protease.

Extracted ion chromatograms (XICs) for the mono, di, tri acetylated peptides were obtained using Thermo Xcalibur Qual Browser. By using the appropriate mass shift, the light acetyl-peptide was also identified, matched and quantified. (z = 2, 1.509 m/z; z = 3, 1.006 m/z) For peptides with more than one lysine, mixed isotopic heavy and light acetylated peptides are generated. These mixed isotopic peptides are identified and quantified using mass shifts corresponding to the number of lysines present in the peptide and used for quantitation of stoichiometry, which is determined by the following equations:

1 lysine containing peptide

$$\frac{\mathrm{XIC_L}}{\mathrm{XIC_L} + \mathrm{XIC_H}}$$

2 lysine containing peptides

$$\frac{\mathrm{XIC_{LL}} + \frac{1}{2}\mathrm{XIC_{LH}}}{\mathrm{XIC_{LL}} + \mathrm{XIC_{LH}} + \mathrm{XIC_{HH}}}$$

3 lysine containing peptides

$$\frac{\mathrm{XIC_{LL}} + \frac{2}{3}\mathrm{XIC_{LLH}} + \frac{1}{2}\mathrm{XIC_{LH}}}{\mathrm{XIC_{LLL}} + \mathrm{XIC_{LLH}} + \mathrm{XIC_{LHH}} + \mathrm{XIC_{HHH}}}$$

where XIC_L is the extracted ion chromatogram of the light acetyl-peptide, XIC_H is the extracted ion chromatogram of the heavy acetyl-peptide and XIC_{LH}, when present, is the extracted ion chromatogram for the double lysine containing peptide with a light and heavy acetyl-lysine

3.3.6 Bioinformatics

3.3.6.1 Acetylation proteomics data aggregation

Proteomics datasets were retrieved from supplemental data tables accompanying.(4, 5, 23, 25) Defunct International Protein Index identifiers from Sol et al. 2012 were mapped onto Uniprot Accession numbers using DAVID's gene ID converter utility.(26, 27) Peptides containing multiple lysines were split into one entry per lysine site and each was assigned the fold-change value for the whole peptide. Lysine site data from all studies was then concatenated into one data table. Conditions comparing $Sirt3^{-/-}$ mice to wild type mice were collapsed onto one data column by averaging values from the $Sirt3^{-/-}$/WT experiment in Mouse embryonic fibroblast cells (Sol 2012), $Sirt3^{-/-}$/WT experiment in mouse liver mitochondria (Rardin 2013), and the $Sirt3^{-/-}$ Control Diet/WT in mouse liver mitochondria from Hebert 2013. We subsequently merged this data with our observed lysine sites by Uniprot accession and lysine site number. If multiple data points for each study matched with one of our observed lysine sites, the fold-change value associated with the smallest peptide was taken as a representative fold-change.

3.3.6.2 In silico pK_a prediction

Prediction of the pK_a of lysine sites of ACAT1 (Model retrieved from SWISS-MODEL Repository for UniProt: Q8QZT1 based on template pdb:2IB8)(28) and GDH (pdb:3MW9) using

the H++ Version 3.1 server.(29–31) Computations were performed using the following initial parameters: external dielectric of 80, internal dielectric of 10, and salinity of 0.15 M. pKs from chain A were taken as the representative value for each lysine.

3.3.6.3 Solvent accessible surface area calculation

Solvent accessible surface area per residue was calculated using POPS* Version 1.6.2a.(32, 33) A probe radius of 1.4 Å was used and the percent of lysine residue solvent accessible from lysines on chain A were plotted against rates.

3.3.6.4 Average B-factor calculation

Temperature factors were extracted from the GDH structure (pdb:3MW9) and for each lysine, the B-factor values corresponding to the constituent atoms were averaged to obtain a per residue average B-factor using PyMol.(34) Values from chain A were used as representative for each lysine site.

3.3.6.5 Three-dimensional lysine motif

Side chain interaction centers for ACAT1 (Model retrieved from SWISS-MODEL Repository for UniProt: Q8QZT1 based on template pdb:2IB8)(28) and GDH (pdb:3MW9) were calculated using PyMol in accordance with Bahar and Jernigan (1996).(34, 35) For each lysine in these two proteins, the distance between the lysine of interest and all side chain interaction centers within 7 Å was calculated.

3.4 Results and Discussion

3.4.1 Acetylation increases in a time- and concentration-dependent manner

To quantify chemical acetylation kinetics, we modified a method previously developed for determining acetylation stoichiometry (**Figure 3-1**) (36). In short, site-specific stoichiometry is determined by reacting native proteins with acetylating reagent acetyl-CoA (AcCoA) or acetyl-phosphate (AcP), followed by denaturation and reaction with heavy-labeled acetic anhydride to chemically acetylate all remaining unmodified lysines. Proteolyzed samples were subjected to high-resolution mass spectrometry and acetylation stoichiometry was quantified. Here, purified protein was incubated with acetylating reagent AcCoA or AcP for various times or at varied reagent concentrations, followed by fast buffer exchange to remove any activated acetylating reagent. Sample processing for quantifying stoichiometry involved denaturation, reduction, alkylation, and isotopic chemical acetylation of the protein sample. Enzymatic digestion using GluC and trypsin generated chemically identical heavy and light peptide pairs, which were resolved by nano-LC-MS/MS, with subsequent quantification of acetylation stoichiometry in a site-specific manner.

To establish progress curves of non-enzymatic acetylation, we first evaluated the time-dependent acetylation of bovine serum albumin (BSA) with AcCoA and AcP. With both acetylating reagents, the site-specific acetylation was linear over 60 minutes (**Figure 3-2**) and displayed vastly different rates for five unique lysine sites, highlighting the robustness of the method over a broad range of chemical reactivities. Once conditions for linearity were established, we determined the second order rate constants of site-specific lysine acetylation for the mitochondrial α-ketoglutarate dehydrogenase complex (αKGDH E1, E2, E3) and glutamate dehydrogenase using both AcCoA and AcP as the acetylating reagent, while acetyl-CoA

acetyltransferase 1 (ACAT1), mitochondrial transcription factor A (TFAM), Histone H3.2, and Histone H4 were treated using AcCoA alone. Rate constants were determined by varying [AcCoA] or [AcP], plotting the pseudo-first-order rate constant (k_{obs}) of site-specific acetylation as a function of concentration, and performing linear regression analysis. The slope of the line designates the second order rate constant. (**Figure 3-3**) shows an example of the rate data for three sites in GDH that were quantified for both AcCoA and AcP reactions.

This kinetic analysis evaluated 90 lysine sites across these eight proteins, including 27 lysine sites that were detected but no significant reactivity was calculable (**Figure 3-4**). Among the 63 sites with quantifiable reactivity, rate constants ranged over three orders of magnitude, with lys-503 (K503) on GDH yielding the second highest reactivity with a second order rate constant of 758 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ (**Figure 3-3**). The least reactive lysine on GDH (K415) displayed a second order rate constant of 7.41 x 10^{-5} $M^{-1}{\cdot}s^{-1}$, representing a 100-fold difference among sites on the same protein. In addition to highlighting the number of sites quantified per protein, dot plot analysis illustrates several important points: AcCoA and AcP show reactivities that are generally similar, and for each mitochondrial protein, the range of lysine reactivities within a given protein was between one and two orders of magnitude (**Figure 3-5**).

Figure 3-1 Kinetic analysis of nonenzymatic lysine acetylation.

Diagram of methodology used to determine non-enzymatic acetylation. Purified mitochondrial and non-mitochondrial proteins were incubated with varied concentrations of acetyl-CoA or acetyl-phosphate. At specific time points, protein is processed for mass spectrometry analysis followed by kinetic and bioinformatic analysis.

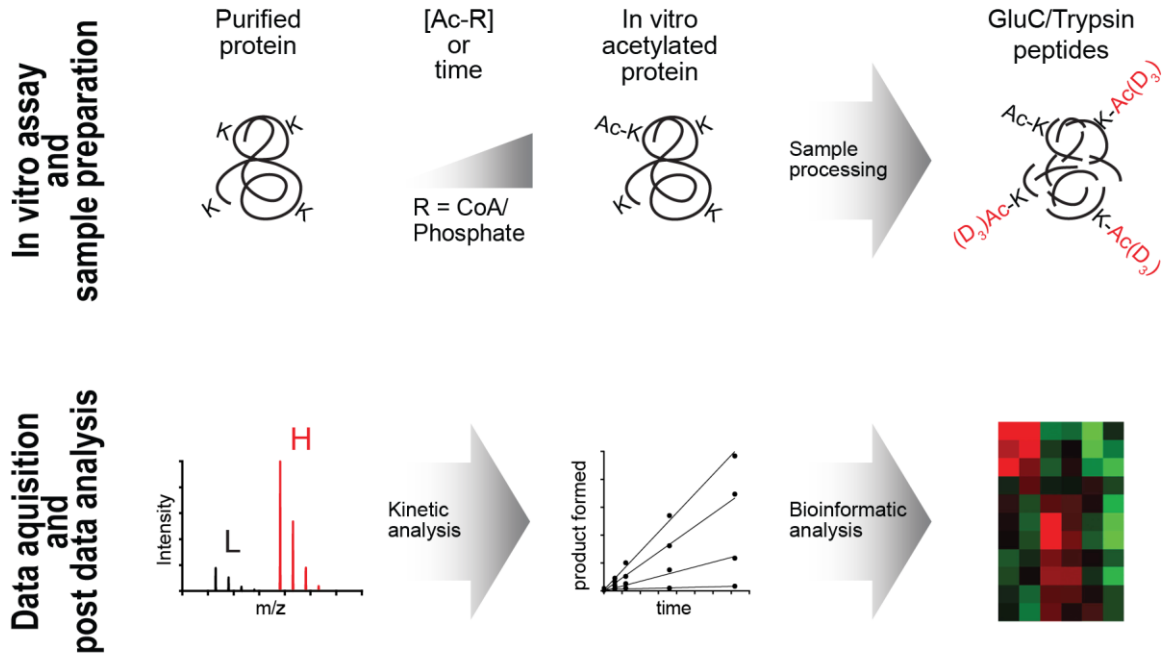


Figure 3-2 Nonenzymatic acetylation is linear with acetyl-CoA and acetyl-phosphate.

Incubation of BSA with 4.5mM acetyl-CoA and 8.5mM acetyl-phosphate over the course of 60 minutes shows linear increase in site-specific lysine acetylation.

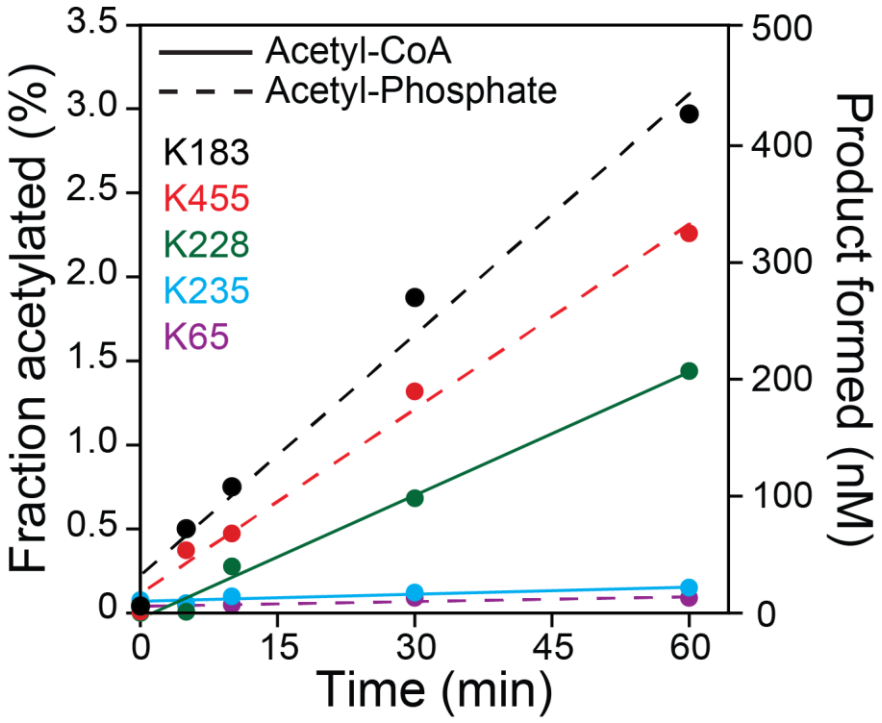


Figure 3-3 Acetyl-CoA and acetyl-phosphate display unique second order rate constants

Concentration-dependent acetylation of glutamate dehydrogenase (GDH). Purified GDH was incubated with acetyl-CoA and acetyl-phosphate for 60 minutes and analyzed to determine rate constants. The rate constants for the GDH sites are: K503 – 412 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ (AcCoA), and 758 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ (AcP), K110 – 138 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ (AcCoA), 37.5 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ (AcP), K415 – 36.4 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ (AcCoA) and 7.4 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ (AcP).

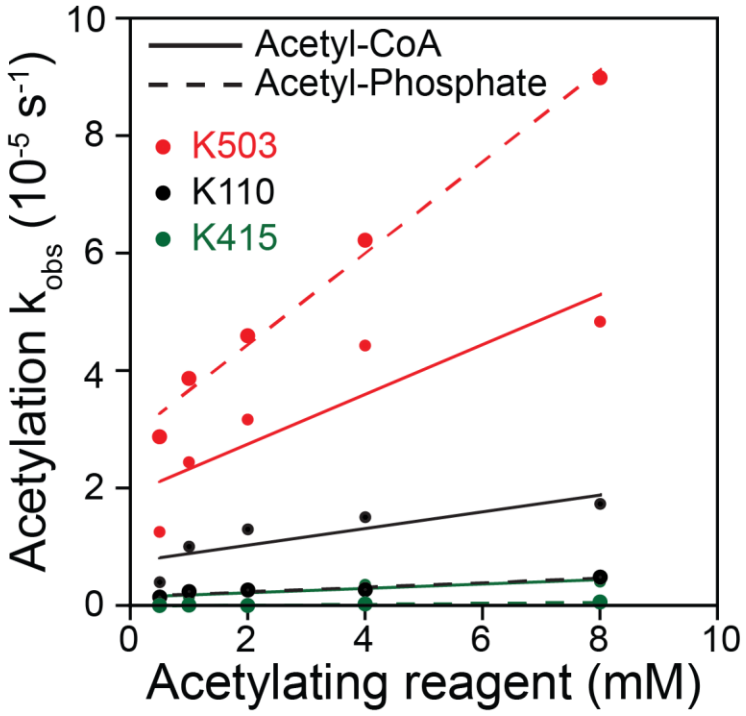


Figure 3-4 Site-specific lysine acetylation rate constants.

(A) Second order rates, standard error, and Pearson correlation coefficient are shown for each measured peptide. (B) Quantified non-reactive peptides are listed for each substrate condition.

A. Peptide Reactivity

		Lysine Site	Rate Constant $10^{-5}M^{-1}\cdot s^{-1}$	Std. Error	R^2
ACAT1	AcCoA	K304	978.2	23.8	0.998
		K332,335,340	397.6	11.9	0.997
		K263,265,270	164.2	56.3	0.739
		K400	120.4	8.1	0.987
		K240, 242, 248	116.6	10.0	0.978
		K260,263,265	106.3	19.9	0.905
		K171	57.6	5.2	0.976
		K227	55.3	0.4	0.999
		K178,187	55.0	15.7	0.803
		K260,263	32.5	9.6	0.792
		K121,128	31.0	10.6	0.740
		K63	24.2	2.7	0.964
		K370	10.5	1.9	0.787
		K220	7.7	2.2	0.807
GDH	AcCoA	K503	412.6	124.3	0.786
		K477, 480	205.8	23.8	0.961
		K110	137.7	51.9	0.701
		K346	89.8	11.6	0.952
		K352,363,365	75.3	10.2	0.948
		K211,212	61.3	21.4	0.732
		K545	54.2	6.8	0.955
		K68	43.0	11.1	0.834
		K415	36.4	9.6	0.826
	AcP	K503	757.9	49.3	0.987
		K527	46.3	20.8	0.622
		K110	37.5	6.9	0.907
		K415	7.4	0.3	0.996
		K147	2.5	0.2	0.976

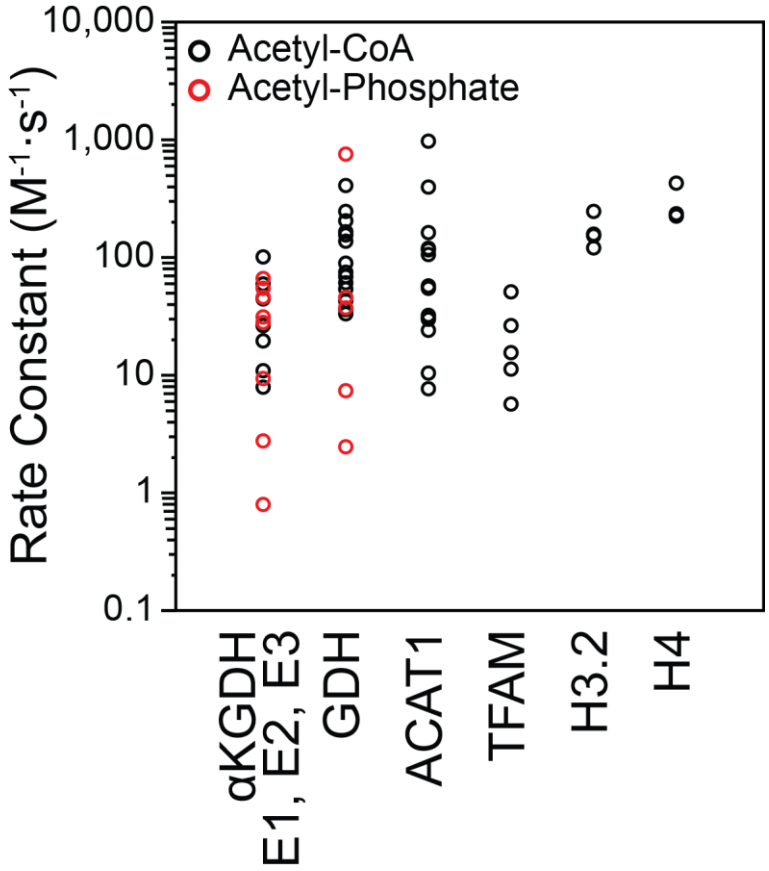
		Lysine Site	Rate Constant $10^{-5}M^{-1}\cdot s^{-1}$	Std. Error	R^2
H3.2	AcCoA	K123	248.0	77.7	0.836
		K80	158.6	64.6	0.667
		K57	154.0	59.5	0.770
		K10,15	120.8	46.3	0.773
H4	AcCoA	K21,32	430.4	107.1	0.890
		K80,92	237.2	69.5	0.853
		K32	226.7	115.9	0.657
TFAM	AcCoA	K52	51.4	1.6	0.997
		K190	26.6	1.9	0.985
		K111	15.7	2.4	0.934
		K69,76	11.3	0.3	0.998
		K69	5.7	1.1	0.906
αKGDH	AcCoA E1	K981	60.0	3.7	0.989
		K813	26.5	1.2	0.994
		K122	8.0	2.0	0.847
		K907	0.8	0.2	0.826
	AcCoA E2	K504	11.0	2.1	0.900
	AcCoA E3	K68,75	101.8	17.5	0.918
		K47	19.7	2.8	0.943
	AcP E1	K981	66.8	4.7	0.980
		K336	46.2	22.2	0.519
		K907	44.9	5.1	0.962
		K74	31.5	7.8	0.801
		K640	9.5	3.8	0.604
	AcP E2	K259	55.0	17.1	0.776
	AcP E3	K47	28.2	3.6	0.940
		K219	2.8	0.9	0.702
Peptide	AcCoA		261.9	6.8	0.832

B. Quantified Non-Reactive Lysines

Protein	Substrate	Lysine Site
αKGDH	AcCoA	K184 / K488 / K528 / K899 / K1020
E1	AcP	K122 / K184 / K348 / K488 / K499 K528 / K768 / K947 / K1020
E2	AcCoA	K177 / K204
	AcP	K13 / K177 / K215 / K505
E3	AcCoA	K134 / K309,308
	AcP	K92 / K134
ACAT1	AcCoA	K80,84 / K362
GDH	AcCoA	K90 / K365
TFAM	AcCoA	K154,156

Figure 3-5 Acetylation reactivities range over three orders of magnitude.

