Current Trends in Microbiology
Vol. 16, 2022

Review

Protein N-terminal acylation: An emerging field in bacterial cell physiology

Anastacia R. Parks and Jorge C. Escalante-Semerena*
Department of Microbiology, University of Georgia, Athens, GA 30606, USA.

ABSTRACT

N-terminal (Nt)-acylation is the irreversible addition of an acyl moiety to the terminal alpha amino group of a peptide chain. This type of modification alters the nature of the N terminus, which can interfere with the function of the modified protein by disrupting protein interactions, function, localization, degradation, hydrophobicity, or charge. Nt acylation is found in all domains of life and is a highly common occurrence in eukaryotic cells. However, in prokaryotes very few cases of Nt acylation have been reported. It was once thought that Nt acylation of proteins, other than ribosomal proteins, was uncommon in prokaryotes, but recent evidence suggests that this modification may be more common than once realized. In this review, we discuss what is known about prokaryotic Nt acetylation and the acetyltransferases that are responsible, as well as recent advancements in this field and currently used methods to study Nt acetylation.

KEYWORDS: N-terminal acetylation, Nα-acetylation, post-translational modification, modulation of protein function, protein degradation.

CONTENTS

*Corresponding author: jcescala@uga.edu

Introduction

Post-translational modifications (PTMs) allow for rapid changes in protein function in response to exogenous or endogenous stimuli. PTMs involve the addition or removal of chemical moieties to biological molecules, as well as can be proteolytic cleavage. A large variety of protein PTMs have been identified that include glycosylation, lipidation, S-nitrosylation, adenylylation, acylation, ubiquitinylation, methylation, phosphorylation, ADP-ribosylation, and SUMOylation, to name

a few. Here we focus on the PTM of protein acetylation, specifically Nt acetylation.

Acyl moieties can be added to amino groups associated with the epsilon carbon of lysyl side chains of proteins, as well as the alpha amino groups on Nt residues of proteins (Figure 1). Nε Acetyltransferases are enzymes that catalyze the transfer of acyl moieties from acyl-CoA to the epsilon amino group of lysyl residues of proteins or small molecules. This type of acetylation has been characterized in all domains of life [1-3].

Nε acetylation was first reported in eukaryotes in the context of gene silencing, where acetyl groups are added to Nε amino groups of lysyl residues of histone proteins, neutralizing the positive charge of the Nε amino groups, resulting in loss of interactions with the negative charges of the phosphate backbone of DNA. Hence, acetylation of the Nε amino group of lysine relaxes the DNA wrapped around histone proteins, which gives RNA polymerase and other proteins involved in gene expression access to previously inaccessible DNA [4].

The field of lysine acetylation in the context of bacterial metabolism and physiology began with studies of the acetyl-CoA synthetase (Acs) enzyme,