Dot plot of rate constants quantified for α-ketoglutarate dehydrogenase complex (E1, E2, E3), glutamate dehydrogenase (GDH), acetyl-coa acetyltransferase 1 (ACAT1), mitochondrial transcription factor A (TFAM), Histone H3.2 (H3) and Histone H4 (H4).



3.4.2 Nonenzymatic acetylation rates track with in vivo acetylation dynamics

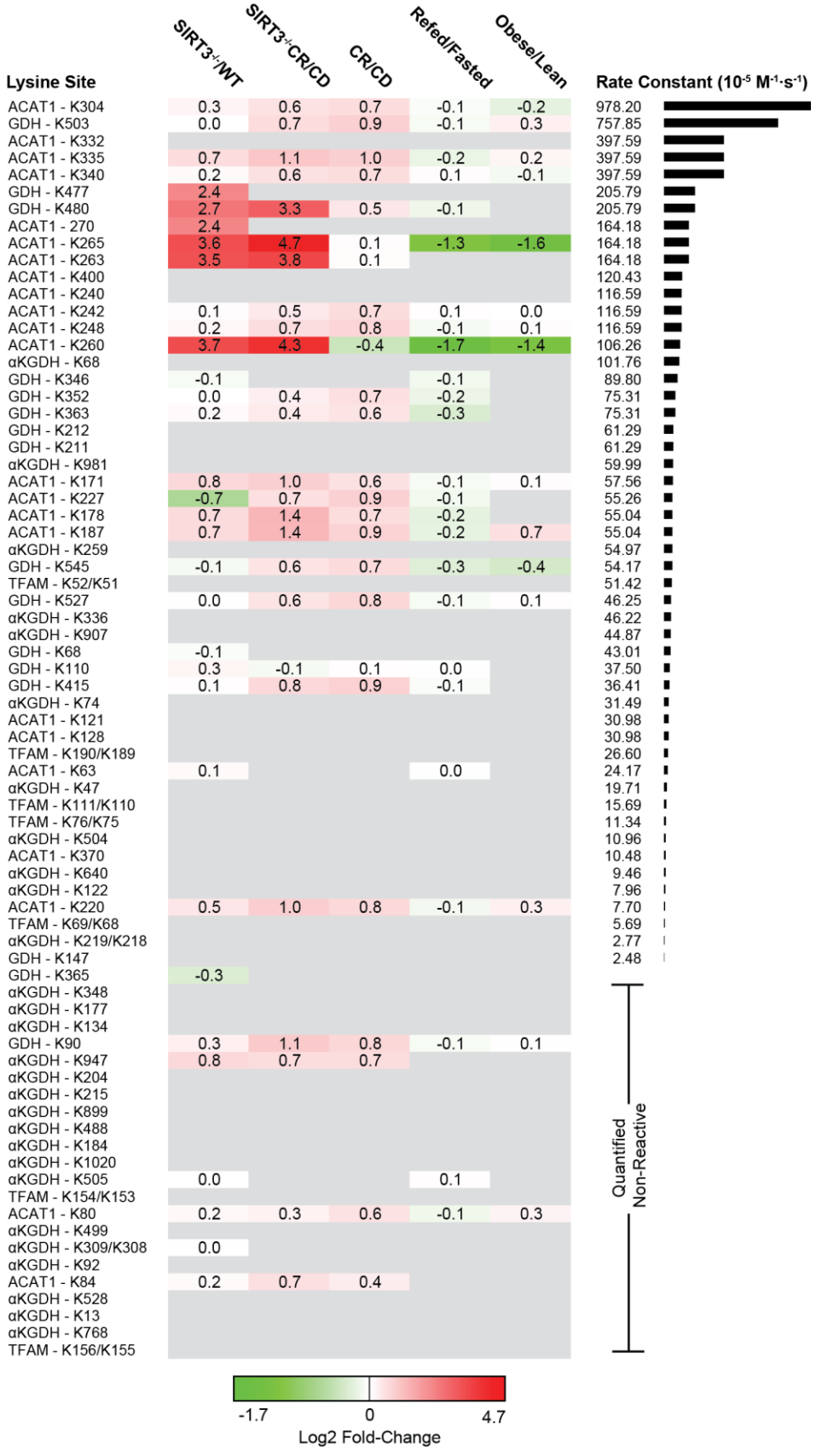
To query the biological relevance of differential lysine acetylation rates, we undertook a meta-analysis of previously published acetylation fold-change datasets generated by immunoenrichment of digested acetyl-peptides. Data from four previous proteome scale studies of mouse mitochondria and embryonic fibroblast cells were gathered, cleaned, and merged into one data table containing over 26,500 data points from 3,836 proteins encompassing five unique physiological perturbations.(4, 5, 7, 8) The datasets reflect a large swath of biological conditions known to interact with mitochondrial acetylation. Assembled data were then merged with the rate constants determined here, and ranked according to decreasing values of reactivities (**Figure 3-6**).

The number of conditions for which we were able to find fold-change data appears to decrease with decreasing rate constants. The immunoenrichment methodology used in these prior in vivo studies does not detect fold-change information for unmodified lysine sites or for acetylated peptides not enriched because of antibody specificity. Missing fold-change data likely reflects low or no in vivo acetylation of these low reactive sites. Here, our method of quantifying site reactivity does not rely on antibodies and therefore permits determination of both poorly and highly reactive lysine residues. In the group of 44 lysine residues with either no detectable reactivity or rates < 30 x 10^{-5} $M^{-1}{\cdot}s^{-1}$, only 9 (or 12%) of these were detected in at least one of the biological data sets. In contrast, 40 sites have second order rate constants > 30 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ with 24 (or 29%) of those sites overlapping with the biological data sets, and importantly 19/24 (or 79%) of those were observed in four different conditions (**Figure 3-6**). For comparison, we determined the second order rate constant of an unstructured histone H3 peptide with acetyl-CoA, yielding a value of 261 x 10^{-5} $M^{-1}{\cdot}s^{-1}$ (**Figure 3-4**). Lysine sites with the highest reactivity (second order rate constant) were found on ACAT1 and GDH, and many of these sites appear to dynamically change in

biological conditions that compare SIRT3$^{-/-}$ to WT and caloric restriction (CR) to a control diet (CD) (**Figure 3-6**). Together, these results suggest that many highly reactive sites are more likely to exhibit larger fold-changes between conditions and more likely to be targets of SIRT3.

Figure 3-6 Lysine reactivities map to dynamic acetylation sites.

Heatmap of observed sites ranked from high to low reactivity across five experimental conditions. CD: control diet. CR: calorie restriction. Acetylation fold-change values were compiled from previous studies of mouse liver mitochondria and MEF cells. Mouse protein sites that differ in sequence number from observed proteins are identified by both sites: $K[site]_{obs}/K[site]_{mus\ musculus}$. Quantified non-reactive sites are randomly ordered. Sites with high reactivities are more likely to be found in the acetylation datasets.


Lysine Site
SIRT3-/-/WT
SIRT3-/-CR/CD
CR/CD
Refed/Fasted
Obese/Lean
Rate Constant (10-5 M-1·s-1)
ACAT1 - K304 0.3 0.6 0.7 -0.1 -0.2 978.20
GDH - K503 0.0 0.7 0.9 -0.1 0.3 757.85
ACAT1 - K332 397.59
ACAT1 - K335 0.7 1.1 1.0 -0.2 0.2 397.59
ACAT1 - K340 0.2 0.6 0.7 0.1 -0.1 397.59
GDH - K477 2.4 205.79
GDH - K480 2.7 3.3 0.5 -0.1 205.79
ACAT1 - 270 2.4 164.18
ACAT1 - K265 3.6 4.7 0.1 -1.3 -1.6 164.18
ACAT1 - K263 3.5 3.8 0.1 164.18
ACAT1 - K400 120.43
ACAT1 - K240 116.59
ACAT1 - K242 0.1 0.5 0.7 0.1 0.0 116.59
ACAT1 - K248 0.2 0.7 0.8 -0.1 0.1 116.59
ACAT1 - K260 3.7 4.3 -0.4 -1.7 -1.4 106.26
αKGDH - K68 101.76
GDH - K346 -0.1 -0.1 89.80
GDH - K352 0.0 0.4 0.7 -0.2 75.31
GDH - K363 0.2 0.4 0.6 -0.3 75.31
GDH - K212 61.29
GDH - K211 61.29
αKGDH - K981 59.99
ACAT1 - K171 0.8 1.0 0.6 -0.1 0.1 57.56
ACAT1 - K227 -0.7 0.7 0.9 -0.1 55.26
ACAT1 - K178 0.7 1.4 0.7 -0.2 55.04
ACAT1 - K187 0.7 1.4 0.9 -0.2 0.7 55.04
αKGDH - K259 54.97
GDH - K545 -0.1 0.6 0.7 -0.3 -0.4 54.17
TFAM - K52/K51 51.42
GDH - K527 0.0 0.6 0.8 -0.1 0.1 46.25
αKGDH - K336 46.22
αKGDH - K907 44.87
GDH - K68 -0.1 43.01
GDH - K110 0.3 -0.1 0.1 0.0 37.50
GDH - K415 0.1 0.8 0.9 -0.1 36.41
αKGDH - K74 31.49
ACAT1 - K121 30.98
ACAT1 - K128 30.98
TFAM - K190/K189 26.60
ACAT1 - K63 0.1 0.0 24.17
αKGDH - K47 19.71
TFAM - K111/K110 15.69
TFAM - K76/K75 11.34
αKGDH - K504 10.96
ACAT1 - K370 10.48
αKGDH - K640 9.46
αKGDH - K122 7.96
ACAT1 - K220 0.5 1.0 0.8 -0.1 0.3 7.70
TFAM - K69/K68 5.69
αKGDH - K219/K218 2.77
GDH - K147 2.48
GDH - K365 -0.3
αKGDH - K348
αKGDH - K177
αKGDH - K134
GDH - K90 0.3 1.1 0.8 -0.1 0.1
αKGDH - K947 0.8 0.7 0.7
αKGDH - K204
αKGDH - K215
αKGDH - K899
αKGDH - K488
αKGDH - K184
αKGDH - K1020
αKGDH - K505 0.0 0.1
TFAM - K154/K153
ACAT1 - K80 0.2 0.3 0.6 -0.1 0.3
αKGDH - K499
αKGDH - K309/K308 0.0
αKGDH - K92
ACAT1 - K84 0.2 0.7 0.4
αKGDH - K528
αKGDH - K13
αKGDH - K768
TFAM - K156/K155
Quantified
Non-Reactive
-1.7
0
4.7
Log2 Fold-Change

3.4.3 Structural analysis of lysine reactivities provides insight for high and low rates

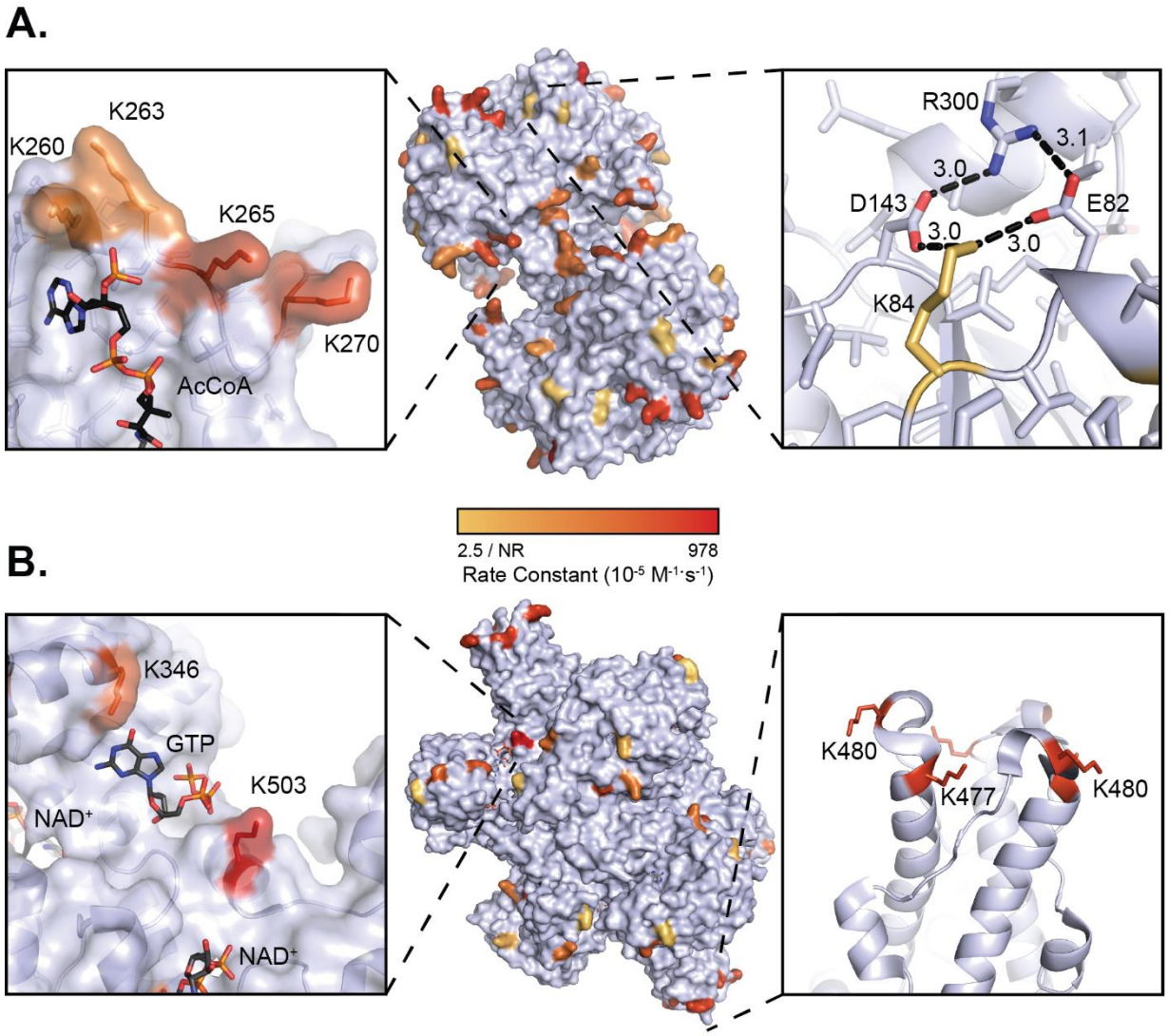
The extensive cataloguing of lysine reactivity allowed us to map these sites onto the protein structures of GDH and ACAT1, displaying a reactivity range that spans over two orders of magnitude. By visual inspection, lysine sites with the highest reactivity (red) tend to protrude away from the surface of the protein, while low reactivity sites (yellow) tend to form electrostatic interactions with neighboring residues (**Figure 3-4 and 3-7**). In ACAT1, this point is illustrated by structurally comparing K84, which yielded no significant reactivity, with K260, K263, K265 and K270, which displayed rate constants ranging from $106 \times 10^{-5} – 164 \times 10^{-5}$ $M^{-1} \cdot s^{-1}$. K84 is part of a network of electrostatic interactions involving aspartate, glutamate, and arginine residues (**Figure 3-7A**) (right panel). As a group, K260, K263, K265 and K270 form a cluster and do not make significant interactions with the protein surface (**Figure 3-7A**) (left panel). Quite remarkably, the acetylation state of K260, K263, K265 and K270 increased ≥10-fold in the SIRT3$^{-/-}$ mice compared with WT (**Figure 3-6**).(4) Equally interesting is the observation that K260 and K265 acetylation decreases when comparing refed/fasted as well as obese/lean (**Figure 3-6**).(8) These residues are located in the CoA binding pocket, within 3 – 5 Å from the ribosyl-phosphate group of CoA. Site-specific acetyl-lysine incorporation and in vitro biochemical analysis provided direct evidence that SIRT3-mediated deacetylation of K260ac and K265ac enhanced ACAT1 activity, likely due to decreased affinity for coenzyme A (CoA) through lost electrostatic interaction between positively-charged lysine and negatively charged 3'-phosphate of CoA.(8) Thus, the high intrinsic reactivity towards acetyl-CoA, described here, can identify functionally relevant acetylation sites, particularly those regulated by SIRT3.

We performed a similar structural analysis of lysine residues from GDH, which exists as a homohexamer featuring stacked dimers of trimers. K503 yielded the highest reactivity with both

acetyl-phosphate and acetyl-CoA, and interestingly, this site is known to be acetylated, succinylated and malonylated in mice.(5, 37) Sourced from bovine liver tissue, we found K503 of GDH was 10.1% acetylated, suggesting that K503 is a highly reactive site both in vivo and in vitro. K503 sits in the allosteric GTP-binding pocket and at the base of the antennae region of GDH (**Figure 3-7B**). The cleft containing K503 has three arginine residues involved in binding GTP. The charge of this cleft might attract negatively-charged acetyl-phosphate and acetyl-CoA. We observed no obvious substrate saturation, suggesting a formal binding does not exist. SIRT5 is reported to function as a mitochondrial de-succinyl and de-malonylase, and would likely remove such acyl groups from K503 of GDH.(37) Succinylated/malonylated K503 would be predicted to prevent binding of the allosteric inhibitor GTP. Interestingly, acetyl-proteomic data for SIRT3 suggests that this sirtuin does not regulate K503 of GDH (**Figure 3-6**). Instead, SIRT3 regulates the acetylation status on the tips of the two antennae, each of which involves the cluster of six lysine residues, K477 and K480 from three monomers (**Figure 3-7B**). K477 and K480 display 5 to 10-fold increase in acetylation in liver mitochondria from $SIRT3^{-/-}$ mice and are also two of the most reactive lysines uncovered in this study (**Figure 3-6**). The 6 lysines of each antenna form a positive cluster that do not engage in electrostatic interaction with other amino acids. The antenna region of GDH is known to undergo conformational changes during the catalytic cycle and functions as the conduit for inter-subunit communication during allosteric regulation.(38, 39) The antennae region projects out from the top of each NAD^+ binding domain, as well as intersecting near the GTP allosteric site (**Figure 3-7B**). Given that the antennae function to transmit catalytic and allosteric information between subunits,(38, 39) acetylation of K477 and K480 might alter the allosteric behavior of GDH. Further studies will be needed to directly investigate the functional role of K477, K480 and K503 acylation.

Figure 3-7 Visualization of lysine reactivity.

(A) Lysine reactivity mapped onto mouse ACAT1 structure (modeled from 2IB8, 87% identity) (center panel). Reactivities of K260, K263, K265, K270 are shown in the acetyl-CoA binding pocket (left panel). Non-reactive K84 shown forming a salt bridge with E82 and D143 (right panel). (B) Lysine reactivity mapped onto bovine glutamate dehydrogenase (pdb: 3MW9) (center panel). Reactivity of K503 shown near the allosteric GTP binding site (left panel). The trimeric antennae of GDH showing K477 and K480 reactivities and their close proximity. Acetylation rate color scale is in Log10 space.



3.4.4 Intrinsic properties do not predict lysine reactivity

To systematically evaluate the structural and chemical features within ACAT1 and GDH that affect lysine reactivity, we assessed the second order rate constants as a function of 1.) predicted pK_a, 2.) B-factor, 3.) surface accessibility and 4.) 3-D motifs (**Figures 3-8 and 3-9**). B-factor and predicted pK_a displayed no significant trend as a function of the determined rate constants. There was a more obvious trend in exposed surface area, as reactivity increased with greater surface exposure. Analysis of neighboring residues (within 7 Å) in three-dimensional space revealed some interesting observations. Among the lysines for which we have quantified reactivity, glutamate was the most abundant residue near lysine, but closer proximity (3.4 – 4 Å) yielded low reactivity, while those 5 – 6.6 Å away tended to display higher reactivity. Similar trends were noted with aspartate. Interestingly, lysine was found within a very narrow distance range 6 – 7 Å, and most of the corresponding paired lysines showed greater than average reactivity. Collectively, our results suggests that surface exposure and local electrostatic interactions influence lysine reactivity toward AcP and AcCoA. Surprisingly, pK_a values computed from structure were not a reliable predictor of lysine reactivity. Similar conclusions were drawn by Kuhn et al., after assessing in vivo acetylation sites likely modified by AcP in bacteria.(18) Among the bacterial proteins presumably acetylated by AcP, the linear sequence around the acetylated lysine tended to favor acidic residues glutamate and aspartate. Curiously, in vitro lysine reactivity toward short immobilized peptides favored lysines with a basic lysine or arginine at -1 and +1 position. Here, our results with ACAT1 and GDH suggest that neighboring basic and acidic residues in 3-D space influence lysine reactivity. Glutamate and aspartate residues engaged in strong salt bridges with lysine yield poor acetylation reactivity, while these acidic groups at distances 5-7 Å permit higher reactivity (**Figure 3-9**). Neighboring lysines found within a range of 6-7 Å appeared among

the more reactive lysines. Future studies will to needed to establish these trends at the proteome level and to chemically evaluate how local electrostatics direct lysine reactivity.

We previously analyzed the substrate specificity of SIRT3 both in vitro and in vivo.(4, 40) These results consistently revealed a strong preference for acetylated lysines in peptides containing basic residues. Here, we noted that many of the most reactive lysine residues exist as clusters in 3-dimensional space within 6-12 Å and are exemplified by K477-K480 in GDH, and K260-K263-K265 and K171-K178-K187 in ACAT1. Strikingly, these lysines exhibited some of the largest fold-changes in acetylation in liver mitochondria when SIRT3 is absent (**Figure 3-6**), suggesting that they represent bona fide targets of SIRT3. Taken together, highly reactive sites that tend to exist within clusters of lysine residues are likely high-affinity substrates of SIRT3. Importantly, the observation that higher chemical reactivity tracks with the likelihood of dynamic modification in vivo, provides evidence that enzyme-catalyzed acylation might not be necessary to explain the prevalence of protein acetylation in mitochondria. Furthermore, the second order rate constants measured in this study are sufficiently fast to be biologically relevant. The quantitative analysis described here can be directly applied to evaluating targeted and proteome-wide reactivities.

Figure 3-8 Comparison of second order rates with intrinsic protein properties

(A) Scatter plot of average B-factor and second order acetylation rate constants with linear fit. B-factor of non-reactive lysines shown below, and not included in regression. Includes observed sites from GDH. (B) Scatter plot of predicted lysine pK_a and second order acetylation rate constants with linear fit. Predicted pK_a value of quantified non-reactive sites represented below and not included in regression. Sites plotted at a pK_a of 12 correspond to H++ server output designated as pK_a of ≥12 (the maximum pK value returned by the computation). Includes all observed sites from ACAT1 and GDH. (C) Scatter plot of solvent accessible surface area (SASA) and second order acetylation rate constants with linear fit. SASA represented as the fraction of each lysine residue accessible. SASA of non-reactive lysines shown below, and not included in regression. Analysis includes all observed sites from ACAT1 and GDH.

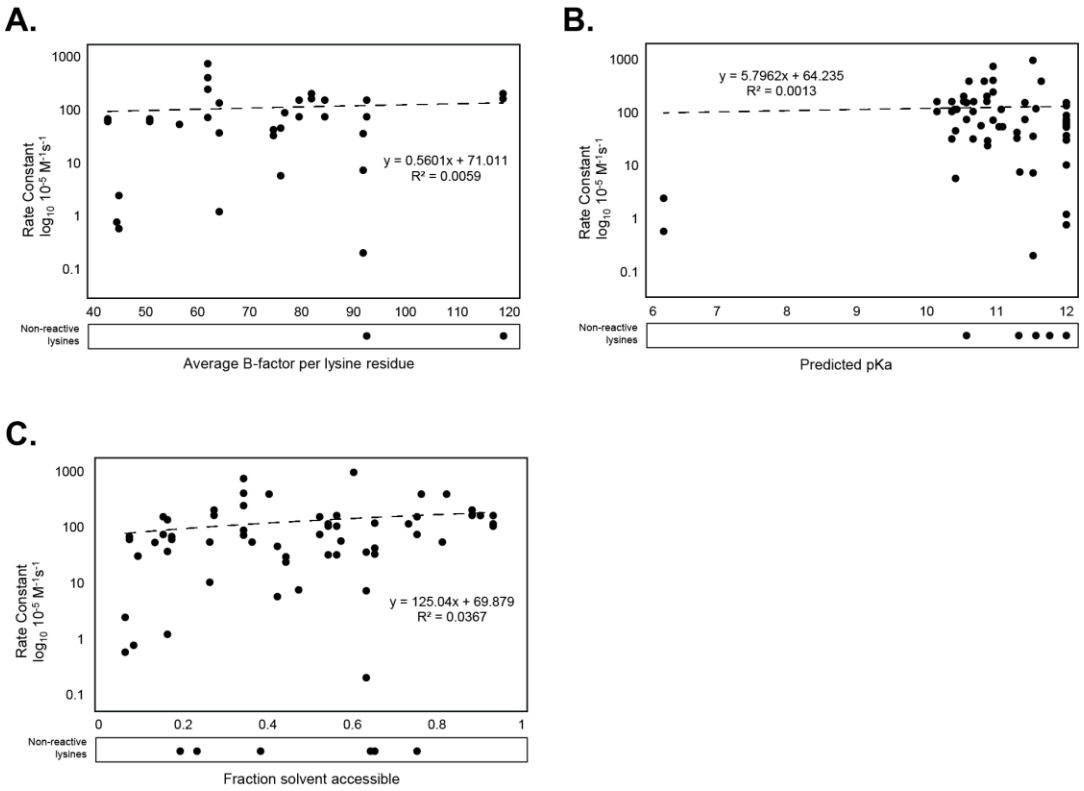
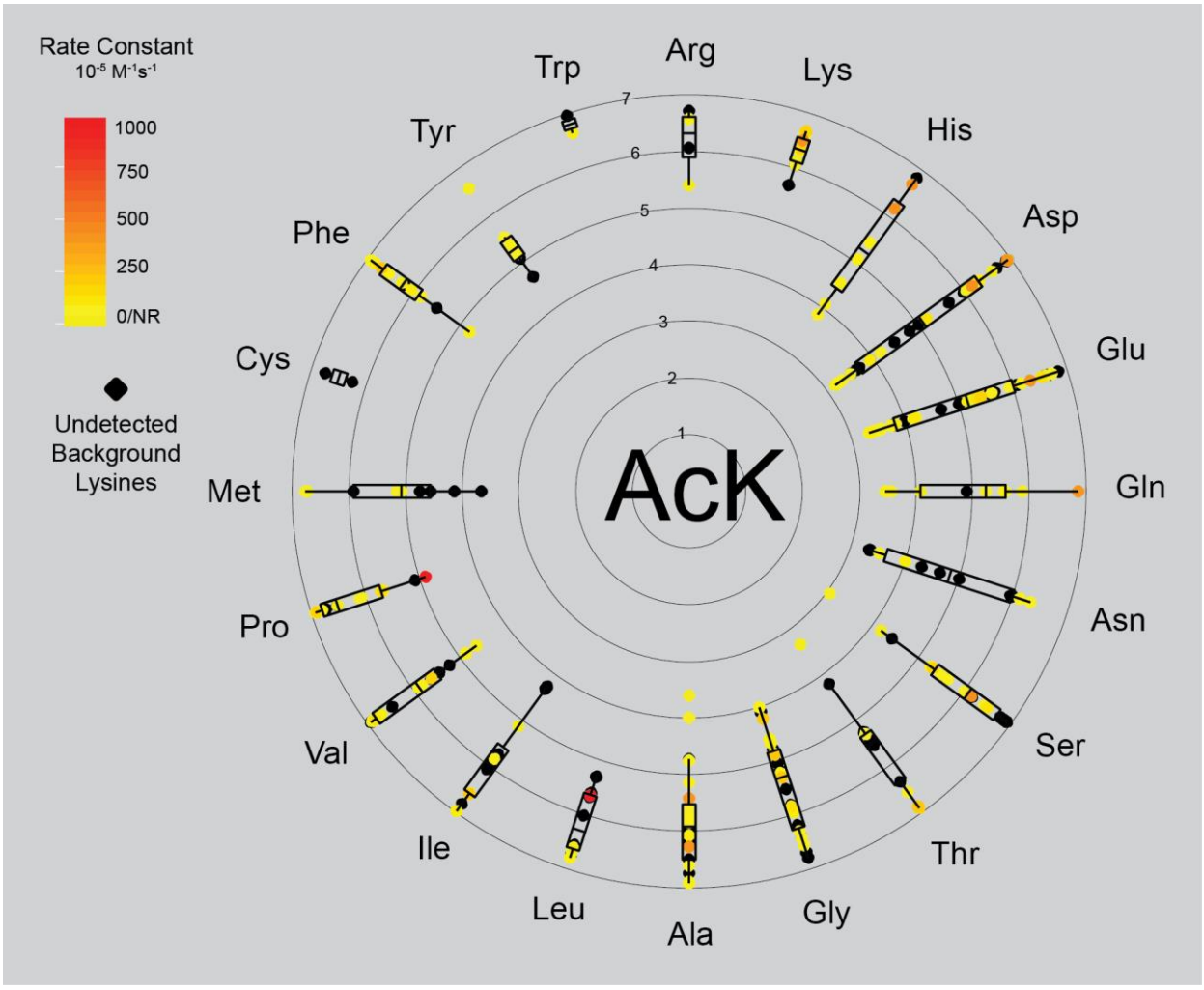


Figure 3-9 Three-dimensional environment of acetyl-lysines in ACAT1 and GDH.

Pairwise distances between the side chain interaction center of each lysine and all neighboring residues in a 7 Å sphere. Distance markers are colored in accordance with the rate constant associated with the neighboring lysine. Lysines not detected in this study are included (colored in grey) as a background to aid in interpretability. NR on rate constant color scale refers to quantified non-reactive lysines.



3.5 References:

1. Xiong, Y., and Guan, K. L. (2012) Mechanistic insights into the regulation of metabolic enzymes by acetylation, J. Cell Biol. 198, 155–164.
2. Yang, X. J., and Seto, E. (2008) Lysine acetylation: codified crosstalk with other posttranslational modifications, Mol. Cell 31, 449–461.
3. 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.
4. Hebert, A. S., Dittenhafer-Reed, K. E., Yu, W., Bailey, D. J., Selen, E. S., Boersma, M. D., Carson, J. J., Tonelli, M., Balloon, A. J., Higbee, A. J., Westphall, M. S., Pagliarini, D. J., Prolla, T. A., Assadi-Porter, F., Roy, S., Denu, J. M., and Coon, J. J. (2013) Calorie restriction and SIRT3 trigger global reprogramming of the mitochondrial protein acetylome, Mol. Cell 49, 186–199.
5. Rardin, M. J., Newman, J. C., Held, J. M., Cusack, M. P., Sorensen, D. J., Li, B., Schilling, B., Mooney, S. D., Kahn, C. R., Verdin, E., and Gibson, B. W. (2013) Label-free quantitative proteomics of the lysine acetylome in mitochondria identifies substrates of SIRT3 in metabolic pathways, Proc. Natl. Acad. Sci. U. S. A. 110, 6601–6606.
6. Lundby, A., Lage, K., Weinert, B. T., Bekker-Jensen, D. B., Secher, A., Skovgaard, T., Kelstrup, C. D., Dmytriyev, A., Choudhary, C., Lundby, C., and Olsen, J. V. (2012) Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns, Cell Rep. 2, 419–431.
7. Sol, E. M., Wagner, S. A., Weinert, B. T., Kumar, A., Kim, H.-S., Deng, C.-X., and Choudhary, C. (2012) Proteomic investigations of lysine acetylation identify diverse

substrates of mitochondrial deacetylase sirt3, PLoS One, Public Library of Science 7, e50545.

8. Still, A. J., Floyd, B. J., Hebert, A. S., Bingman, C. A., Carson, J. J., Gunderson, D. R., Dolan, B. K., Grimsrud, P. A., Dittenhafer-Reed, K. E., Stapleton, D. S., Keller, M. P., Westphall, M. S., Denu, J. M., Attie, A. D., Coon, J. J., and Pagliarini, D. J. (2013) Quantification of Mitochondrial Acetylation Dynamics Highlights Prominent Sites of Metabolic Regulation, J. Biol. Chem. 288, 26209–26219.
9. Feldman, J. L., Dittenhafer-Reed, K. E., and Denu, J. M. (2012) Sirtuin catalysis and regulation, J. Biol. Chem. 287, 42419–42427.
10. Lombard, D. B., Alt, F. W., Cheng, H.-L., Bunkenborg, J., Streeper, R. S., Mostoslavsky, R., Kim, J., Yancopoulos, G., Valenzuela, D., Murphy, A., Yang, Y., Chen, Y., Hirschey, M. D., Bronson, R. T., Haigis, M., Guarente, L. P., Farese, R. V., Jr, Weissman, S., Verdin, E., and Schwer, B. (2007) Mammalian Sir2 homolog SIRT3 regulates global mitochondrial lysine acetylation, Mol. Cell. Biol. 27, 8807–8814.
11. Hallows, W. C., Yu, W., Smith, B. C., Devries, M. K., Devires, M. K., Ellinger, J. J., Someya, S., Shortreed, M. R., Prolla, T., Markley, J. L., Smith, L. M., Zhao, S., Guan, K.-L., and Denu, J. M. (2011) Sirt3 promotes the urea cycle and fatty acid oxidation during dietary restriction, Mol. Cell 41, 139–149.
12. Newman, J. C., He, W., and Verdin, E. (2012) Mitochondrial protein acylation and intermediary metabolism: regulation by sirtuins and implications for metabolic disease, J. Biol. Chem. 287, 42436–42443.
13. Wagner, G. R., and Hirschey, M. D. (2014) Nonenzymatic protein acylation as a carbon stress regulated by sirtuin deacylases, Mol. Cell 54, 5–16.
14. Paik, W. K., Pearson, D., Lee, H. W., and Kim, S. (1970) Nonenzymatic acetylation of

histones with acetyl-CoA, Biochim. Biophys. Acta 213, 513–522.

15. Tanner, K. G., Langer, M. R., and Denu, J. M. (2000) Kinetic mechanism of human histone acetyltransferase P/CAF, Biochemistry 39, 15652.
16. Kuo, Y.-M., and Andrews, A. J. (2013) Quantitating the specificity and selectivity of Gcn5-mediated acetylation of histone H3, PLoS One, Public Library of Science 8, e54896.
17. 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.
18. Kuhn, M. L., Zemaitaitis, B., Hu, L. I., Sahu, A., Sorensen, D., Minasov, G., Lima, B. P., Scholle, M., Mrksich, M., Anderson, W. F., Gibson, B. W., Schilling, B., and Wolfe, A. J. (2014) Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation, PLoS One, Public Library of Science 9, e94816.
19. Weinert, B. T., Iesmantavicius, V., Wagner, S. A., Schölz, C., Gummesson, B., Beli, P., Nyström, T., and Choudhary, C. (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in E. coli, Mol. Cell 51, 265–272.
20. Schwer, B., Eckersdorff, M., Li, Y., Silva, J. C., Fermin, D., Kurtev, M. V., Giallourakis, C., Comb, M. J., Alt, F. W., and Lombard, D. B. (2009) Calorie restriction alters mitochondrial protein acetylation, Aging Cell 8, 604–606.
21. Hirschey, M. D., Shimazu, T., Jing, E., Grueter, C. A., Collins, A. M., Aouizerat, B., Stančáková, A., Goetzman, E., Lam, M. M., Schwer, B., Stevens, R. D., Muehlbauer, M. J., Kakar, S., Bass, N. M., Kuusisto, J., Laakso, M., Alt, F. W., Newgard, C. B., Farese, R. V., Jr, Kahn, C. R., and Verdin, E. (2011) SIRT3 deficiency and mitochondrial protein hyperacetylation accelerate the development of the metabolic syndrome, Mol. Cell 44, 177–

190.

22. Berndsen, C. E., Albaugh, B. N., Tan, S., and Denu, J. M. (2007) Catalytic mechanism of a MYST family histone acetyltransferase, Biochemistry 46, 623–629.
23. Still, A. J., Floyd, B. J., Hebert, A. S., Bingman, C. A., Carson, J. J., Gunderson, D. R., Dolan, B. K., Grimsrud, P. A., Dittenhafer-Reed, K. E., Stapleton, D. S., Keller, M. P., Westphall, M. S., Denu, J. M., Attie, A. D., Coon, J. J., and Pagliarini, D. J. (2013) Quantification of mitochondrial acetylation dynamics highlights prominent sites of metabolic regulation, J. Biol. Chem. 288, 26209–26219.
24. Su, Z., Boersma, M. D., Lee, J.-H., Oliver, S. S., Liu, S., Garcia, B. A., and Denu, J. M. (2014) ChIP-less analysis of chromatin states, Epigenetics Chromatin 7, 7.
25. Sol, E. M., Wagner, S. A., Weinert, B. T., Kumar, A., Kim, H. S., Deng, C. X., and Choudhary, C. (2012) Proteomic investigations of lysine acetylation identify diverse substrates of mitochondrial deacetylase sirt3, PLoS One 7, e50545.
26. 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.
27. 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.
28. Kopp, J., and Schwede, T. (2004) The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models, Nucleic Acids Res. 32, D230–4.
29. Anandakrishnan, R., Aguilar, B., and Onufriev, A. V. (2012) H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations, Nucleic Acids Res. 40, W537–41.

30. Myers, J., Grothaus, G., Narayanan, S., and Onufriev, A. (2006) A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules, Proteins 63, 928–938.

31. Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S., and Onufriev, A. (2005) H++: a server for estimating pKas and adding missing hydrogens to macromolecules, Nucleic Acids Res. 33, W368–71.

32. Cavallo, L., Kleinjung, J., and Fraternali, F. (2003) POPS: A fast algorithm for solvent accessible surface areas at atomic and residue level, Nucleic Acids Res. 31, 3364–3366.

33. Fraternali, F., and Cavallo, L. (2002) Parameter optimized surfaces (POPS): analysis of key interactions and conformational changes in the ribosome, Nucleic Acids Res. 30, 2950–2960.