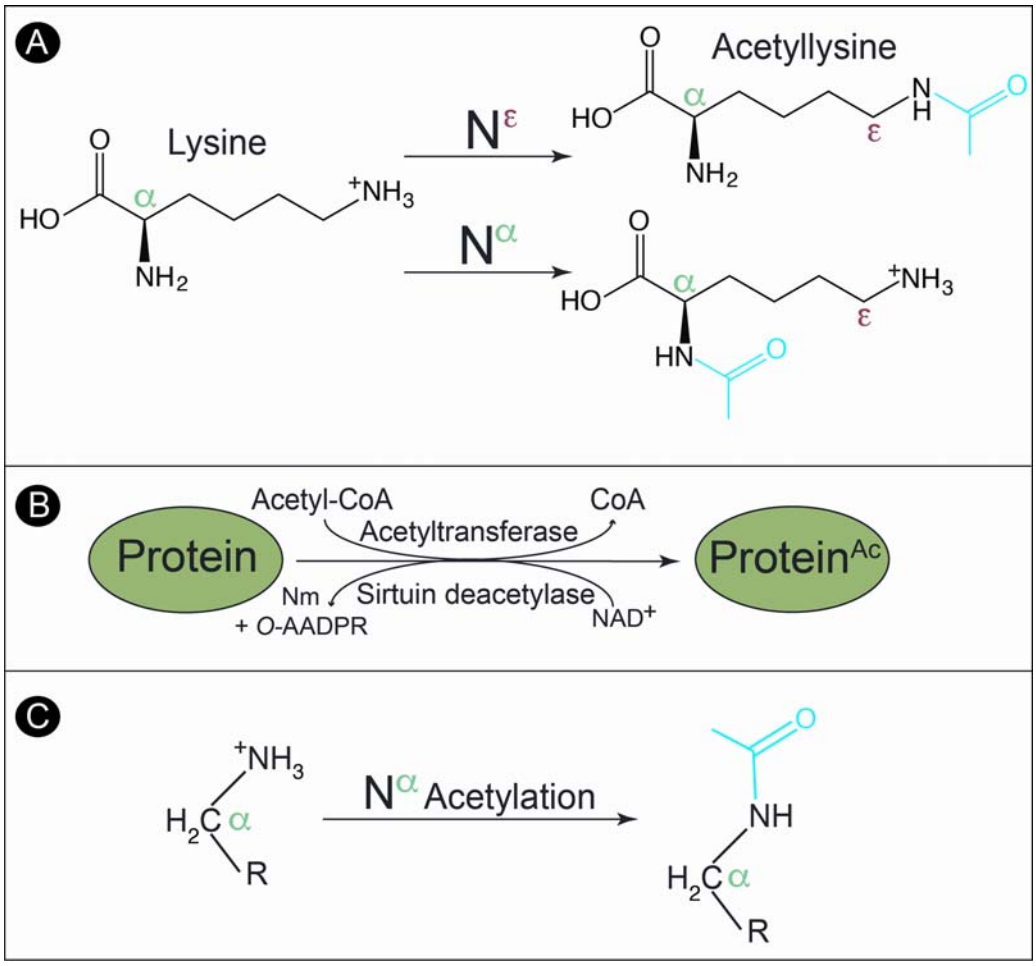

Figure 1. Schematic of different modes of acetylation. **A.** The amino acid lysine can be acetylated on two different amine groups; the amine associated with the alpha carbon (teal, α) or the amine associated with the epsilon carbon (orange, ε). **B**. Depiction of reversible lysine acetylation (RLA) of proteins, which involves acetylation of a protein at a lysyl residue by an acetyltransferase and deacetylation by a deacetylase. **C**. Example of alpha acetylation of an amine of a small molecule or the N-terminal amine of a peptide chain.

which is acetylated on its active-site lysine by the **p**rotein **a**cetyl**t**ranferase (Pat), effectively blocking the first half reaction catalyzed by Acs, *i.e.,* the conversion of acetate to acetyl-AMP [1]. The acetyl moiety on Acs^{Ac} is removed by the NAD^+-dependent CobB sirtuin deacylase, restoring Acs activity [5]. Reversible lysine acetylation (RLA) of Acs appears to be controlled by the energy charge and carbon status of the cell. Acs appears to be acetylated (*i.e.,* inactivated) when the carbon status is poor, that is, when Ac-CoA-consuming processes (*e.g.,* lipid synthesis, anabolism, etc) slow down resulting in elevated levels of Ac-CoA. When the cell becomes active again, Ac-CoA levels go down and Acs^{Ac} is deacetylated (*i.e.,* activated) and so it can synthesize more Ac-CoA to satisfy the anabolic needs of the cell. Similarly, if the cell is not metabolically active, the NAD^+:NADH ratio is <1, reflecting on a slow-down of the flux of electrons through the electron transport chain with the concomitant reduction in the proton motive force (PMF) that fuels the synthesis of ATP by the membrane-bound ATPase; a low PMF means reduced levels of ATP. In sum, high levels of Ac-CoA (reflects carbon status) deactivate Acs, whilst high levels of NAD^+ (reflects energy status) reactivate Acs. For a comprehensive review on lysine acetylation in bacteria, refer to the following references [6-8].

Some acetyltransferases can acetylate primary amino groups of small molecules, which can serve a variety of functions for the cell. The first characterizations of small molecule acetylation involved bacterial acetylation of antibiotic compounds, which resulted in antibiotic resistance [9]. Since this discovery, a wide range of small molecules have been shown to be acetylated, which impacts a variety of cellular processes such as detoxification, toxin-antitoxin systems, anabolism of specific compounds, and transcription to name a few. For a recent review of small molecule acetylation in bacteria, see [10].

In this review we discuss current findings on Nt acetylation of proteins in bacteria, and address methodologies for the study of Nt acetylation and their limitations.

Nt acetylation involves the addition of an acetyl group to the alpha amino group of the N-terminal residue of a peptide chain. This type of modification is ubiquitous, with ~60-90% of proteins in lower and higher eukaryotes (including humans) being acetylated on their N-termini [11-13]. In stark contrast, Nt acetylation in bacteria has been studied mostly in the context of the ribosome, where some of the proteins are N-terminally acetylated by the Rim family acetyltransferases (discussed below). More recently, N-terminomics studies in bacteria show that ~3-10% of total cellular proteins are acetylated at the N terminus [14, 15].

Role of Nt acetylation in eukaryotes

Hypotheses including the Ac/N-End rule

Even though the occurrence of Nt acetylation appears to be widespread, the physiological purpose it serves remains under debate. Some studies suggest that N-terminal acetylation of proteins in eukaryotes provides stability to proteins and prevents protein degradation [16, 17]. In contrast, other studies characterized the Ac/N-End rule pathway, suggesting that Nt acetylation acts as a degradation signal to target Nt acetylated proteins to be labeled with ubiquitin by ubiquitin-ligase Doa10, which commits the ubiquitin-labeled proteins for proteasomal degradation [18]. In *Sacchromyces cerevisiae*, some proteins that have P2'-residues of Ala, Val, Ser, Thr, and Cys can be Nt acetylated, thus termed Ac/N-degrons, and committed to the Ac/N-End rule degradation pathway if the acetylated N-terminus is accessible to be recognized by proteins termed Ac/N-recognins, while proteins acetylated in the initiator methionine can either be committed to the Ac/N-end rule pathway or routed through a separate degradation pathway termed Arg/N-end rule pathway, depending on the hydrophobicity of the protein's P2 residue [19, 20]. A large caveat with these hypotheses is that only a few proteins have been verified to be a part of the aforementioned pathways, begging the question of whether the Ac/N-End rule pathway is relevant to most proteins or to just the few that have been found to be a part of the protein degradation pathways. A second caveat to the Ac/N-end rule is that it has been fully characterized in yeast, but has not been expanded to other eukaryotic organisms.

Other N-terminal modifications to consider

Formylation, deformylation, and initiator methionine cleavage

As mentioned above, Nt modification occurs in all domains of life and is one of the most common protein modifications, especially in eukaryotic cells. In eukaryotes, mRNA translation in the cytosol uses free methionyl-tRNA (Met-tRNA) to begin translation, whilst in mitochondria, eukaryotic plastids, and bacteria N-formylated methionyl-tRNA (fMet-tRNA) is used to initiate translation. In these organelles and cells, the methionyl-$tRNA_f^{Met}$ formyltransferase esterifies methionine to the 3' end of the initiator tRNA [21, 22]. The N-formyl group on iMet can be retained after translation or can be removed rapidly after translation begins by the essential enzyme peptide deformylase (PDF) [23-26]. After deformylation of iMet, the iMet residue usually, but not always, is cleaved from mature proteins through catalysis by methionine aminopeptidase (MAP), *via* a process called N-terminal methionine excision or NME, which occurs in eukaryotes and prokaryotes and is thought to process 55-70% total cellular proteins depending on the organism [27]. The ability of the PDF enzyme to deformylate iMet is not influenced by the nature of amino acid residues occurring after iMet [28]. In contrast, substrate preference of the MAP enzyme seems influenced by the size and nature of the second residue, with a preference for residues containing small side chains, such as Gly, Ala, Ser, Thr, Cys, Pro, or Val [29-31]. Notably, the MAP enzyme cannot excise iMet that retains its formyl group [24], and it has decreased catalytic activity *in vitro* on methionine residues that have been oxidized to methionine sulfoxide [29]. Most importantly, the physiological function of NME is controversial, with some groups arguing that NME provides protein stability, while others favor the idea of NME triggering protein degradation by providing substrates for the N-End Rule (NER), which is a protein degradation process that uses specific N-terminal amino acid residues as signals called N-degrons [32-34]. This hypothesis however, is less applicable to bacteria where the NER system differs considerably as compared to eukaryotes (for a review on N-end rule see [35]).

In the case of Rim acetyltransferases, NME is required for the acetylation of Ser or Ala if they are the second residue (P2) of the protein, as discussed under section *Rim acetyltransferases*.