34. Delano, W. L. (2002) The PyMOL Molecular Graphics System.

35. Bahar, I., and Jernigan, R. L. (1996) Coordination geometry of nonbonded residues in globular proteins, Fold. Des. 1, 357–370.

36. 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.

37. Du, J., Zhou, Y., Su, X., Yu, J. J., Khan, S., Jiang, H., Kim, J., Woo, J., Kim, J. H., Choi, B. H., He, B., Chen, W., Zhang, S., Cerione, R. A., Auwerx, J., Hao, Q., and Lin, H. (2011) Sirt5 Is a NAD-Dependent Protein Lysine Demalonylase and Desuccinylase, Science 334, 806–809.

38. Smith, T. J., Schmidt, T., Fang, J., Wu, J., Siuzdak, G., and Stanley, C. A. (2002) The structure of apo human glutamate dehydrogenase details subunit communication and allostery, J. Mol. Biol. 318, 765–777.

39. Banerjee, S., Schmidt, T., Fang, J., Stanley, C. A., and Smith, T. J. (2003) Structural Studies on ADP Activation of Mammalian Glutamate Dehydrogenase and the Evolution of Regulation†,‡, Biochemistry 42, 3446–3456.

40. Smith, B. C., Settles, B., Hallows, W. C., Craven, M. W., and Denu, J. M. (2011) SIRT3 substrate specificity determined by peptide arrays and machine learning, ACS Chem. Biol. 6, 146–157.

Chapter 4: Towards decoding cellular regulatory acetylation networks using quantitative acetylation stoichiometry

This chapter is a manuscript in preparation for submission with the following authors/affiliations:

Josue Baeza[1, 2], Jing Fan[2, 3], Ian Lienert[4], Tejas Gandhi[4], Oliver Bernhardt[4], Lukas Reiter[4], John M. Denu[1, 2]

[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]Biognosys AG, Wagistrasse 25, CH-8952 Schlieren, Switzerland

Author contributions: J.B., J.F., and J.M.D. designed the research study. J.B. performed acetylation stoichiometry MS experiments and analyzed data. J.F. performed single amino acid MS experiments and analyzed data. I.L., T.G., O.B., and L.R. developed software to analyze DIA MS data. J.B. performed bioinformatic analysis, designed figures and drafted the manuscript. J.B., J.F. and J.M.D edited manuscript. J.B. and J.F. are co-first authors. J.M.D. is the corresponding author.

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 the National Science Foundation Graduate Research Fellowship Program (NSF-GRFP) DGE-1256259 (J.B.).

4.1 Abstract

Protein acetylation is a widespread and regulatory modification in cells controlled by enzymatic and nonenzymatic mechanisms. With over 20,000 acetyl-lysine sites identified in mammalian cells, a way to distinguish between the majority of functional and spurious acetylation is needed. In this study, we developed a method to quantify site-specific and global acetylation stoichiometry using complementary mass spectrometry approaches. We also demonstrate how lysine acetylation stoichiometry is distributed across the cell, with the nuclear compartment displaying the highest amount of proteins with high stoichiometry. These results highlight the functional role played by acetyltransferases and deacetylases in this compartment. Using a metabolic tracer of ^{13}C-glutamine to generate a cytoplasmic-specific ^{13}C-acetyl-CoA, we were able to identify proteins acetylated by this compartment-specific metabolite and demonstrate that some mitochondrial proteins can be modified by a cytoplasmic acetyl-CoA. Quantifying acetylation stoichiometry is becoming a powerful tool to investigate proteome-wide acetylation, regardless of whether the reaction is enzyme catalyzed.

4.2 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 tails 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. Additionally, bromodomains recognize and bind to acetyl-lysine residues for recruitment of larger multisubunit complexes (8). Acetylation of cytoplasmic proteins affect diverse cellular processes to include cell migration, cytoskeleton dynamics, metabolism, and aging, while mitochondrial acetylation has been linked to metabolic regulation, oxidative stress, OXPHOS, and mitochondrial gene expression (5). Acetylation was previously thought to be under the direct control of acetyltransferases, however recent evidence suggests that some sites could be the result of a nonenzymatic mechanism. Accordingly, the measured second order rate constants of nonenzymatic acetylation can explain some of the in vivo acetylation sites, especially those with low stoichiometry acetylation (9). With the number of in vivo acetylation sites reaching over 20,000 (10), a way to prioritize the functional and regulatory acetylation sites is needed, whether or not acetylation is enzyme catalyzed.

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. In fact, evidence suggests that acetylation exerts its effect by modulating protein-protein and protein-DNA interactions, cellular localization, enzyme activity and stability. In mitochondria,

acetylation generally acts as an inhibitory mark for enzyme function (5). In this regard, acetylation appears to function as a rheostat to modulate the degree of a biochemical process. Therefore, quantifying acetylation stoichiometry would provide critical information to understand regulatory acetylation networks and to predict and model global cellular effects.

In this study, we used complementary mass spectrometry methods to quantify global and site-specific acetylation stoichiometry in MCF7 cells, ranging from less than 1% to 97%. We also determined the cellular distribution of stoichiometry, spanning cellular compartments such as the cytoplasm, mitochondria, and nucleus including histone and non-histone proteins, as well as proteins residing in the endoplasmic reticulum. The nuclear compartment, which contains the majority of documented lysine acetyltransferases, displayed the highest acetylation stoichiometry, while mitochondrial and cytoplasmic proteins displayed the lowest stoichiometry. Additionally, RNA binding proteins including ribosomal proteins were enriched in high acetylation stoichiometry. Using a novel form of monitoring cellular acetylation stoichiometry, which we termed, cellular reverse-stoichiometry, a metabolic tracer of glutamate is used to generate a cytoplasmic-specific pool of isotopic acetyl-CoA, Using this strategy, we determine that a small subset of mitochondrial-localized proteins can be acetylated using a cytoplasmic pool of acetyl-CoA, suggesting that some mitochondrial proteins can be acetylated prior to import. This quantitative strategy, to quantify acetylation stoichiometry, is a critical tool for prioritizing the ever-increasing number of detected lysine acetylation sites and towards a deeper understanding of this regulatory modification.

4.3 Experimental procedures

4.3.1 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. For experiments where the metabolic tracer of L-glutamine-5-^{13}C was added, MCF7 cells were grown in standard DMEM media with 10% FBS. To begin the experiment, MCF7 cells were split in the morning to a low confluency and cultured in standard DMEM with 10% FBS. Once adhered to the plate, the cells were washed with PBS twice and media was replaced to glutamine-free DMEM with 10% FBS supplemented with L-glutamine-5-^{13}C (Sigma). 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.

4.3.2 Sample preparation

4.3.2.1 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 (9, 11, 12). 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.

4.3.2.2 Digesting proteomic sample to single amino acids

The enzymatic digestion of proteomic samples converts proteins to single amino acids. N-ε-acetylation of lysine residues are converted to N-ε-acetyl lysine. Lysine is an essential amino acid and is therefore not labeled. Additionally, acetyl-lysine is isobaric with Gly-Leu, Gly-Ile, and Ala-Glu, so it is important to achieve complete digestion. Begin by diluting 20 µg of each sample into 50 µL of digestion buffer (50 mM ammonium bicarbonate, pH 7.5, 5 mM DTT, using LC–MS grade water). Include a procedural blank with 50 µL digestion buffer and protein. Next add 0.4 µg Pronase to each sample and incubate for 24 hr at 37°C. To quench the reaction, heat samples to 95°C for 5 min and allow to cool down to ambient temperature. Then, add 0.8 µg aminopeptidase and incubate at 37°C for 18 hr. Heat samples to 95°C for 5 min and cool down. Finally, digest with 0.4 µg prolidase and incubate at 37°C for 3 hr. To each sample, add 200 µL LC–MS grade acetonitrile (ACN) to quench the reaction and vortex for 5 sec. Spin at maximal speed for 5 min and transfer supernatant to a clean LCMS vial.

4.3.2.3 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, 110Å, 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 were dried down using a speedvac and pooled by concatenation into 6 final fractions as described previously (13).

4.3.3 Mass spectrometry

4.3.3.1 Liquid chromatography

Peptides were separated with a Dionex Ultimate 3000 RSLCnano HPLC using a Waters Atlantis dC18 (100 µm x 150 mm) 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.

4.3.3.2 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 20, dynamic exclusion was set for 30 seconds, and a loop count of 20.

4.3.3.3 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.

4.3.3.4 LCMS analysis of single amino acids

To measure the abundance of lysine, acetyl-lysine, and other amino acids, samples including the procedural blank, are analyzed by LC–MS. To quantify absolute levels of lysine and acetyl-lysine, an external calibration curve is used in the same sequence with the experimental samples. Lysine ranges between 10 to 200 μM, while acetyl-lysine ranges between 0.5 to 10 μM. Amino acids are 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. Solvent A is ACN and solvent B is 10 mM ammonium acetate in water, pH 5.5. The gradient profile for chromatography is as follows: 10% solvent B for 2 min, linear increase in solvent B to 90% over 12 min, isocratic 90% solvent B for 3 min, and then equilibration with 10% solvent B for 2 min at a flow rate of 0.3 mL/min. Samples are introduced to the mass spectrometer by heated electrospray ionization. Analysis is performed under positive ionization mode. 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. Data-dependent acquisition mode with a dynamic exclusion of 10 s is enabled. In every cycle, one full MS scan is collected with a scan range of 88–500 m/z, resolution of 70 K, maximum injection time of 40 ms, and AGC

of 1E6. Then three MS2 scans are followed on parent ions from the most intense peaks, with NCE 25. Resolution is set at 17,500, AGC target 2E4, maximum injection time 40 ms, and isolation width 1 m/z.

4.3.4 Data Processing

4.3.4.1 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 methione oxidation as variable modifications and cysteine carbamidomethylation as fixed modification. The MaxQuant search results were imported into Spectronaut to build the ^{12}C-AcK library. The ^{12}C-AcK spectral library was then exported from Spectronaut and imported into a custom spectral library modifier (developed by Spectronaut), which generates all the corresponding heavy fragmentation spectra. With this in silico approach to inflate the ^{12}C-AcK library, every acetyl-lysine peptide precursor will contain the light and heavy fragmentation spectra needed for quantitation.

4.3.4.2 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 quantitaion were used with slight modification: Identification-Qvalue score was set to 0.1 and Workflow-Unify peptide peaks was

selected. The spectral library modified to contain the light and heavy acetyl lysine fragment ions was selected. 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.

4.3.4.3 Stoichiometry data processing

Data processing was performed in R v3.3.2 (Sincere Pumpkin Patch) (http://www.r-project.org/) using an in-house made R script, which is available in the supplementary information. The stoichiometry pipeline consists of two major steps: quantifying fragment specific stoichiometry and natural abundance isotopic correction.

4.3.4.4 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.

4.3.4.5 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/) (14). To 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 $^{Corr}XIC_H$.

4.3.5 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.

4.3.6 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.

4.3.7 Bioinformatics

4.3.7.1 Functional annotation

Functional annotation of enriched gene ontology (GO) terms were assessed using DAVID v6.8 (15, 16). 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.

4.3.7.2 Acetylation site sequence analysis

The amino acid sequence (± 6 amino acids) flanking each acetyl-lysine from each of the five quantiles were compared to the background data set, which includes the sequence space of all lysine stoichiometry sites identified in this study. Acetylation site sequence patterns from each of the five quantiles were computed using the web-based, Java application, iceLogo (17).

4.3.7.3 Subcellular localization assignment

To assign protein subcellular localization, we used the MitoCarta (18, 19) 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.

4.4 Results

4.4.1 DIA acetylation stoichiometry

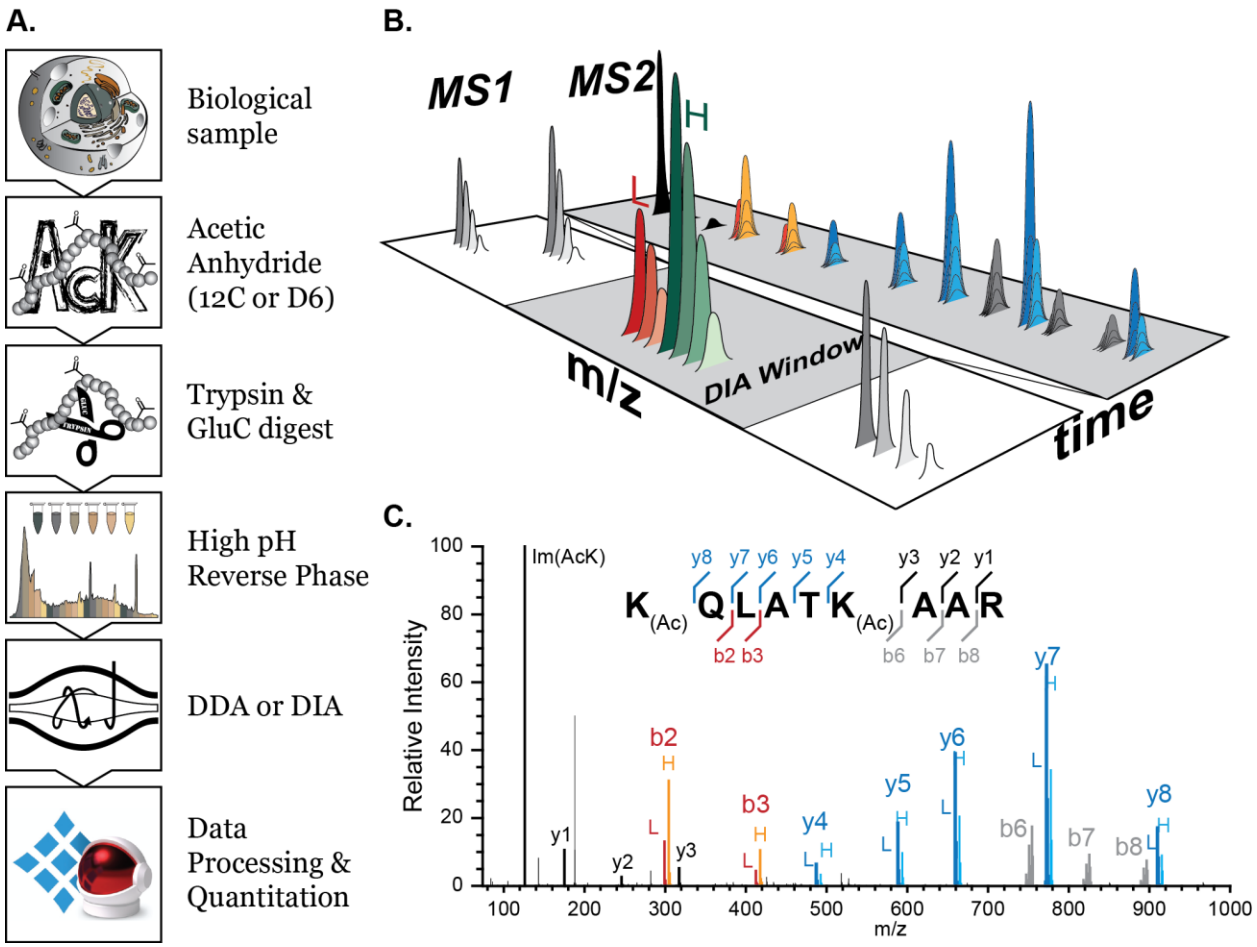
4.4.1.1 DIA acetylation stoichiometry workflow

We previously reported a method to determine lysine acetylation stoichiometry across an entire proteome (11). 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 (11, 20–22) as well as in vitro, nonenzymatic acetylation kinetics (9). Here, we report an improved method to quantify acetylation stoichiometry using data-independent acquisition (DIA) mass spectrometry to quantify global acetylation stoichiometry (**Figure 4-1A**). Briefly, a sample is chemically acetylated using isotopic acetic anhydride and digested with trypsin and GluC, which the sequential digestion of the acetyl-proteome generates shorter peptides for MS analysis. Peptides are then prefractionated using high pH reverse phase (HPRP) chromatography and analyzed in DIA mode using nano-LC-MS/MS and analyzed using Spectronaut™ (23–25). The spectral library for the DIA analysis is generated from a ^{12}C-acetic anhydride labeled sample analyzed in data-dependent acquisition (DDA) mode. For the spectral library to contain both the light and heavy acetyl-lysine fragment ions, a novel, standalone software was built to be used with Spectronaut™ for inflating the ^{12}C-AcK spectral library. This process guarantees that both, light and heavy fragments, are present in the spectral library for identification and quantitation. Combining offline prefractionation and DIA analysis addresses unique limitations of the original study as previously discussed (22). HPRP prefractionation reduces interferences caused by coeluting peptides and has the added benefit of increasing the depth of acetylomic 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 (**Figure 4-1B, C**).

Figure 4-1: DIA acetylation stoichiometry workflow

(A) Diagram for measuring acetylation stoichiometry from whole cell lysate. A biological sample is chemically acetylated using isotopic acetic anhydride followed by proteolytic digestion using trypsin and GluC. Offline fractionation using high pH reverse phase chromatography is used to decrease sample complexity and enables deep acetylomic coverage. MS spectra are acquired using DIA mass spectrometry and analyzed with Spectronaut™. To generate the spectral library, the same workflow is used, but substituting with ^{12}C-acetic anhydride and MS spectra acquired in data-dependent acquisition (DDA) mode. (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.