A ground-breaking study for N-terminal proteomes was conducted by Bonissone *et al*. [36]. This group of investigators analyzed proteomics data that included mass spectrometry data and a comparative genomics approach of yeast, mammals and 112 million spectra from 57 bacterial species. In this study, the authors challenged assumptions in the field surrounding the physiological role and specificity of the first three amino acids of proteins and their involvement in NME by comparing datasets across all domains of life. Methionine aminopeptidases (MAPs) remove the initiator methionine (P1) in proteins that have either Gly, Ala, Ser, Pro, Val, Thr, or Cys at position P2, yielding proteins with either one of these residues as their new N-terminus after cleavage. Bonissone *et al*. found elevated levels of conservation of only two of the seven MAP-targeted position 2 (P2) residues, Ser or Ala, in a diverse array of prokaryotic and eukaryotic species, suggesting an evolutionary pressure for Ser or Ala residues to be conserved at the P2 position. This finding contrasts with the previous observation by Arfin and Bradshaw [37], who first noticed that the specific target residues of NME directly correlated with the exact residues that were considered "stabilizing" amino-terminal residues" (*i.e.,* Gly, Ala, Ser, Cys, Thr, Pro, and Val), which were found on proteins with longer half-lives than proteins containing terminal residues of "destabilizing" residues (*i.e.,* Glu, Gln, Asp, Asn, Ile, Leu, Phe, Trp, Tyr, His, Lys, Arg) [34, 38]. Bonissone *et al*. argued that their data set suggested that only Ser and Ala were important for NME and that the remaining target residues were evolutionarily inconsequential for NME [36].

An alternative hypothesis for the function of NME is that it primes the N-termini of proteins so that other post-translational modifications can occur, such as N-terminal acetylation [18]. Recent studies report data in support of a connection between Nt acetylation of proteins and NME and NER in the yeast *S. cerevisiae*, where N-terminally acetylated proteins (either on the iMet or the remaining

P2 residue) can be targeted as N-degrons for degradation by the Doa10 E3 ligase, with the Ubc6 or Ubc E2 enzymes acting as a N-recognin [18]. This system is limited to select proteins and has yet to be characterized in higher eukaryotic cells. Further discussion on NME and N-terminomics studies are continued under sub-section 'N-terminal acetylomes in prokaryotes'.

Rim N-acetyltransferasess (NATs)

This review only discusses bacterial Nt acetyltransferases. For a full review on eukaryotic Nt acetyltransferases, see [39]. For a full summary of characterized bacterial NATs and corresponding references, refer to Table 1.

Some of the first examples of N(alpha) (also referred to as Nα) acetylation of bacterial proteins implicated ribosomal proteins modified by the Rim (**ri**bosomal **m**odification) acetyltransferases in *Escherichia coli*. In the 1970s, two-dimensional polyacrylamide gel electrophoresis was used to characterize the ribosomal proteins of more than 1500 thermosensitive *E. coli* mutant strains, which led to the discovery of mutations outside of ribosomal structural genes that altered the physical characteristics of some ribosomal proteins. Utilizing the method of mapping by the gradient of transmission [40], the mutations identified were in the genes encoding the acetyltransferases RimJ, RimI, and RimL [41-44].

The acetyltransferase RimJ acetylates the N-terminal alanine of ribosomal protein S5 in *E. coli* [44-46]. To date, the function of this acetylation event is unknown. Overexpression of *rimJ* was found to suppress the growth defect of an *E. coli* strain with a mutation in ribosomal protein S5 that resulted in cold sensitivity and ribosome biogenesis [47]. Not only has RimJ been implicated in the thermal stability of the ribosome, RimJ has been found to be involved in thermoregulation of the pyelonephritis-associated pili or *pap* operon, an operon that is repressed at 23 ºC [48]. In mutants that can transcribe the *pap* operon at non-permissive temperature of 23 ºC, overexpression of *rimJ* restores thermoregulation of the *pap* operon, and prevents transcription at 23 ºC. It is suggested that RimJ controls *pap* transcription in a temperature-dependent manner by preventing phase variation to turn the genes into the "ON" state; however, the exact mechanism by which RimJ regulates *pap* is unknown [49]. RimJ has also been shown to N-terminally acetylate the recombinantly expressed proteins of thymosin alpha1 and the Z-domain of staphylococcal protein A during co-overexpression in *E. coli* [50-52]; however, the lack of physiological relevance of these protein substrates may just indicate N-terminal amino acid preferences of RimJ.

The second ribosomal modification acetyltransferase is RimI, which acetylates the N-terminal alanine of the 30S ribosomal subunit protein S18, following cleavage of the initiator formyl-methionine [45, 53]. Even though protein S18 has been shown to be acetylated by RimI, the physiological impact of this acetylation event is still unknown. There is speculation that since the S18 protein is located near the E-site of the ribosome, acetylation of S18 could influence the translation initiation process by changing the structure of S18 near the mRNA landing site [54]. Vetting and co-workers performed the biochemical characterization studies of the *Salmonella typhimurium* RimI acetyltransferase in complex with acetyl-CoA, a peptide substrate, and solved the crystal structure of RimI to understand the reaction mechanism of RimI [55]. RimI acts on acetyl-CoA through a nucleophilic addition-elimination mechanism, where Glu103 acts as the catalytic base and accepts a proton from ribosomal protein S18. This causes a nucleophilic attack on the carbonyl group of acetyl-CoA by the N-terminal nitrogen of alanine in protein S18 which causes the formation of a tetrahedral intermediate. Lastly, Tyr115 of RimI serves as the acid by providing a proton and causing decomposition of the tetrahedral intermediate [55].