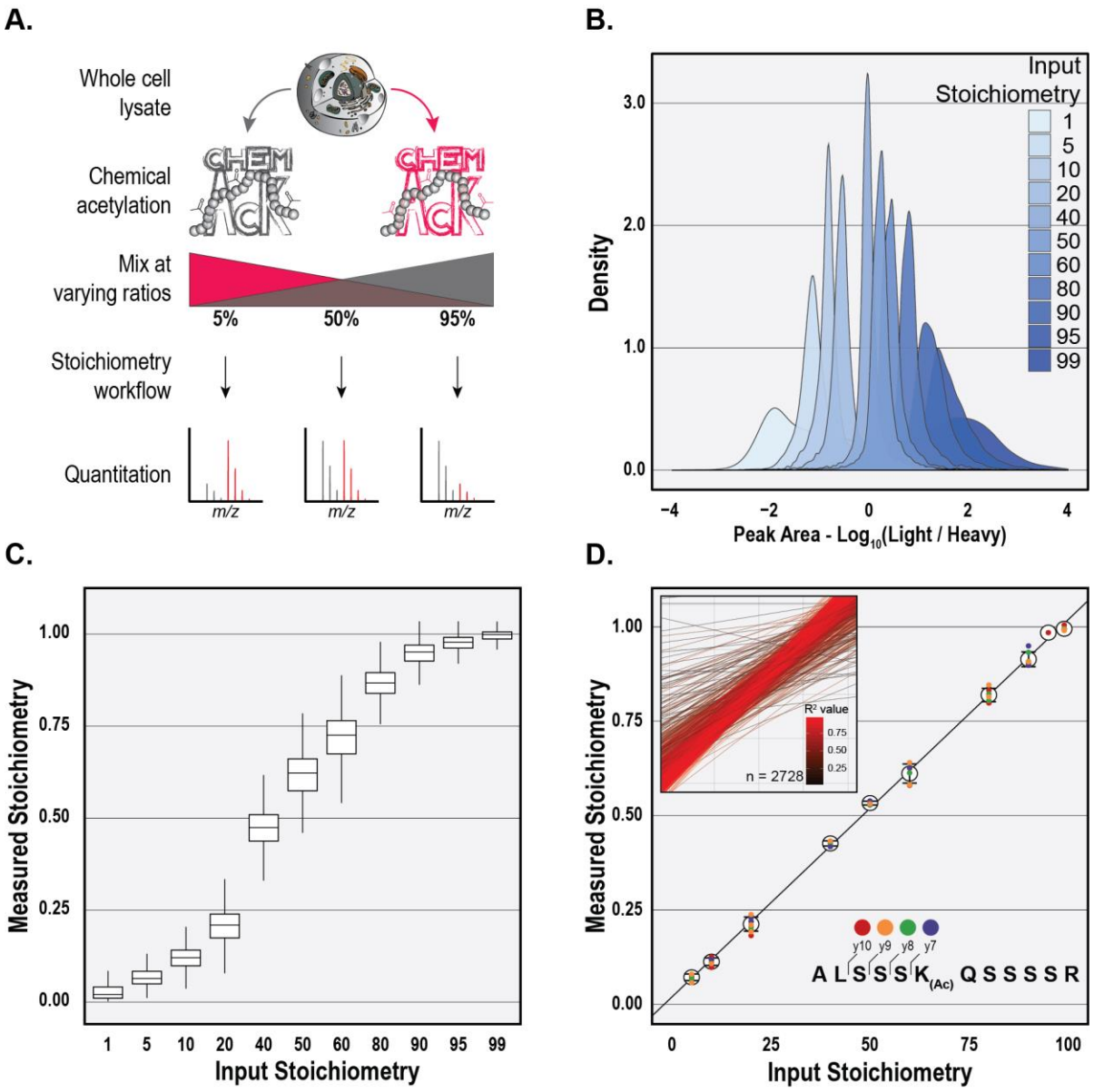

A.
Biological sample
Acetic Anhydride (12C or D6)
Trypsin & GluC digest
High pH Reverse Phase
DDA or DIA
Data Processing & Quantitation
B.
MS1
MS2
H
L
m/z
DIA Window
time
C.
Im(AcK)
y8 y7 y6 y5 y4 y3 y2 y1
K(Ac) Q L A T K(Ac) A A R
b2 b3
b6 b7 b8
Relative Intensity
m/z

4.4.1.2 Stoichiometry curve across an entire proteome

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 **(Figure 4-2A**). As expected, stoichiometry between 40 and 60% displayed the lowest variance (**Figure 4-2B**). This is due to the similar relative abundance 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 highest variance since quantitation in these conditions require the measurement of fragment ions greater than 20-fold difference (**Figure 4-2B**). We next assessed high confidence acetyl lysine sites, defined as sites quantified by multiple fragment ions with ± 5% of the mean stoichiometry. Figure 4-2C shows the distribution of stoichiometry in each condition and generally agrees with the input stoichiometry (**Figure 4-2C**). Linear regression analysis of individual fragment ions plotted as a function of input stoichiometry displays high accuracy and precision for an individual lysine site (**Figure 4-2D**). Additionally, plotting the linear regression of AcK sites having more than nine data points (n = 2728 AcK sites) shows high accuracy of the stoichiometry method (**Figure 4-2D inset**). This global analysis using high confidence fragment ions with defined input stoichiometries highlights the quantitative nature of this method and is applicable to query acetylation stoichiometry of an entire proteome.

Figure 4-2: Validation of our DIA stoichiometry workflow

(A) Diagram representing the experimental setup of the proteome-wide stoichiometry curve. A HEK293 lysate was chemically acetylated using ^{12}C- or D_6-acetic anhydride, mixed at varying ratios, and processed using our workflow. (B) Density plot of the Log_{10}-transformed (light / heavy) peak area every fragment ion in all the input stoichiometry conditions. (C) Boxplot showing the global distribution of each input stoichiometry. High-confidence AcK sites were used, which correspond to sites with multiple fragment ions quantified ± 5%. Note that x-axis is not to scale. (D) Scatter plot displaying the measured stoichiometry plotted as a function of input stoichiometry of acetyl-lysine fragment ions for a single peptide. Fragment ions, y7 – y10, were measured across the majority of input stoichiometries. Circles represent the mean stoichiometry for that input stoichiometry. Linear regression analysis of the measured and input stoichiometry shows high correlation (R^2 = 0.9977 using Pearson correlation). To plot the global linear regression analysis, we used AcK sites with greater than nine input stoichiometry data points (n = 2728). This analysis reveals the high correlation between the measured and expected stoichiometry as the majority of AcK sites display high R^2 values (red lines).


A.
Whole cell lysate
Chemical acetylation
Mix at varying ratios
5%
50%
95%
Stoichiometry workflow
Quantitation
m/z
m/z
m/z
B.
Input Stoichiometry
1
5
10
20
40
50
60
80
90
95
99
Density
3.0
2.0
1.0
0.0
−4
−2
0
2
4
Peak Area - Log10(Light / Heavy)
C.
Measured Stoichiometry
1.00
0.75
0.50
0.25
0.00
1
5
10
20
40
50
60
80
90
95
99
Input Stoichiometry
D.
Measured Stoichiometry
1.00
0.75
0.50
0.25
0.00
R² value
0.75
0.50
0.25
n = 2728
y10 y9 y8 y7
A L S S S K(Ac) Q S S S S R
0
25
50
75
100
Input Stoichiometry

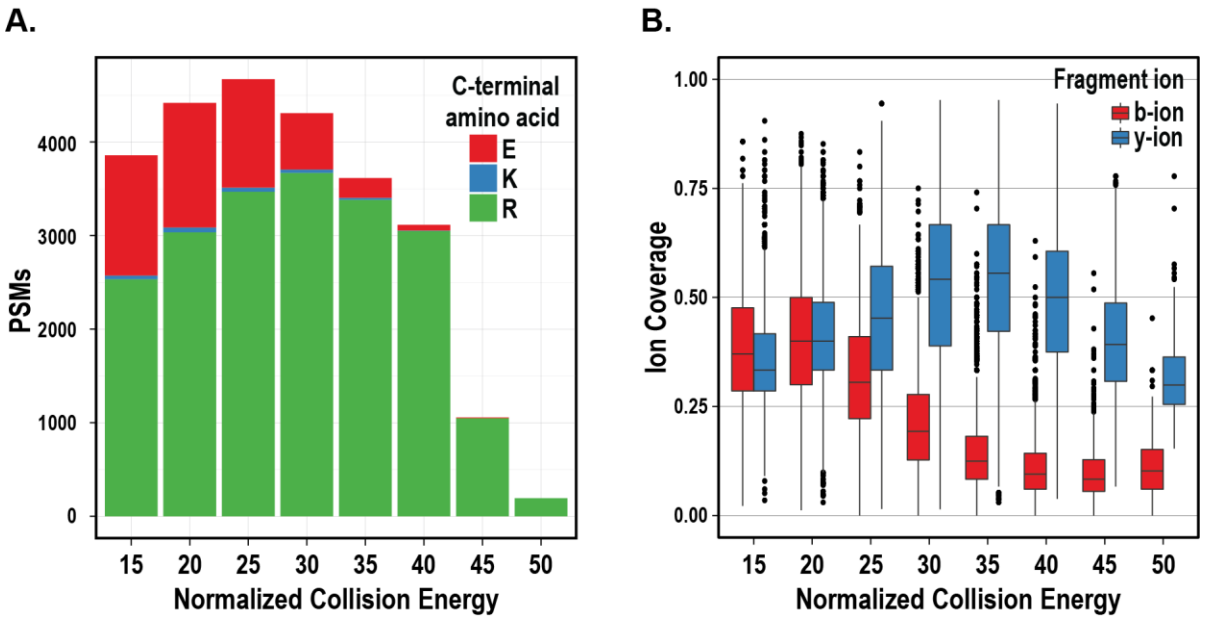
4.4.1.3 NCE Optimization for deep proteomic coverage of acetyl-lysine sites

Our improved stoichiometry workflow enables the quantitation of different lysines from a single peptide, thus removing any ambiguity of site quantitation. As an example, the histone H3 peptide, K_{Ac}QLATK$_{Ac}$AAR (residues K18-R26), has unique fragment ions specific to K18Ac and K26Ac. K18Ac is quantified by the fragment ions b2-b3, while y4-y8 are specific for K18Ac (**Figure 4-1C**). For this strategy to be effective at the proteomic level, obtaining high quality and high coverage of b- and y-ions during MS2 fragmentation is necessary. Using our Q-Exactive system, we optimized the normalized collision energy (NCE) for higher energy collisional dissociation (HCD) fragmentation (26) and evaluated the number of peptide spectral matches (PSMs) as well as the global b- and y-ion coverage across a wide range of NCEs (15-50 in 5 unit increments) using a chemically acetylated, trypsin and GluC digested proteome. Optimal NCE settings were different for trypsin- and gluC-digested peptides corresponding to c-terminal arginine and glutamate, respectively. However, using the combined digestion in our workflow, NCE 25 is optimal for obtaining the maximum amount of PSMs (**Figure 4-3A**). 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.

To determine the global b- and y- fragment ion coverage, we counted the number of each fragment ion identified for each PSM and normalized to the peptide length. As y-ions increase at higher NCE, b-ion coverage begins to decline at a similar rate (**Figure 4-3B**). Using our setup, the optimal setting is NCE 25, which balances b-ion coverage as well as the number of PSMs identified. This type of analysis reveals the importance of optimizing MS conditions for a unique proteomic sample.

Figure 4-3: Optimization of MS2 setting 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, followed by HPRP fractionation and analyzed in DDA mode using varying NCE settings. (B) Global fragment ion coverage for b- and y-ions of the data in A.



4.4.2 Subcellular distribution of acetylation stoichiometry

Studies measuring (or estimating) acetylation stoichiometry in mammalian systems are limited (21, 27) and have not provided a clear picture of the acetylation stoichiometry distribution across the cell. How does acetylation stoichiometry compare across subcellular compartments? To address this question, we utilized our quantitative stoichiometry approach across MCF7 cells using three biological replicates. As we showed using the proteome-wide stoichiometry curve, we were able to quantify a wide range of stoichiometry (< 1% up to 97%) using an MCF7 cell lysate with high correlation between biological replicates when considering every acetyl-lysine fragment ion (grey), acetyl-lysine site (blue), and high confidence acetyl-lysine sites (red) (**Figure 4-4A**). Additionally, when considering the high confidence acetyl-lysine sites (red), the bulk of quantified stoichiometry is less than 10%, however, a small subset of sites seems to cluster at a high stoichiometry range, > 85% (**Figure 4-4A**).

To determine the subcellular distribution of acetylation stoichiometry in our dataset, we grouped each uniprot accession identifier into nuclear, mitochondrial, or cytoplasmic compartments based on which fraction that specific compartment would appear during a crude differential centrifugation protocol. Using the Uniprot subcellular location (http://www.uniprot.org) and the Mitocarta database (18, 19) for protein localization assignments, the nuclear fraction contains significantly more proteins with a higher acetylation stoichiometry compared to the mitochondrial and cytoplasmic fractions (**Figure 4-4B**). Mass spectrometry, by its nature, identifies and quantifies high abundant precursor ions. Therefore, our comparison of nuclear, cytoplasmic, and mitochondrial acetylation stoichiometry may not be accounting for high stoichiometry on low abundant proteins in these subcellular compartments. To address this potential limitation, we utilized a complementary approach to measure global stoichiometry, which

accounts for every lysine within the sample by digesting the proteome into single amino acids and quantifying the abundance of modified and unmodified lysines in a sample (12) (**Figure 4-4C**). This approach reduces sample complexity and directly measures the global stoichiometry. We performed a crude cellular fractionation using differential centrifugation of MCF7 cells, and enriched for cytoplasmic, mitochondrial and nuclear fractions. The nuclear fraction was further divided into histone and non-histone pools by acid extraction. The histone and non-histone nuclear pool displayed the highest acetylation, while the cytoplasm and mitochondrial displayed lower acetylation levels, corroborating our stoichiometry results (**Figure 4-4D**).

4.4.3 Acetyl and methyl groups higher on lysine residues than their metabolite counterparts

Our amino acid quantitation approach (**Figure 4-4C**) allowed us to quantify not only acetyl lysine residues, but other lysine post translational modifications (PTMs). Surprisingly, other lysine PTMs such as mono-, di-, and tri-methylation also displayed a profile similar to that of acetyl lysine: highest in the nuclear compartment and lower in the cytoplasm and mitochondria (**Figure 4-4E, G**). To compare the relative abundance of lysine PTMs to the levels of their small molecule metabolite counterparts, we measured the cellular levels of acetyl-CoA, CoA-SH, S-adenosyl-methionine (SAM), and S-adenosyl-homocysteine (SAH) from MCF7 cells. Unexpectedly, there is more acetyl and methyl groups present on lysine residues than as the small molecule metabolites, acetyl-CoA and SAM, from which they are derived (**Table 4-1**).

Figure 4-4: 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 every fragment ion stoichiometry (grey), every acetyl-lysine site stoichiometry (blue), and high confidence acetyl-lysine site stoichiometry (red). (B) Distribution of acetylation stoichiometry across cytoplasmic, mitochondrial, and nuclear compartments. Stoichiometry dot plot and boxplot of each subcellular compartment is shown. Statistical analysis was performed using the posthoc kruskal nemenyi test using high confidence acetyl stoichiometry sites (p-values are listed for each comparison, ns = not significant). Colored circles represent individual subcellular location assignment based on Uniprot location and mitocarta database (18, 19). (C) Diagram representing workflow for quantifying single amino acid abundances of modified and unmodified lysines. Nuclear, mitochondrial and cytoplasmic fractions are enriched from MCF7 cells by differential centrifugation. Histones are isolated from the nuclear fraction by acid extraction, separating the nuclear fraction into histone and non-histone pool. 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. (D-G) lysine-PTM:lysine ratio across cytoplasmic, mitochondrial, histone, and non-histone proteins for acetyl, mono-. di-, and tri-methyl lysines.

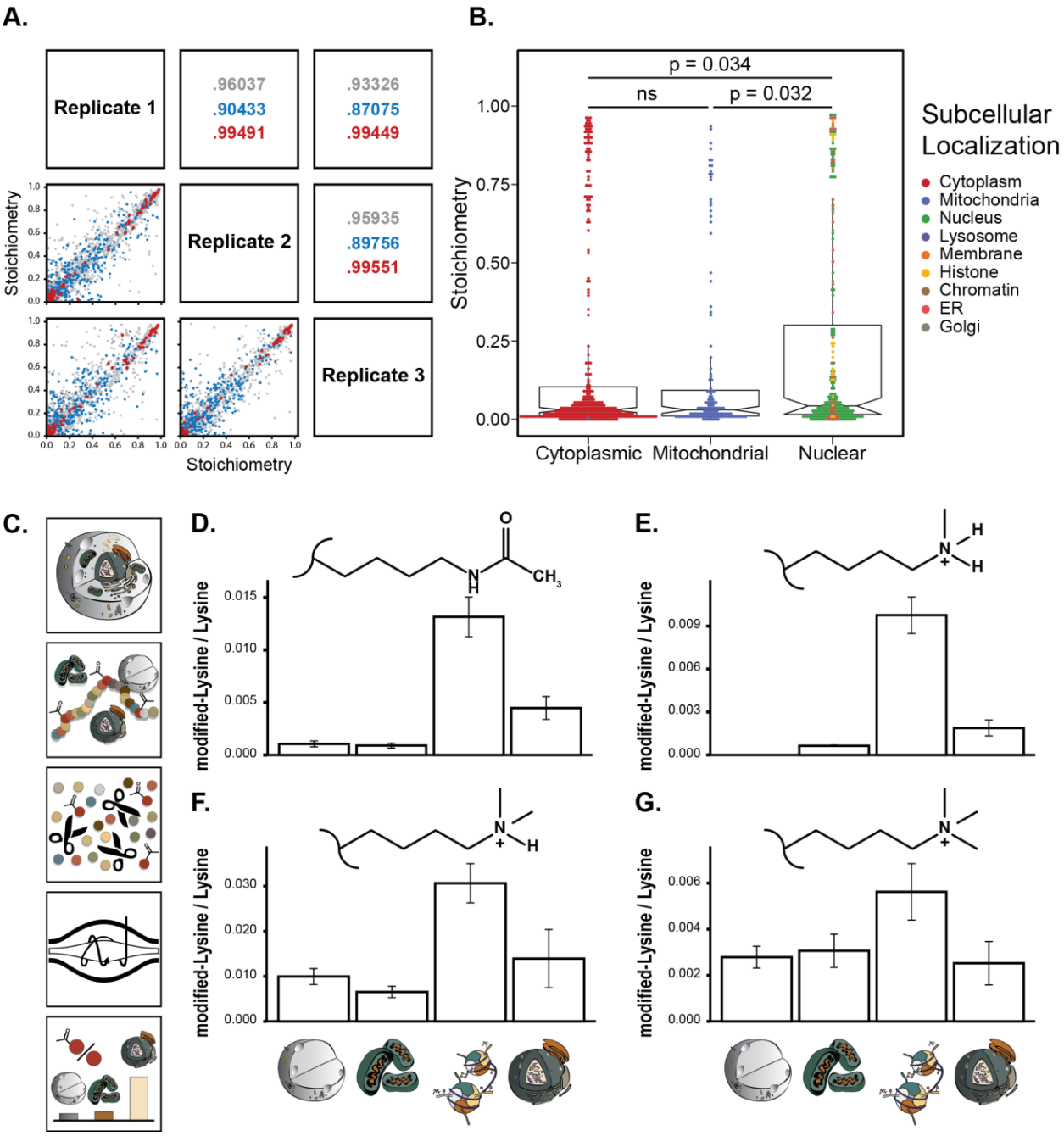

A.
Replicate 1
.96037
.90433
.99491
.93326
.87075
.99449
Replicate 2
.95935
.89756
.99551
Replicate 3
Stoichiometry
Stoichiometry
B.
p = 0.034
ns
p = 0.032
Stoichiometry
Cytoplasmic
Mitochondrial
Nuclear
Subcellular Localization
Cytoplasm
Mitochondria
Nucleus
Lysosome
Membrane
Histone
Chromatin
ER
Golgi
C.
D.
modified-Lysine / Lysine
E.
modified-Lysine / Lysine
F.
modified-Lysine / Lysine
G.
modified-Lysine / Lysine

Table 4.1 Quantitation of small molecular metabolites in MCF7 cells

Metabolite	Abundance (pmol/uL)	Standard Deviation
Acetyl-Coenzyme A	11.7	5.50
Coenzyme A	3.60	0.20
S-adenosyl-methionine (SAM)	78.5	10.7
S-adenosyl-homocysteine (SAH)	0.60	0.04

4.4.4 Bioinformatics

4.4.4.1 Genes involved in RNA binding are enriched in high stoichiometry

To understand which biological processes are enriched in high acetylation stoichiometry, we performed a functional annotation analysis using DAVID (15, 16). First, we divided the high confidence stoichiometry gene list in five equal quantiles (n = 96 or 97). Quantile A consists of very low stoichiometry (< 1%) and Quantile E consists of high stoichiometry (> 27%). For enrichment analysis, each of the five quantiles were compared to the complete gene list identified in the DDA-MS spectral library (2400 unique gene ids). Gene ontology (GO) terms associated with RNA-binding, translation, ribosome, spliceosome were significantly enriched in our acetylation dataset (**Figure 4-5A**). Interestingly, when a GO term was significantly enriched in more than one quantile, it typically displayed a higher fold enrichment and coverage in quantile E (high stoichiometry). For example, the biological process: "translational initiation" and "SRP−dependent cotranslational protein targeting to membrane," as well as the molecular function: "poly (A) RNA binding," were identified in quantile A and quantile E, however, these GO terms displayed a higher fold enrichment in quantile E.