A study in *E. coli* attempted to characterize the potential for RimI to be a Nε-lysine acetyltransferase through the construction of a strain of *E. coli* where the primary protein lysine acetyltransferase YfiQ (Pat homologue from *S. enterica*) gene was deleted from the chromosome, as well as the gene encoding Pta, which enables *E. coli* to generate acetyl-phosphate. Proteins can be non-enzymatically acetylated on lysyl residues by acetyl-phosphate [56]. The sole lysine deacetylase, CobB, was also deleted along with the gene *acs* encoding

Table 1. Known bacterial N terminal acetyltransferases.

Organism	Enzyme	Substrate	Acetylated Nt residue	Physiological relevance	PDB	References
Escherichia coli	RimJ	Ribosomal protein S5	P' Alanine	Thermal stability		[45-47]
Escherichia coli, *Salmonella* Typhimurium, *Mycobacterium tuberculosis*	RimI	Ribosomal protein S18	P' Alanine	Unknown	5ISV, 2CNS	[46, 54], [60]
Escherichia coli, *Salmonella* Typhimurium	RimL	Ribosomal protein L7/12	P' Serine	Implicated in adaptation of change in growth condition or phase variation	1S7N	[44, 54, 62, 64], [65], [66]
Salmonella Typhimurium	NatA	$CobB_L$ sirtuin deacylase	iMethionine	Decreases enzymatic activity, regulates acetate metabolism		[69]

acetyl-CoA synthetase, which in the absence of Pat causes detrimental effects on cell growth [57]. The authors call this strain the 'gutted strain', and provide the *E. coli rimI* gene *in trans*, then assay cell lysates in anti-acetyllysine antibody western blots to look for changes in acetylated protein patterns [58]. Cell lysates enriched for acetyllysine were subjected to LC-MS/MS to identify protein targets of RimI. The authors found 11 unique proteins with acetylation in the RimI-overexpressed lysate with 14 unique lysine sites acetylated [58]. However, no further validation of acetylation of the identified proteins by RimI was conducted.

In an attempt to confirm acetylation of ribosomal protein S18 and expand the number of potential acetylation targets of RimI, the biochemical characterization of the *Mycobacterium tuberculosis* homologue of RimI (~32% sequence identity to *S.* Typhimurium RimI) was performed [59]. The $RimI^{Mtb}$ enzyme was purified to homogeneity and tested for its ability to acetylate small peptide substrates that represent the first six residues of ribosomal protein S18, as well as a small-scale tryptic peptide library (STPL) was a tested for acetylation by $RimI^{Mtb}$ by using MS/MS, MALDI-TOFF, or the DTNB assay. A caveat of this study was that even though the STPL assay demonstrated that $RimI^{Mtb}$ acetylated peptides identical to the N-terminus of proteins whose genes were located near *rimI* in the chromosome (*e.g.,* GroES, GroEL1, TsaD), purified GroES and TsaB-TsaD were not acetylated by $RimI^{Mtb}$ in an *in vitro* [^{14}C]-acetyl-CoA radiolabel transfer assay [59]. The authors argue that $RimI^{Mtb}$ has more relaxed substrate specificity as compared to Rim acetyltransferases commonly being considered as highly substrate specific. However, this flexibility in N-terminal substrates is only seen using the peptide library and not with full length protein targets [59]. The authors provide *in vivo* support of protein-protein interactions between $RimI^{Mtb}$ and genes that are located near *rimI* in the genome through Mycobacterial Protein Fragment Complementation (MPFC) assay, where successful protein-protein interactions *in vivo* allowed reconstitution of mDHFR (dihydrofolate reductase), which aids in survival of cells exposed to trimethoprim, an antibiotic that blocks the

reduction of dihydrofolate. Survival of cells where co-transformants yielded trimethoprim resistance occurred in pairings with $RimI^{Mtb}$ and GroES chaperone proteins, and $RimI^{Mtb}$ with TsaD (a neighboring tRNA-A_{37}-t^6A transferase) [59].