Because the functional analysis indicated an enrichment of RNA binding proteins, we hypothesized that there may be a common protein interaction network. For this, we performed a network analysis, using all uniprot accession identifiers with quantified stoichiometry, queried for known interactions using STRING v10.0 (28, 29), and visualized the interactions with Cytoscape 3.4.0 (30) (**Figure 4-5B**). The value of the highest acetylation stoichiometry was superimposed onto each node given that the high stoichiometry quantile had a higher fold enrichment in our functional annotation analysis (**Figure 4-5A**). Subnetworks were evident from the global interaction map, including RNA binding proteins as well as ribosomal proteins (**Figure 4-5C, D**).

The highly connected ribosomal protein subnetwork displayed many nodes with high acetylation stoichiometry sites including: RPL3, RPL9, RPL10, RPL13, RPS19, RPL23A, RPL35, RPL35A, RPS12, and EIF3D. Ongoing studies are aimed at identifying whether acetylation is a structural component of the ribosome or if acetylation has a dynamic, regulatory role. Interestingly, we previously quantified acetylation stoichiometry in *E. coli* and observed high stoichiometry of ribosomal proteins as well (11) indicating a possible conserved function.

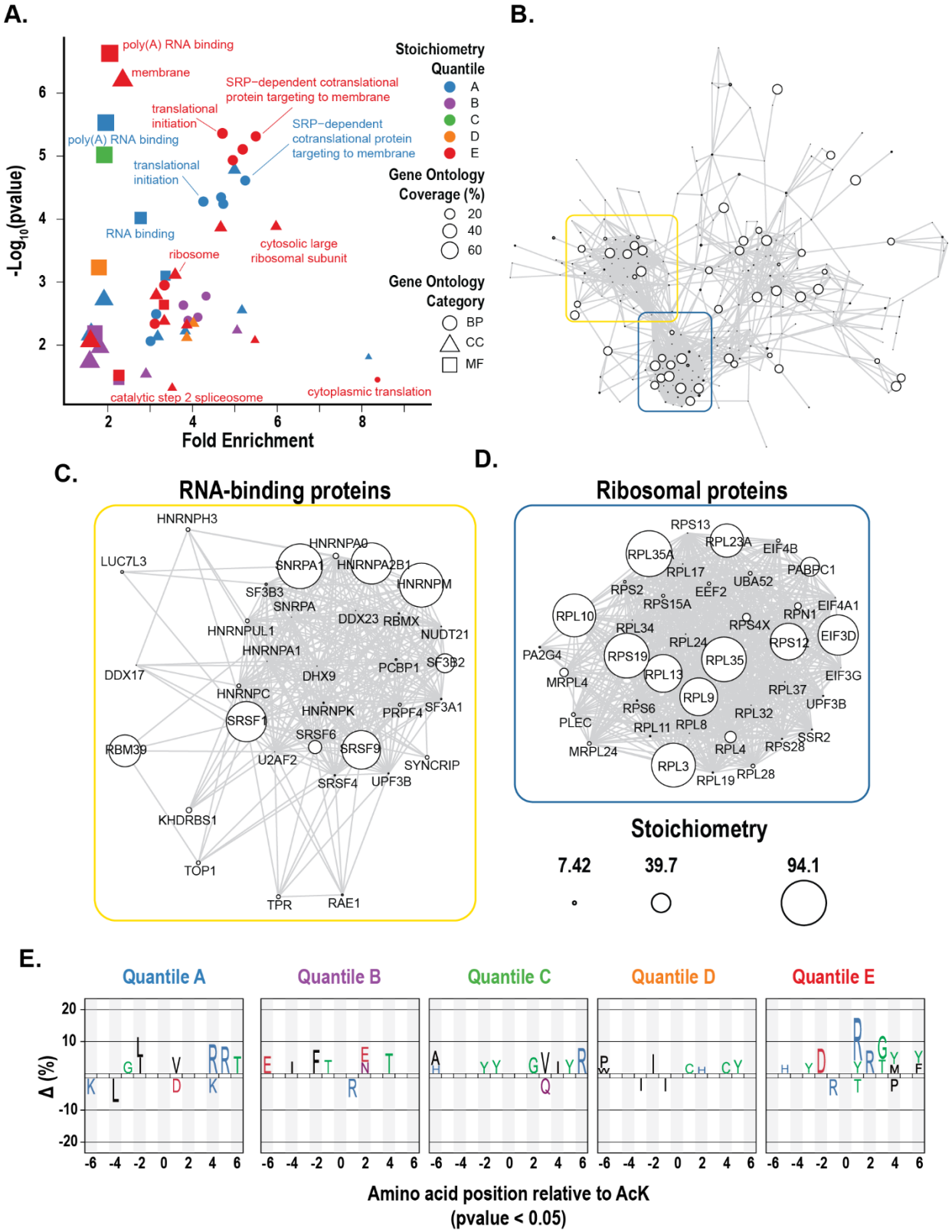
4.4.4.2 Amino acid sequence alignment

We postulated that the high stoichiometry sites may have a unique sequence space surrounding the acetyl-lysine and reasoned that high stoichiometry could be the result of enzyme-targeted acetylation. To address this, we compared the sequence space (± 6 amino acids) around the acetyl-lysine of each of the five quantiles to every acetyl-lysine in the DDA spectral library using iceLogo (17). Indeed, the high stoichiometry quantile displayed a unique amino acid sequence motif than the other four quantiles (**Figure 4-5E**). Interestingly, the sequence motif of high acetylation stoichiometry reflects the Class I acetylation sites identified by Hebert et al. (31) with basic residues at the +1 and +2 positions as well as a tyrosine in the -3 position. In our analysis, we also detected an aspartate residue in the -2 position. The low stoichiometry quantile also displayed a resemblance to the Class II acetylation sites with hydrophobic residues surrounding the acetyl-lysine. In their study, Hebert et al. described Class I sites as acetyl-lysine sites that respond to SIRT3 deacetylation while Class II sites responded to calorie restriction, but not SIRT3 (31). In our study, we measured a single snapshot of acetylation stoichiometry across the cell. However, it is compelling to speculate that the high stoichiometry sites might be functionally relevant and dynamically regulated by acetyltransferases and deacetylases, whether this

mechanism is mediated through Sirtuins or other KDACs. Future studies are aimed at determining dynamic acetylation sites under various conditions.

Figure 4-5: Functional annotation and network analysis reveals RNA binding proteins with high stoichiometry

(A) Global view of the functional annotation analysis of the high-confidence acetyl-lysine site stoichiometry. Site stoichiometry was divided into five equal quantiles (Quantile A: ≤ 0.95%, n = 97; Quantile B: 0.96 – 2.46%, n = 97; Quantile C: 2.47 – 4.64%, n = 96; Quantile D: 4.65 – 27.41%, n = 96; Quantile E: ≥ 27.47, n = 96). Functional annotation was performed using DAVID for each stoichiometry quantiles (15, 16). For enrichment analysis, the background included all uniprot accession identifiers identified in the DDA spectral library. (B) Network analysis of high confidence stoichiometry sites. The size of each node represents the highest acetylation stoichiometry quantified. (C, D) Subnetwork of RNA-binding and ribosomal proteins. (E) Amino acid sequence analysis surrounding the acetyl-lysine site (± 6 amino acids) for each quantile. Analysis was performed using iceLogo (17). For statistical analysis of enrichment, the background included every acetyl-lysine site identified in the DDA spectral library.


A.
poly(A) RNA binding
membrane
SRP−dependent cotranslational protein targeting to membrane
translational initiation
SRP−dependent cotranslational protein targeting to membrane
poly(A) RNA binding
translational initiation
RNA binding
ribosome
cytosolic large ribosomal subunit
catalytic step 2 spliceosome
cytoplasmic translation
-Log10(pvalue)
Fold Enrichment
Stoichiometry Quantile
A
B
C
D
E
Gene Ontology Coverage (%)
20
40
60
Gene Ontology Category
BP
CC
MF
B.
C.
RNA-binding proteins
HNRNPH3
LUC7L3
HNRNPA0
SNRPA1
HNRNPA2B1
HNRNPM
SF3B3
SNRPA
DDX23
RBMX
HNRNPUL1
NUDT21
HNRNPA1
DDX17
DHX9
PCBP1
SF3B2
HNRNPC
SRSF1
HNRNPK
PRPF4
SF3A1
SRSF6
SRSF9
RBM39
U2AF2
SYNCRIP
SRSF4
UPF3B
KHDRBS1
TOP1
TPR
RAE1
D.
Ribosomal proteins
RPS13
RPL23A
EIF4B
RPL35A
RPL17
PABPC1
UBA52
RPS2
EEF2
RPS15A
EIF4A1
RPN1
RPL10
RPL34
RPS4X
EIF3D
RPS12
RPL24
PA2G4
RPS19
RPL35
RPL13
EIF3G
MRPL4
RPL37
RPL9
UPF3B
RPS6
PLEC
RPL32
RPL11
RPL8
SSR2
RPS28
MRPL24
RPL4
RPL3
RPL28
RPL19
Stoichiometry
7.42
39.7
94.1
E.
Quantile A
Quantile B
Quantile C
Quantile D
Quantile E
Δ (%)
Amino acid position relative to AcK
(pvalue < 0.05)

4.4.5 Lysine acetylation derived from cytoplasmic acetyl-Coenzyme A

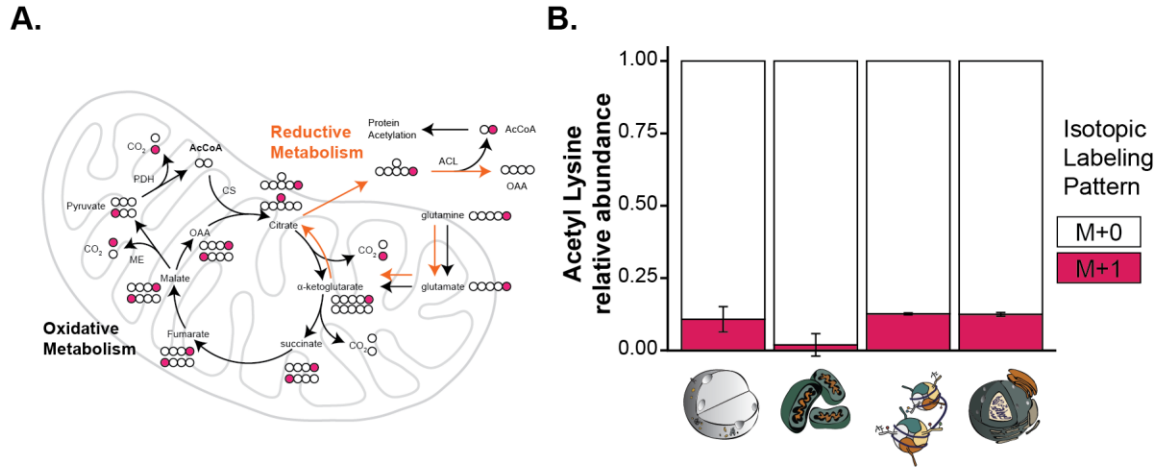
We next asked if mitochondrial proteins can be acetylated outside of mitochondria, whether this occurs prior to import due to translocation across cellular compartments. To address this question, we leveraged the cell's metabolic processes to convert glutamine into cytoplasmic acetyl-CoA through reductive metabolism (32). Using the isotopic tracer, L-glutamine-5-^{13}C, a cytoplasmic specific ^{13}C-acetyl-CoA is generated (**Figure 4-6A**). Using this approach, lysine residues are modified with a ^{13}C-acetyl group in the cytoplasm, which can be quantified using our complimentary MS approaches. If mitochondrial-localized proteins are acetylated in the cytoplasm, an increase in isotopic acetyl incorporation would be observed. First, we utilized our single amino acid quantitation to measure acetyl-lysine across subcellular regions. As expected, cytoplasmic, histone and non-histone pool of proteins represent the majority of ^{13}C-acetyl incorporation (**Figure 4-6B**).

Surprisingly, the mitochondrial pool also displayed a small amount of ^{13}C-acetyl-lysine. To determine whether this observation is due to contamination of non-mitochondrial proteins of if mitochondrial-localized proteins can have ^{13}C-acetyl incorporation, we used a version of the acetylation stoichiometry method, which quantifies based on the acetyl-lysine diagnostic peak during MS/MS fragmentation (12, 20). This approach avoids mixed spectra caused by ^{13}C-amino acid incorporation of newly synthesized proteins since glutamate, proline, aspartate, and asparagine be also be derived from glutamine metabolism (12). After correcting for the natural abundance of ^{13}C-isotope, we quantified acetylation stoichiometry despite the presence of isotopically labeled amino acid (**Figure 4-6C**). While the vast majority of mitochondrial-localized proteins contain no detectable ^{13}C-acetylation, there were a few sites quantified, albeit with low stoichiometry (**Table 4-2**). These results demonstrate that acetylation of mitochondrial proteins

can occur using a cytoplasmic pool of acetyl-CoA. Whether this acetylation event is enzyme catalyzed or nonenzymatic is still unknown.

Figure 4-6: ^{13}C-5-glutamine supplementation generates an isotopic acetyl-CoA pool, which is used to acetylate proteins in the cytoplasm

(A) Metabolic network illustrating the flow of ^{13}C-carbon (orange) derived from ^{13}C-5-L-glutamine through oxidative (black arrows) and reductive (purple arrows) metabolism. The ^{13}C-carbon is expelled as CO_2 via oxidative metabolism through malic enzyme, pyruvate dehydrogenase, and isocitrate dehydrogenase. During reductive metabolism, the ^{13}C-carbon is exported out of mitochondria as citrate to the cytoplasm, where ATP citrate lyase generates ^{13}C-acetyl-CoA. (B) Fraction of acetyl-CoA across subcellular fractions. M+0 is unlabeled acetyl-CoA. M+1 is isotopic acetyl-CoA derived from L-glutamine-5-^{13}C. (C) MS/MS fragmentation spectra of the ribosomal protein L8 peptide containing lysine 60 (RPL8-K60) - GAPLAK$_{Ac}$VVFR. (Left zoom) Immonium ion used to quantify K60 acetylation stoichiometry after correcting for ^{13}C natural abundance isotopic distribution. The measured stoichiometry is 2.9%. (Right zoom) Fragment ion y8 showing a different isotopic distribution than the immonium ion isotopic distribution due to the incorporation of ^{13}C-proline into the peptide.


A.
Reductive Metabolism
Protein Acetylation
AcCoA
ACL
OAA
PDH
CS
Pyruvate
Citrate
glutamine
OAA
ME
Malate
α-ketoglutarate
glutamate
Oxidative Metabolism
Fumarate
succinate
B.
Acetyl Lysine relative abundance
1.00
0.75
0.50
0.25
0.00
Isotopic Labeling Pattern
M+0
M+1

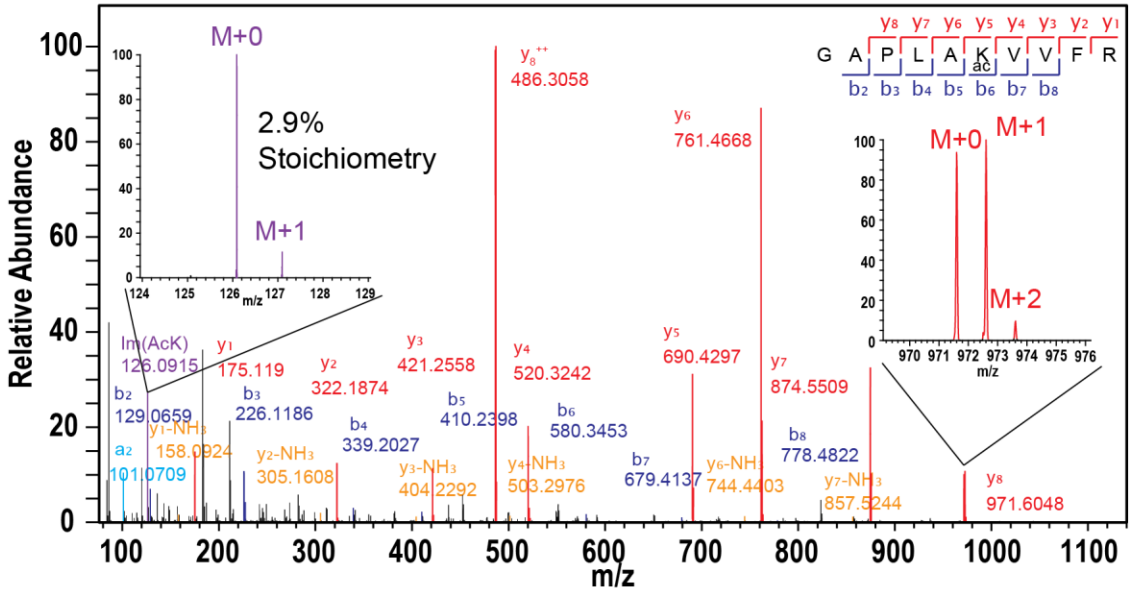

C.
Relative Abundance
m/z
M+0
2.9%
Stoichiometry
M+1
Im(AcK)
126.0915
y1
175.119
y2
322.1874
y3
421.2558
y4
520.3242
y8++
486.3058
y6
761.4668
y5
690.4297
y7
874.5509
b2
129.0659
b3
226.1186
b4
339.2027
b5
410.2398
b6
580.3453
b8
778.4822
b7
679.4137
a2
101.0709
y1-NH3
158.0924
y2-NH3
305.1608
y3-NH3
404.2292
y4-NH3
503.2976
y6-NH3
744.4403
y7-NH3
857.5244
y8
971.6048
G A P L A K V V F R
b2 b3 b4 b5 b6 b7 b8
y8 y7 y6 y5 y4 y3 y2 y1
ac
M+0
M+1
M+2

Table 4-2: List of mitochondrial-localized proteins acetylated with cytoplasmic acetyl-CoA

Replicate	Uniprot Accession	Protein Description	K-site	Stoichiometry
MCF7-2	P10809	60 kDa heat shock protein	K191; K192	0.01377
MCF7-1	Q9UIJ7	GTP:AMP phosphotransferase AK3	K20	0.01299
MCF7-2	P00505	Aspartate aminotransferase	K404	0.01029
MCF7-2	P38117	Electron transfer flavoprotein subunit beta	K19	0.00983
MCF7-2	Q99798	Aconitate hydratase	K520; K521	0.00834
MCF7-1	P10809	60 kDa heat shock protein	K191; K192	0.00740
MCF7-2	P12532	Creatine kinase U-type	K199	0.00464
MCF7-1	P00505	Aspartate aminotransferase	K404	0.00297
MCF7-2	P53007	Tricarboxylate transport protein	K97	0.00288
MCF7-1	P53007	Tricarboxylate transport protein	K97	0.00284
MCF7-2	Q9BXW7	Cat eye syndrome critical region protein 5	K69	0.00223
MCF7-1	P43304	Glycerol-3-phosphate dehydrogenase	K510; K517	0.00207
MCF7-1	P54886	Delta-1-pyrroline-5-carboxylate synthase	K761	0.00145
MCF7-2	Q99798	Aconitate hydratase	K304; K305; K309	0.00038
MCF7-1	P12532	Creatine kinase U-type	K199	0.00036

4.5 Discussion

4.5.1 Acetylation stoichiometry throughout the cell

In this study, we developed an improved method for quantifying acetylation stoichiometry by optimizing MS settings and incorporating an offline, high pH reverse phase prefractionation coupled to DIA-MS/MS. This workflow allows for deeper coverage and quantitation of acetylation stoichiometry for multiple lysine-containing peptides. Using this approach, we determined the acetylation stoichiometry distribution (global and site-specific acetylation stoichiometry) across MCF7 cells, observing a significantly higher proportion of proteins with higher acetylation stoichiometry within the nuclear compartment. Amino acid quantitation from digested subcellular proteomes agreed with the site-specific acetylation distribution. Measuring global PTM stoichiometry of other lysine modifications including mono-, di-, and tri-methylation displayed similar cellular distribution profiles as acetylation. Interestingly, well-characterized post translational modifying enzymes responsible for these modifications reside in the nuclear compartment, highlighting the strong regulatory role of epigenetic and gene expression systems.