The $RimI^{Mtb}$ protein was further investigated using biophysical approaches to characterize variants of $RimI^{Mtb}$ to generate a more stable protein. The most stable variant of $RimI^{Mtb}$, called MtRimIC21A4-153 (a variant containing a cysteine to alanine point mutation with both N-terminal and C-terminal truncations) was compared to the full-length native $RimI^{Mtb}$ using 2D ^{1}H-^{15}N heteronuclear single quantum coherence (HCQC) NMR spectroscopy and circular dichroism (CD) and it was found that the truncations and mutation had very little effect on the enzymatic activity of this enzyme as compared to the parental protein [60].

The third known prokaryotic Rim acetyltransferase is RimL [61], which acetylates the N-terminal serine residue in ribosomal protein L12 in *E. coli* [62]. This acetylation event results in the formation of L12's acetylated form, thus called L7 [63]. Since the amino acid sequence of these proteins is identical other than the N-terminal acetylation of L7, this protein is referred to as ribosomal protein L7/L12. *E. coli* RimL can post-translationally acetylate L12 *in vitro* and acetylate L12 *in vivo*, when ribosomal protein L12 is in complex with the 50s ribosomal subunit [43, 53, 63]. The *Salmonella* Typhimurium RimL protein can also acetylate ribosomal protein L12 [64].

The population of ribosomal proteins S5 and S18 are always found in their acetylated state, but protein L12 is unique in that the ratio of L7/L12 protein (L12 = unacetylated L7) differs depending on the rate and growth phase of the cells, suggesting a role of acetylation in modulation during the growth cycle. During growth in rich medium, the L12 (*i.e.,* unacetylated L7) is present in increasing concentrations over the L7 form, from early logarithmic to mid-logarithmic phase, where 85% of L7 is unacetylated [65]. However, as cells enter stationary phase the acetylated form, L7, becomes the dominant form over the unacetylated L12 [65]. Another important feature of the aforementioned study is that the shift in increased acetylation seen in L7/L12 proteins requires the production of new ribosomes and not modification of already formed ribosomes containing unacetylated L12 proteins [65]. The authors suggest that the acetylated L7 protein may aid the survival of a nongrowing culture or in the adaptation to a different condition or cell phase [65]. In contrast, when cells are grown in minimal medium, L12 protein lacking its N-terminal methionine is 100% acetylated, that is, in its L7 form. The authors' data suggest that during conditions of stress, acetylation of L12 (*i.e.,* L7) increases the stability of the ribosomal stalk complex where acetylated L7 protein has a tighter interaction with its associated ribosomal subunit L10 [66, 67]. Even with consideration of the aforementioned experiments, further experimentation is required to elucidate the physiological relevance of RimL acetylation of L12.

The studies of *in vitro* acetylation of L12 by *E. coli* RimL further investigated several factors aiming to compare the acetylation flexibility of RimL to its eukaryotic Nα-acetyltransferase counterparts. To test the specificity of RimL for L12's N-terminal amino acid sequence, the ability of the enzyme to acetylate L12 protein whose N-terminal serine was changed to alanine was conducted. First, the authors determined that RimL acetylated the $L12^{I2A}$ variant, albeit at a slower rate than the wildtype L12 protein. Second, the ability of RimL to acetylate a L12 variant that has the second residue changed from isoleucine to aspartic acid ($L12^{I2D}$) was tested, and the authors showed that RimL could acetylate the $L12^{I2D}$ variant, but to a much lesser degree than the wildtype L12 protein [63]. The authors note this is in stark contrast to eukaryotic Nα-acetyltransferases that typically acetylate N-termini with acidic residues such as aspartic acid, in the P2 position [11]. The final comparison noted is that eukaryotic Nα-acetyltransferases require accessory protein subunits for activity, whilst RimL homodimers do not [64].

Non-ribosomal N-terminally acetylated proteins

Salmonella NatA and CobB

The *Salmonella enterica enterica* sv Typhimurium str. LT2 (*S.* Typhimurium) genome possesses

~26 GCN5-related acetyltranserase genes, of which the *in vivo* function of half of their gene products has been elucidated. Our group recently determined that the *S.* Typhimurium putative acetyltransferase NatA (formerly YiaC) acetylates the N-terminus of the NAD^+-dependent CobB sirtuin deacylase [68]. In this bacterium, the CobB deacylase is present in two biologically active isoforms, which differ in size by a 37-amino acid N-terminal extension. NatA can acetylate the long isoform of CobB (hereafter $CobB_L$), resulting in a decrease in the enzymatic ability of $CobB_L$ to deacylate its *bona fide* protein substrate Acs^{Ac} (Figure 2) [68]. The NatA and $CobB_L$ are the first non-ribosomal, Nα-acetyltransferase/protein substrate pair to be identified in bacteria, as well as being the first example of physiological function control of a bacterial sirtuin deacylase by N-terminal acetylation.

N-terminally acetylated proteins without known acetyltranferases

Several other proteins have been identified as N-terminally acetylated in bacteria but do not have a cognate acetyltransferase, nor has the relevance of the modification been elucidated. For example, in *Escherichia coli*, the ribosomal elongation factor Tu (Ef-Tu) was found to be N-terminally acetylated on Ser after iMet cleavage [69].

Another protein of interest is the mycobacterial Esx-1 (ESAT-6 system 1) transporter, which is involved in the pathogenesis of mycobacterial species. Several Esx-1 protein substrates have been found to be N-terminally acetylated [70-72]. Two mycobacterial proteins transported by Esx-1, called EsxA (ESAT-6, which stands for early secreted antigenic target, 6 kDa) and EsxB (a.k.a. culture filtrate protein, 10 kDa, CFP-10), have been shown to interact with each other in an acetylation-dependent manner, where N-terminal acetylation of EsxA decreases the binding interaction with EsxB, or possibly changes complex oligomerization [73]. A genetic screen aimed to identify genes involved in Esx-1 export revealed a deletion of a putative N-acetyltransferase called MMAR_0039 that caused hyper N-terminal acetylation of bacterial surface-associated EsxA. The authors noted that seeing an increase in acetylation when an acetyltransferase was deleted was unexpected, and proposed that alternative proteins in this pathway that could be modified by MMAR_0039 are yet to be identified [72].

A study analyzing changes in the proteome of the radiation-tolerant bacterium *Deinococcus deserti* when exposed to gamma radiation over time found not only an increase in the presence of GyrA (DNA gyrase subunit A), the enzyme responsible for introducing negative supercoiling to DNA, but also a mixture of N-terminally acetylated and unacetylated populations of GyrA that had the iMet removed and were modified on the remaining Thr residue. The authors also note upregulation of a putative GCN5-related acetyltransferase, deide_20140, but do not have direct evidence that deide_20140 modifies GyrA [74].

N-terminal acetylomes in prokaryotes

Prokaryotic NTA is more prevalent than previously thought

In bacteria, the long-held paradigm of N-terminal acetylation occurring rarely and only in the instances of ribosomal protein acetylation has been challenged through the study of N-terminal protein modifications by mass spectrometry. In recent years, N-terminomics studies have provided evidence that N-terminal protein modification is common in bacteria. For example, work performed in the early 2010s provided insights into the conservation of NME motif recognition, which was discovered through blind PTM searches of 68 million bacterial mass spectra in 45 different bacterial species, obviating the fact that the prevalence of N-terminal protein acetylation was much higher than previously realized. Remarkably, 16% of the N-terminal modifications analyzed were N-terminal acetylations, and most of the N-terminal residues that were acetylated were Ser, Ala, or Thr as in the yeast data set [36].

To visualize the most common N-terminally acetylated residues found in the N-terminomics data from different bacterial species described here as well as corresponding references, refer to Table 2.

Shortly after the Bonnissone *et al*. paper was published, a dedicated N-terminomics study of *Pseudomonas aeruginosa* PA14 was published [75].

Table 2. Most common N-terminally acetylated residues based on N-terminomics data.

Organism	Most common residue with NTA	Second most common residue with NTA	Third most common residue with NTA	Fourth most common residue with NTA	References
E. coli	Ser	Ala	Met	Thr	[14, 15]
P. aeruginosa	Met	Ser	Ala	Thr	[76]
M. tuberculosis	Thr	Met	Ser	Ala	[82]
A. baumannii	Met	Ser/Ala/Val	-	-	[81]