4.5.2 Site-specific acetylation stoichiometry

Lysines sites with high stoichiometry were notable throughout the cell. Many of the high stoichiometry sites localized to the cytoplasm were enriched on ribosomal and RNA binding proteins (**Figure 4-5**), an observation that we have also made in *E. coli* (11). Interestingly, Liew et al. demonstrated that ribosomes from rat liver are rapidly acetylated on the N-ε-lysine after intraperitoneal injection with radiolabeled acetate (33). Overlaying our stoichiometry data with the solved structure for the ribosome revealed many lysine sites were within hydrogen bonding distance (data not shown). Ongoing studies are aimed to determine whether acetylation plays a structural role or as a regulatory modification based on the nutritional state of the cell, modulating

protein synthesis. In line with a regulatory function of ribosomal acetylation, in yeast, increased nuclear acetyl-CoA levels are a pro-growth signal for gene activation (34, 35).

High acetylation stoichiometry would be predicted for those lysine sites targeted by KAT enzymes as well as autoacetylation sites directly on KAT enzymes as has been identified previously as a regulatory mechanism (36). In our study, we measured two proteins that have lysine sites targeted by KATs and a lysine residue, which may be a novel autoacetylation event on a recently described lysine acetyltransferase. Glucose-6-phosphate dehydrogenase (G6PD) K386 displayed a stoichiometry of 48%. G6PD serves as the branch point for glucose entering glycolysis or the pentose phosphate pathway and functions as a homodimer with K386 positioned at the dimer interface. Acetylation of this site by the acetyltransferase, KAT9, consequently alters this interaction (37). Acetylation of nearly half of G6PD enzyme population suggests that MCF7 cells are undergoing both pentose phosphate pathway and glycolysis for nucleotide biosynthesis and ATP production, respectively. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which has a well characterized role in glycolysis also has an apoptotic function in the nucleus (38). Nuclear translocation is in part, mediated through acetylation of three lysine residues, K117, K251, and K227 by P300/CBP-associated factor (PCAF) (39). We measured the acetylation of K117 and K251 to be 10% and < 1%, respectively. This low stoichiometry suggests that the majority of GAPDH is not nuclear-localized and serves its canonical function as a glycolytic enzyme. Lastly, another site with high stoichiometry was hnRNPA2 K59 at 88% acetylation. This enzyme was recently shown to be a lysine acetyltransferase, targeting histone H4K8 (40). The authors identified that acetyl-CoA binding and acetyltransferase activity was in part mediated by residues R60 and R62, but these residues were not completely sufficient for KAT activity as a fraction of activity was detectable even in a double mutant. Perhaps lysine 59, which we measured at 86%

stoichiometry is another residue critical for activity and the level of stoichiometry we observed could be the result of autoacetylation.

4.5.3 Higher acetate load on protein than acetyl-Coenzyme A

Acetate is more abundant in the form of covalent lysine modification than the amount conjugated to Coenzyme A (as acetyl-CoA). What role does this form of acetate have on cellular metabolism? Acetate from lysine deacetylation reactions has two immediate fates: free acid and the sirtuin-derived reaction product, O-acetyl-ADP-ribose (OAADPr) (41). The catabolic fate of OAADPr is unclear and warrants future investigation, while the free acid can be reesterified to generate acetyl-CoA by the acetyl-CoA synthetase family of enzymes. Recently, Comerford et al., identified that a major source of acetyl-CoA in cancer cells came from acetate, derived from the histone deacetylation reaction. Generation of acetyl-CoA was dependent on the nucleocytosolic enzyme, acetyl-CoA synthetase 2 (ACSS2) and was highly expressed in cancer tissues (42). MCF7 cells are derived from an adenocarcinoma of mammary glands, thus linking our measured stoichiometry with the role of acetate on tumor metabolism. While we did not detect ACSS2 in our study, probably owing to the stochastic nature of mass spectrometry-based techniques, this enzyme has been detected previously in MCF7 cells at low abundance (43, 44).

4.5.4 Quantifying acetylation through cellular reverse-stoichiometry

In this study, we used a reverse-stoichiometry method, where cells are supplemented an isotopic metabolite that generates the corresponding isotopic acetyl-CoA in vivo and through enzymatic or nonenzymatic processes, lysine residues can incorporate this isotopic acetyl group. During MS sample prep, all unmodified lysines are labeled with the light acetyl group using acetic anhydride. Using our MS workflow, we can quantify the stoichiometry, which originated from the isotopic metabolite. This method will allow us to measure kinetic incorporation of heavy acetyl

groups onto proteins. A similar approach has been utilized to measure acetylation kinetics onto histone proteins (45), however understanding acetylation dynamics across the proteome is still lacking. Determining proteome-wide acetylation kinetics will help decode functional vs. spurious acetylation events and prioritize those regulatory lysine sites.

4.6 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. Baeza, J., Smallegan, M. J., and Denu, J. M. (2016) Mechanisms and Dynamics of Protein Acetylation in Mitochondria. Trends Biochem. Sci. 41, 231–244
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. Baeza, J., Smallegan, M. J., and Denu, J. M. (2015) Site-specific reactivity of nonenzymatic lysine acetylation. ACS Chem. Biol. 10, 122–128
10. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E.

(2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. 43, D512–20

11. 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
12. Fan, J., Baeza, J., and Denu, J. M. (2016) Investigating Histone Acetylation Stoichiometry and Turnover Rate. Methods Enzymol. 574, 125–148
13. 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
14. 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
15. 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
16. 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
17. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009) Improved visualization of protein consensus sequences by iceLogo. Nat. Methods. 6, 786–787
18. 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

19. 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
20. 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
21. 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
22. 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
23. 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
24. 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

25. 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

26. Greer, T., Lietz, C. B., Xiang, F., and Li, L. (2015) Novel isotopic N,N-dimethyl leucine (iDiLeu) 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. Snel, B., Lehmann, G., Bork, P., and Huynen, M. A. (2000) STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. Nucleic Acids Res. 28, 3442–3444

29. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., and von Mering, C. (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 43, D447–52

30. Saito, R., Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., Lotia, S., Pico, A. R., Bader, G. D., and Ideker, T. (2012) A travel guide to Cytoscape plugins. Nat. Methods. 9, 1069–1076

31. Hebert, A. S., Dittenhafer-Reed, K. E., Yu, W., Bailey, D. J., Selen, E. S., Boersma, M. D., Carson, J. J., Tonelli, M., Balloon, A. J., Higbee, A. J., Westphall, M. S., Pagliarini, D. J., Prolla, T. A., Assadi-Porter, F., Roy, S., Denu, J. M., and Coon, J. J. (2013) Calorie restriction and SIRT3 trigger global reprogramming of the mitochondrial protein acetylome. Mol. Cell.

49, 186–199

32. Metallo, C. M., Gameiro, P. A., Bell, E. L., Mattaini, K. R., Yang, J., Hiller, K., Jewell, C. M., Johnson, Z. R., Irvine, D. J., Guarente, L., Kelleher, J. K., Vander Heiden, M. G., Iliopoulos, O., and Stephanopoulos, G. (2011) Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature. 481, 380–384
33. Liew, C. C., and Gornall, A. G. (1973) Acetylation of ribosomal proteins. I. Characterization and properties of rat liver ribosomal proteins. J. Biol. Chem. 248, 977–983
34. Cai, L., Sutter, B. M., Li, B., and Tu, B. P. (2011) Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. Mol. Cell. 42, 426–437
35. Shi, L., and Tu, B. P. (2015) Acetyl-CoA and the regulation of metabolism: mechanisms and consequences. Curr. Opin. Cell Biol. 33, 125–131
36. Albaugh, B. N., Arnold, K. M., Lee, S., and Denu, J. M. (2011) Autoacetylation of the histone acetyltransferase Rtt109. J. Biol. Chem. 286, 24694–24701
37. Wang, Y.-P., Zhou, L.-S., Zhao, Y.-Z., Wang, S.-W., Chen, L.-L., Liu, L.-X., Ling, Z.-Q., Hu, F.-J., Sun, Y.-P., Zhang, J.-Y., Yang, C., Yang, Y., Xiong, Y., Guan, K.-L., and Ye, D. (2014) Regulation of G6PD acetylation by SIRT2 and KAT9 modulates NADPH homeostasis and cell survival during oxidative stress. EMBO J. 33, 1304–1320
38. Chuang, D.-M., Hough, C., and Senatorov, V. V. (2005) Glyceraldehyde-3-phosphate dehydrogenase, apoptosis, and neurodegenerative diseases. Annu. Rev. Pharmacol. Toxicol. 45, 269–290
39. Ventura, M., Mateo, F., Serratosa, J., Salaet, I., Carujo, S., Bachs, O., and Pujol, M. J. (2010) Nuclear translocation of glyceraldehyde-3-phosphate dehydrogenase is regulated by

acetylation. Int. J. Biochem. Cell Biol. 42, 1672–1680

40. Guha, M., Srinivasan, S., Guja, K., Mejia, E., Garcia-Diaz, M., Johnson, F. B., Ruthel, G., Kaufman, B. A., Rappaport, E. F., Glineburg, M. R., Fang, J.-K., Szanto, A. K., Nakagawa, H., Basha, J., Kundu, T., and Avadhani, N. G. (2016) HnRNPA2 is a novel histone acetyltransferase that mediates mitochondrial stress-induced nuclear gene expression. Cell Discov. 2, 16045
41. Tong, L., and Denu, J. M. (2010) Function and metabolism of sirtuin metabolite O-acetyl-ADP-ribose. Biochim. Biophys. Acta. 1804, 1617–1625
42. Comerford, S. A., Huang, Z., Du, X., Wang, Y., Cai, L., Witkiewicz, A. K., Walters, H., Tantawy, M. N., Fu, A., Charles Manning, H., Horton, J. D., Hammer, R. E., McKnight, S. L., and Tu, B. P. (2014) Acetate Dependence of Tumors. Cell. 159, 1591–1602
43. Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A.-K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P.-H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., and Pontén, F. (2015) Tissue-based map of the human proteome. Science. 347, 1260419
44. Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., Wernerus, H., Björling, L., and Ponten, F. (2010) Towards a knowledge-based Human Protein Atlas. Nat. Biotechnol. 28, 1248–1250
45. Evertts, A. G., Zee, B. M., Dimaggio, P. A., Gonzales-Cope, M., Coller, H. A., and Garcia, B. A. (2013) Quantitative dynamics of the link between cellular metabolism and histone

acetylation. J. Biol. Chem. 288, 12142–12151

Chapter 5: Conclusions, ongoing projects, and future directions

5.1 Overall conclusions from research

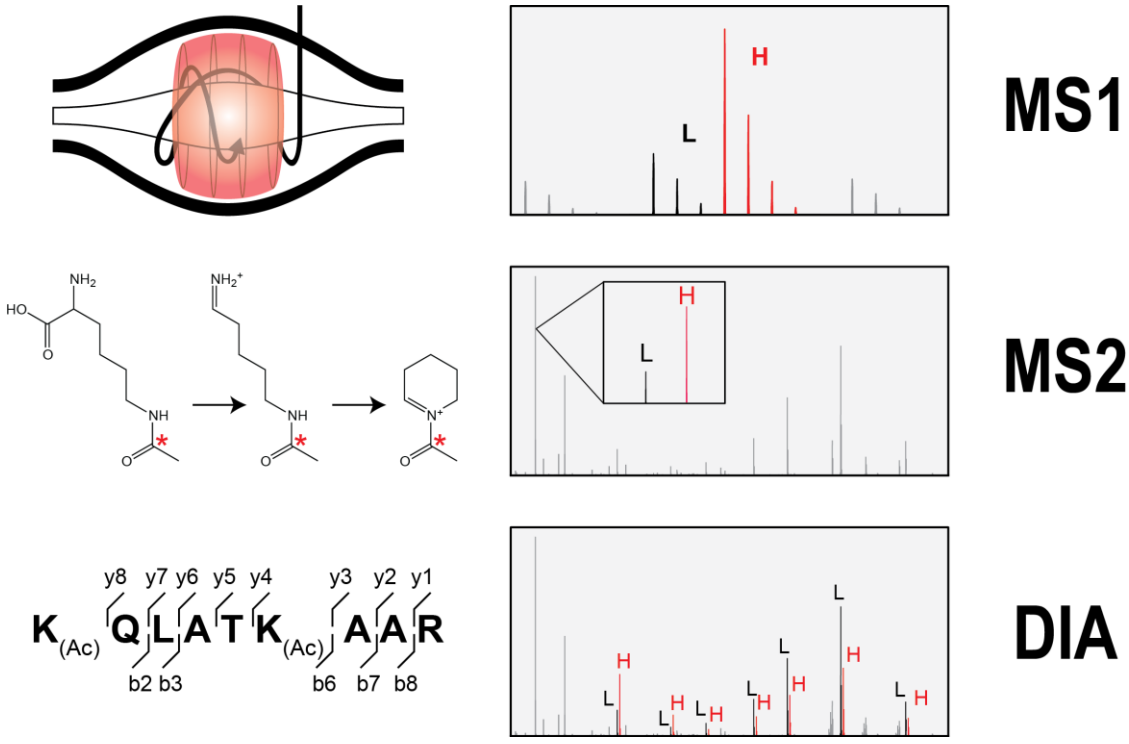
5.1.1 Quantifying acetylation stoichiometry

At the beginning of my PhD thesis, researchers in the field of protein acetylation, including the Denu lab, were asking how to quantify acetylation stoichiometry. This was the time when mass spectrometry was rapidly being utilized to identify and quantify the ever increasing lysine acetylation sites throughout the cell and early functional studies began to paint a picture that acetylation was mostly inhibitory, at least within mitochondria. Therefore, knowing the level of stoichiometry would provide critical information to understand this level of inhibition in cells.

A major component to my thesis work was the development of a mass spectrometry method to quantify lysine acetylation stoichiometry. The challenge in quantifying stoichiometry arises when measuring peptides of different chemical composition due to their varying ionization efficiencies in electrospray ionization-mass spectrometry. Additionally, the widely used digestion method, trypsin, cleaves c-terminal to the unmodified lysine, but does not cleave when acetyl-lysine is present. To circumvent these issues, I devised a chemical labeling approach, which equalizes the ionization efficiency and chemical composition of the modified and unmodified acetyl-lysine peptides. This strategy can involve various acetic anhydride isotopologues to produce an isotopic acetyl-lysine, generating chemically identical light and heavy peptides upon digestion. Throughout my research, I have used different isotopic variants for various forms of quantitation (**Figure 5-1**). After our initial publication, this approach has also been adopted by other groups using different mass spectrometry platforms, highlighting the broad utility and applicability of this method (1, 2).

Figure 5-1: Methods used for quantifying acetylation stoichiometry

Diagramed are the different forms of acetylation stoichiometry quantitation. Top row includes MS1 quantitation which measures the relative abundance of the light and heavy acetyl precursor ion. Middle row outlines an MS2-based method which utilizes the acetyl-lysine immonium ion. Bottom row represents acetylation stoichiometry quantified in data-independent acquisition mode. Here, every acetyl-lysine containing fragment ions are measured across the elution time and used to quantify site-specific acetylation stoichiometry.



5.1.2 Mechanistic understanding of mitochondrial protein acetylation

The field of mitochondrial acetylation was rapidly evolving with the use of immunoenrichment and MS-based methods for detecting and quantifying acetylation (3–5). Investigations by various groups in search for a mitochondrial-localized protein acetyltransferase were underway with no success. Early evidence suggested that lysine acetylation could be the result of a nonenzymatic mechanism especially within the conditions of the mitochondrial matrix: high acetyl-CoA levels and alkaline pH (6, 7). A systematic, analytical approach was needed to validate this hypothesis.

My research was critical in demonstrating that nonenzymatic acetylation was plausible in the context of the cell. There are several factors controlling the level of acetylation in vivo: protein synthesis and degradation rates, rate of acetylation, and the rate of deacetylation. Taking into consideration the protein turnover rates (8) as well as the range of measured second order rate constants of native, purified proteins (9), mathematical estimation of in vivo acetylation corresponds to the measured stoichiometry levels in cells (Chapter 4). Interestingly, not all mitochondrial acetylation sites can be explained by the rates of nonenzymatic acetylation. Interestingly, the nuclear localized acetyltransferase, KAT8 (a.k.a. hMOF and MYST), was recently shown to translocate to mitochondria (10). While the authors demonstrated a functional role for KAT8 in regulating mitochondrial transcription and translation, no lysine sites were identified for its regulatory effect. Interestingly, I quantified 6 lysine sites to have high stoichiometry on proteins involved in the OXPHOS system. The sites are involved in Complex I and Complex V function and/or assembly. It will be interesting to determine whether these sites might be targets of the newly identified acetyltransferase, which translocates to mitochondria.

5.2 Ongoing work

My thesis research as well as prior studies quantifying (or estimating) mammalian acetylation stoichiometry report overall low levels, especially in mitochondria (1, 11). While it might appear reasonable to label this modification as spurious, or nonfunctional due to low stoichiometry, much work is to be done to obtain a broad understanding of this modification. This point is highlighted by the very limited biological conditions used to quantify acetylation stoichiometry. For example, stoichiometry was quantified from MCF7 (Chapter 4) and Hela cells (1). These cells are highly proliferative cells grown with high glucose and optimal growth factors. Additionally, stoichiometry estimation from mouse liver mitochondria were performed using young mice ranging from 1 - 16 months fed an ad libitum diet. To understand protein acetylation in the context of health and disease, we must investigate this modification using a broader set of biological conditions. Ongoing projects in the Denu Lab are aimed with this focus in mind. Below, I outline collaborative projects that span aging, heart failure, and neurodegeneration.

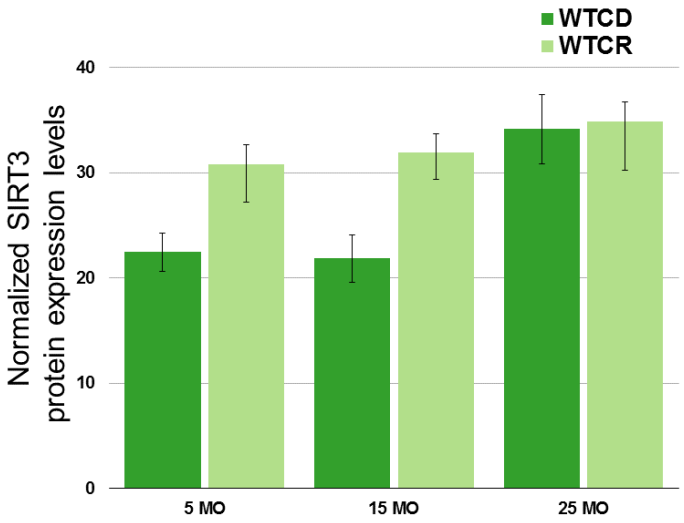
5.2.1 A longitudinal study understanding effects of SIRT3 and calorie restriction on mitochondrial protein acetylation

Major progress has been made by Kristin Dittenhafer-Reed, a previous graduate student in the Denu Lab in collaboration with Tom Prolla, Josh Coon, and Dave Pagliarini's research groups, to understand the role of SIRT3, the NAD^+-dependent mitochondrial deacetylase, and calorie restriction (CR) on mitochondrial protein acetylation (12, 13). To understand the longitudinal effects of calorie restriction and SIRT3, Rashpal Dhillon, a postdoctoral fellow in our lab, has systematically measured different components of the electron transport system across four different tissues in an aging mouse model. His work has revealed striking age-, diet-, and tissue-dependent effects on the activity of the oxidative phosphorylation system including complex I and

complex II activity. While these studies establish a link between SIRT3, diet, and mitochondrial enzyme function, little is known of how acetylation levels change with age and its possible functional consequences. Ongoing work in the Denu lab is aimed at quantifying the level of acetylation stoichiometry in this complex aging mouse model to understand how lifestyle (including diet and exercise) affects acetylation with age. It is interesting to note, that most acetylomic studies thus far have utilized young and middle-aged mouse model systems. Perhaps the enzymatic activity of SIRT3 in regulating mitochondrial function is more critical as we age when other protein homeostasis systems begin to decline. In line with this hypothesis, SIRT3 protein levels are induced with CR in young and middle age mice, however display an elevated level in both a control diet and CR in older mice, suggesting that SIRT3 is needed at an older age (**Figure 5-2**).

Figure 5-2 Diet induced SIRT3 protein levels in an aging mouse liver mitochondria

SIRT3 protein levels were quantified by immunoaffinity capillary electrophoresis (ICE) using WES, the automated western immunoassay system by Protein Simple. Quantified values represent the SIRT3 peak area normalized to total protein levels in each condition. Data is courtesy of Rashpal Dhillon.



5.2.2 The cardiac mitochondrial acetylome en route to heart failure

Cardiomyocytes rely on fatty acid oxidation to meet its massive energy demands. As such, mutations in genes involved in fatty acid oxidation as well as the OXPHOS system are causative for cardiomyopathies (14). Dr. Dan Kelly's lab at Sanford Burnham Prebys Medical Discovery Institute is interested in the early events prior to heart failure (HF). His lab has shown that the mouse failing heart undergoes a reprogramming of metabolic fuel from fatty acids to ketone bodies to meet its energy demands and is accompanied with elevated mitochondrial acetyl-CoA levels (suggested by the increase in acetyl-carnitine) (15, 16). Additionally, failing heart also displays protein hyperacetylation (16). To understand the changes of mitochondrial protein acetylation en route towards heart failure, we will use our quantitative acetyl-proteomic approach to quantify acetylation stoichiometry. Identifying key acetylation sites may help elucidate metabolic derangements and molecular mechanism leading to heart failure and unravel the complex etiology of heart failure, which affects about 5.7 million Americans (www.cdc.gov).

5.2.3 Dynamic acetylation in a neurodegenerative disease model

Studies in the Puglielli lab in the Department of Medicine at UW-Madison have linked protein acetylation in the endoplasmic reticulum (ER) with neurodegenerative disease (17). The Puglielli lab has identified two ER-resident acetyltransferases, ATase1 and ATase2, which function to modify correctly folded proteins as proteins traffic through the secretory pathway. Misfolded proteins not acetylated by the ER acetyltransferases are degraded (18–20). Acetyl-CoA flux to the ER by the acetyl-CoA transporter, AT-1, is required for protein acetylation in the ER (19). Genetic studies have revealed a missense, point mutation of AT-1 (S116R), is associated with a familial form of spastic paraplegia and heterozygous mice display a form of neurodegeneration.

In collaboration with the Puglielli lab, we aim to compare global changes in acetylation stoichiometry using two AT-1 mouse models, an AT-1 knockout and AT-1 overexpressing mouse models. This work is in collaboration with Inca Dietrich, a graduate student in the Puglielli lab. She has performed subcellular fractionation on liver samples to enrich for nuclear, mitochondrial, and cytoplasmic fractions. I will perform the stoichiometry and bioinformatic analysis of the mass spectrometry findings. I am hopeful that results from this collaboration will shed light into the role of mitochondrial protein acetylation within the endoplasmic reticulum. In line with this collaborative work, I have quantified proteins localized to the ER with high stoichiometry (Chapter 4).

5.3 Future directions

5.3.1 Stoichiometry of acylation

The technological advancements of mass spectrometry have enabled the identification of numerous lysine post-translational modifications including acetylation, butyrylation, propionylation, malonylation, succinylation, glutarylation, etc (21, 22). As these modifications likely arise from their corresponding acyl-CoA derivatives, the list of potential modifications has not reached its limit, given that many other acyl-CoAs are present (For more information, see Chapter 1). We now face the daunting task to quantify stoichiometry of combined acylation. As a first step, Beatriz Camacho, a graduate student in the Denu lab, is working to quantify the global stoichiometry of acylation by digesting whole proteomes down to single amino acids (23). This method will simplify the proteomic sample and allow the quantification of these acyl-lysine modifications. Using an external calibration curve, Beatriz will be able to calculate acylation stoichiometry in various subcellular fractions.

5.3.2 Regulation of translation by reversible acetylation

My thesis work reveals high stoichiometry of ribosomal proteins in bacteria (Chapter 2) and mammalian systems (Chapter 4) likely representing a conserved function for acetylation on translational systems. To investigate the functional role of acetylation on translation, I propose to use global acetylation stoichiometry profiling as well as in vitro translation systems to study how protein acetylation affects ribosomal function. A collaborator at UT Southwestern, Benjamin Tu, has characterized the “Yeast Metabolic Cycle” in a continuous, nutrient limited, culture system (**Figure 5-3** and Ref. (24)). Benjamin Tu’s work reveals that yeast in batch culture undergo synchronous metabolic and growth oscillations when switched to nutrient-limited conditions. This system is well adept to provide a homogenous population of yeast undergoing synchronized

regulatory mechanisms, including translational regulation. If acetylation functions as a regulatory modification, I would expect to quantify oscillations in acetylation stoichiometry, much like the gene expression oscillations.

To determine the role of acetylation on ribosomal function, ribosomes can be purified at different timepoints throughout the YMC and used to measure in vitro translation rates using a fluorescent reporter gene construct encoding SNAP enzyme (25), which is an O6-alkylguanine DNA O6-alkyltransferase. When SNAP is reacted with a quenched fluorogenic guanine substrate, it produces a covalently linked fluorescent-SNAP molecule. Incubating purified ribosomes with the SNAP mRNA, translation factors, and a quenched fluorogenic guanine can be used to monitor the production of the SNAP enzyme by measuring the increase in fluorescence over time (**Figure 5-4**). Treating purified ribosomes with deacetylases and acetyltransferases to alter acetylation levels, the effect of acetylation on ribosome function can be determined.

Figure 5-3 Temporal regulation of gene expression of the Yeast Metabolic Cycle.

Yeast undergo metabolic oscillations during continuous, nutrient-limited conditions as depicted by dO_2 levels. (Figure is from Tu et al. (24))

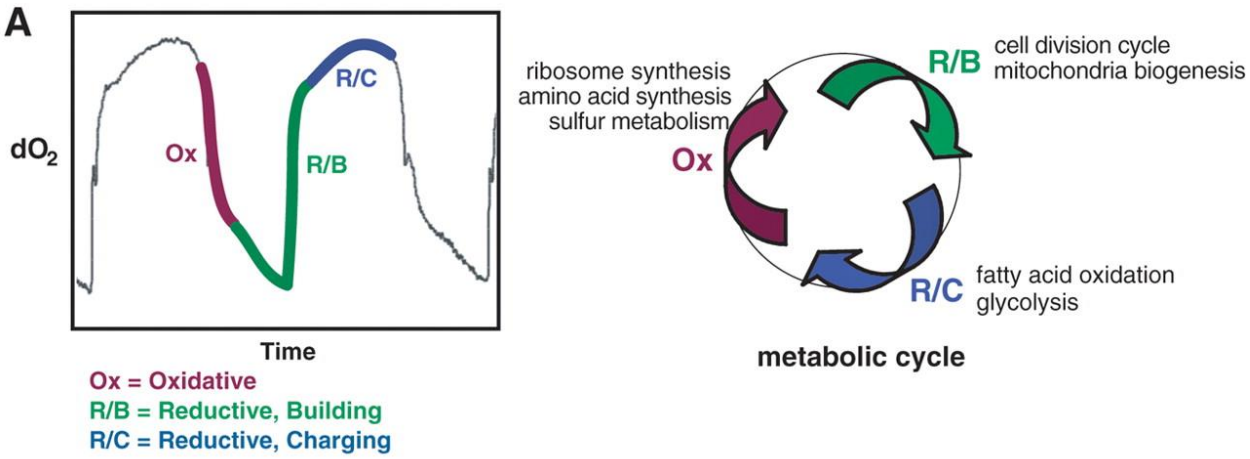
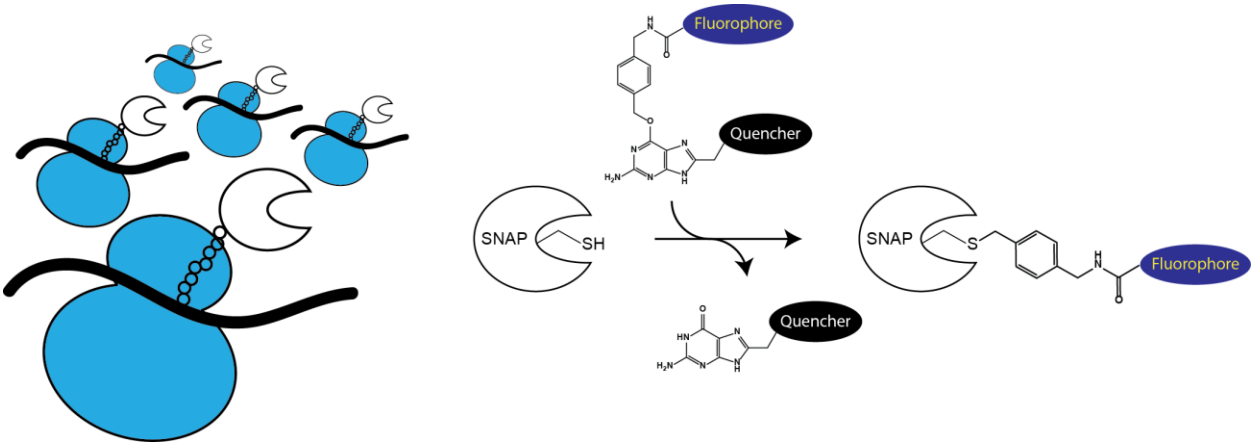


Figure 5-4 Ribosomal activity profiling

SNAP mRNA is translated by ribosomes purified at different timepoints throughout the Yeast Metabolic Cycle. Active SNAP enzyme reacts with a quenched fluorogenic guanine to produce a fluorescently labeled SNAP protein. Change in fluorescence is proportional to ribosome translation rates (Figure adapted from Capece et al. (25)).



5.3.3 Deciphering between enzymatic and non-enzymatic acetylation

A major challenge still faced is in differentiating between functional and spurious acetylation events. If enzyme-catalyzed acetylation represents targeted, functional acetylation and nonenzymatic representing spurious acetylation events (though some nonenzymatic acetylation may be functional), then measuring and comparing cellular acetylation kinetics to nonenzymatic acetylation (second order rate constants) will likely distinguish between the two class of sites. Cellular reverse-stoichiometry is a method, which can be used to determine cellular acetylation kinetics (Chapter 4). [U-^{13}C] glucose supplementation will generate $^{13}C_2$-acetyl-CoA, which can be used by acetyltransferases for lysine $^{13}C_2$-acetylation. Determining cellular acetylation kinetics is accomplished by quantifying the change in stoichiometry over time for the cellular acetylome as has been done previously for histone acetylation (26). To do this, cultured cells will be grown in RPMI media (to supplement with essential and nonessential amino acids) and serum starved for a short length of time. At the beginning of the experiment, cells will be supplemented with growth factors (serum) as well [U-^{13}C] glucose. Cells will be harvested at 10, 20, 40, 80, 160 minute time points, and processed for acetylation stoichiometry by cellular reverse-stoichiometry. As a comparison, cells without growth factors will be used. Acetylation kinetics will be fit using a differential equations based model (23). Enzyme-catalyzed acetylation kinetics is expected to increase rapidly and reach a maximum. Since the time course is done at a relatively short time scale (~3 hours), nonenzymatic acetylation is expected to be much slower.

5.4 Impact

It has been such an exciting time to be part of the field of protein acetylation. This field has rapidly evolved during this time owing to the technical advancements in mass spectrometry and immunoaffinity enrichment methods. During my time in graduate school, I witnessed as reports of novel acylation modifications were discovered (27–30) and linked to metabolic regulation. Further understanding of the combinatorial effects of multi acyl modifications, regulatory crosstalk, and functional impact will help elucidate the complex mitochondrial regulatory mechanisms. With the numerous fascinating biological questions to address in protein acetylation, I look forward towards the significant progress in the future.

5.5 References

1. 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
2. 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
3. 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
4. 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
5. Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H., Li, Y., Shi, J., An, W., Hancock, S. M., He, F., Qin, L., Chin, J., Yang, P., Chen, X., Lei, Q., Xiong, Y., and Guan, K.-L. (2010) Regulation of cellular metabolism by protein lysine acetylation. Science. 327, 1000–1004
6. Paik, W. K., Pearson, D., Lee, H. W., and Kim, S. (1970) Nonenzymatic acetylation of histones with acetyl-CoA. Biochim. Biophys. Acta. 213, 513–522
7. 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
8. Kim, T.-Y., Wang, D., Kim, A. K., Lau, E., Lin, A. J., Liem, D. A., Zhang, J., Zong, N. C.,

Lam, M. P. Y., and Ping, P. (2012) Metabolic labeling reveals proteome dynamics of mouse mitochondria. Mol. Cell. Proteomics. 11, 1586–1594

9. Baeza, J., Smallegan, M. J., and Denu, J. M. (2015) Site-specific reactivity of nonenzymatic lysine acetylation. ACS Chem. Biol. 10, 122–128
10. Chatterjee, A., Seyfferth, J., Lucci, J., Gilsbach, R., Preissl, S., Böttinger, L., Mårtensson, C. U., Panhale, A., Stehle, T., Kretz, O., Sahyoun, A. H., Avilov, S., Eimer, S., Hein, L., Pfanner, N., Becker, T., and Akhtar, A. (2016) MOF Acetyl Transferase Regulates Transcription and Respiration in Mitochondria. Cell. 167, 722–738.e23
11. 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
12. Hebert, A. S., Dittenhafer-Reed, K. E., Yu, W., Bailey, D. J., Selen, E. S., Boersma, M. D., Carson, J. J., Tonelli, M., Balloon, A. J., Higbee, A. J., Westphall, M. S., Pagliarini, D. J., Prolla, T. A., Assadi-Porter, F., Roy, S., Denu, J. M., and Coon, J. J. (2013) Calorie restriction and SIRT3 trigger global reprogramming of the mitochondrial protein acetylome. Mol. Cell. 49, 186–199
13. 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
14. Kelly, D. P., and Strauss, A. W. (1994) Inherited Cardiomyopathies. N. Engl. J. Med. 330, 913–919
15. Aubert, G., Martin, O. J., Horton, J. L., Lai, L., Vega, R. B., Leone, T. C., Koves, T., Gardell, S. J., Krüger, M., Hoppel, C. L., Lewandowski, E. D., Crawford, P. A., Muoio, D. M., and

Kelly, D. P. (2016) The Failing Heart Relies on Ketone Bodies as a Fuel. Circulation. 133, 698–705

16. Horton, J. L., Martin, O. J., Lai, L., Riley, N. M., Richards, A. L., Vega, R. B., Leone, T. C., Pagliarini, D. J., Muoio, D. M., Bedi, K. C., Jr, Margulies, K. B., Coon, J. J., and Kelly, D. P. (2016) Mitochondrial protein hyperacetylation in the failing heart. JCI Insight. 10.1172/jci.insight.84897
17. Peng, Y., Li, M., Clarkson, B. D., Pehar, M., Lao, P. J., Hillmer, A. T., Barnhart, T. E., Christian, B. T., Mitchell, H. A., Bendlin, B. B., Sandor, M., and Puglielli, L. (2014) Deficient import of acetyl-CoA into the ER lumen causes neurodegeneration and propensity to infections, inflammation, and cancer. J. Neurosci. 34, 6772–6789
18. Costantini, C., Ko, M. H., Jonas, M. C., and Puglielli, L. (2007) A reversible form of lysine acetylation in the ER and Golgi lumen controls the molecular stabilization of BACE1. Biochem. J. 407, 383–395
19. Ko, M. H., and Puglielli, L. (2009) Two endoplasmic reticulum (ER)/ER Golgi intermediate compartment-based lysine acetyltransferases post-translationally regulate BACE1 levels. J. Biol. Chem. 284, 2482–2492
20. Ding, Y., Dellisanti, C. D., Ko, M. H., Czajkowski, C., and Puglielli, L. (2014) The endoplasmic reticulum-based acetyltransferases, ATase1 and ATase2, associate with the oligosaccharyl-transferase to acetylate correctly folded polypeptides. J. Biol. Chem. 10.1074/jbc.M114.585547
21. Baeza, J., Smallegan, M. J., and Denu, J. M. (2016) Mechanisms and Dynamics of Protein Acetylation in Mitochondria. Trends Biochem. Sci. 41, 231–244
22. Choudhary, C., and Mann, M. (2010) Decoding signalling networks by mass spectrometry-

based proteomics. Nat. Rev. Mol. Cell Biol. 11, 427–439

23. Fan, J., Baeza, J., and Denu, J. M. (2016) Investigating Histone Acetylation Stoichiometry and Turnover Rate. Methods Enzymol. 574, 125–148
24. Tu, B. P., Kudlicki, A., Rowicka, M., and McKnight, S. L. (2005) Logic of the Yeast Metabolic Cycle: Temporal Compartmentalization of Cellular Processes. Science. 310, 1152–1158
25. Capece, M. C., Kornberg, G. L., Petrov, A., and Puglisi, J. D. (2015) A simple real-time assay for in vitro translation. RNA. 21, 296–305
26. Evertts, A. G., Zee, B. M., Dimaggio, P. A., Gonzales-Cope, M., Coller, H. A., and Garcia, B. A. (2013) Quantitative dynamics of the link between cellular metabolism and histone acetylation. J. Biol. Chem. 288, 12142–12151
27. Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y., and Zhao, Y. (2011) Identification of lysine succinylation as a new post-translational modification. Nat. Chem. Biol. 7, 58–63
28. Peng, C., Lu, Z., Xie, Z., Cheng, Z., Chen, Y., Tan, M., Luo, H., Zhang, Y., He, W., Yang, K., Zwaans, B. M. M., Tishkoff, D., Ho, L., Lombard, D., He, T.-C., Dai, J., Verdin, E., Ye, Y., and Zhao, Y. (2011) The first identification of lysine malonylation substrates and its regulatory enzyme. Mol. Cell. Proteomics. 10, M111.012658
29. Tan, M., Peng, C., Anderson, K. A., Chhoy, P., Xie, Z., Dai, L., Park, J., Chen, Y., Huang, H., Zhang, Y., Ro, J., Wagner, G. R., Green, M. F., Madsen, A. S., Schmiesing, J., Peterson, B. S., Xu, G., Ilkayeva, O. R., Muehlbauer, M. J., Braulke, T., Mühlhausen, C., Backos, D. S., Olsen, C. A., McGuire, P. J., Pletcher, S. D., Lombard, D. B., Hirschey, M. D., and Zhao, Y. (2014) Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. Cell Metab. 19, 605–617

30. Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., Lu, Z., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B., and Zhao, Y. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell. 146, 1016–1028