To enrich for N-terminal peptides, the authors utilized protein fractionation by isoelectric points (pI) and then analyzed digested protein fractions of nanoLC-MS/MS. It was found that 63% of the peptides analyzed had populations with complete iMet cleavage, 31.5% of peptides fully retained their iMet, and 4.9% of specific peptide sequences were found both with retained iMet and with iMet cleavage [75]. To select for N-terminally acetylated peptides, three strategies were incorporated into the workflow. The first strategy was to perform a direct analysis of extracted proteins after tryptic digestion. The second strategy involved enrichment of N-terminally acetylated peptides through CNBr-activated Sepharose resin [76]. The final strategy encompassed in-solution isoelectric focusing (IEF) fractionation using the Microrotofor IEF cell. The combination of the aforementioned strategies yielded 117 N-terminally acetylated peptides corresponding to proteins involved in a wide range of cellular functions, including transcription, translation, metabolism, transport, and also a large set of hypothetical proteins, showing that Nt acetylation may be integrated into multiple facets of *P. aeruginosa* cellular functioning [75]. A set of N-terminally acetylated proteins whose iMet was cleaved had P2 residues that fit within the NME-specific criteria, such as Ser, Ala, Gly, Thr or Val, yet surprisingly peptides were found acetylated with and without iMet that largely contained Asn and to a smaller extent several other residues at P2 (*e.g.,* Ile, Ser, Thr, Gln, Gly, Lys Leu, Ala, Val). Finally, a subset of N-terminal sequences was found to have acetylation on the iMet that also possessed P2 residues with small radius of gyration such as Ser, Ala, Thr, and Gly, but were not limited to these P2 residues, with Asn, Leu, Gln, Ile and Lys being more abundant as P2 residues retaining iMet. The authors hypothesized that even though these N-termini should be substrates for NME, perhaps Nt acetylation blocked MAP catalysis of the target protein's iMet. A large set of proteins' N-termini that were acetylated on the iMet possessed P2 residues that are considered 'destabilizing' (*e.g.,* such as Glu, Asp, Gln, His, Ile, Lys, Leu, or Asn) where retainment of the iMet is expected because these P2 residues are not compatible with MAP, but they also are not compatible substrates for Rim acetyltransferases, suggesting alternative Nt acetyltransferases could be involved. *P. aeruginosa* possesses 50 genes annotated as potential acetyltransferases which could be alternative N-terminal acetyltransferases. The *S.* Typhimurium protein deacylase CobB possesses an N-terminal Met-Glu sequence which is recognized and acetylated by the NatA acetyltransferase [69], lending support to the idea that GNATs with similar N-terminal substrates could exist in *P. aeruginosa*.

Bienvenut *et al*. used SILProNAQ (stable-isotope protein N-terminal acetylation quantification) methodology [77] to conduct a proteome-wide analysis of the effect on the N-terminome of *E. coli* proteins when iMet deformylation was inhibited by the PDF-specific antibiotic actinonin [15]. The authors' data suggested that 10% of the proteins identified in their study were N-terminally acetylated, and they found over 100 distinctly acetylated N-termini. This dataset of Nt acetylation consisted of three subsets due to NME; a set of 27 proteins (3% of the 10% total acetylated protein) with acetylation on the iMet; acetylation on the P2 residue (75 proteins, 7% of

the Nt acetylation data set); and proteins with two different subpopulations where acetylation occurred on both the iMet and on the P2 residue after NME (two proteins, 0.2% of dataset) [15].

A large-scale quantitative *E. coli* proteomics study was conducted utilizing peptide mixtures from 22 different growth conditions [14]. The proteins were digested with trypsin and either directly analyzed by shotgun LC-MS/MS or were fractionated by their theoretical isoelectric points using OFFGEL electrophoresis [78]. The authors used stable isotopic dilution (SID) and selected reaction monitoring (SRM) mass spectrometry to quantify the number of copies of proteins of interest per cell. Their control experiments included the quantification of 41 proteins, which were used as references in an effort to accurately estimate the abundance of proteins. The Intensity Based Absolute Quantification (iBAQ) approach was used for protein quantification [79]. The resulting mass spectra were analyzed for all types of protein modifications, but of interest to this review, 31 proteins were found to be acetylated at the N-terminus [14]. In this report, the authors showed that Nt acetylation occurred primarily on N-terminal Ser, Ala, Met, or Thr. Strikingly, the authors also noted that the abundance of Nt acetylation decreased as the average growth rate of a culture increased, with the most Nt acetylation abundance seen in cells that had reached and maintained stationary phase for three days, and the least abundance of Nt acetylation was found in exponential phase in rich medium. Collectively, these data provided a glimpse into the diversity of proteins modified by Nt acetylation and into the abundance of Nt acetylation observed under different growth conditions.

The authors further analyzed the N-terminal acetylation pattern of peptides isolated from three independent mutants of *E. coli*, each lacking either *rimJ*, *rimI*, or *rimL*. The only change found from the parent strain was in the *rimJ* mutant strain, where a significant decrease in Nt acetylation at Ser or Thr residues were seen [14], suggesting that RimJ was modifying proteins other than ribosomal protein S5 *in vivo*, which supported the studies on regulation of the *pap* operon by RimJ [50].

A proteomics study enriching for both Nα-acetylated and Nε-acetylated proteins in *Acinetobacter baumannii* strain ATCC 17978 was conducted using a similar work flow as the one used to study *P. aeruginosa* [75], where direct LS-MS/MS analysis was compared to IEF fractionated peptides and CNBr-activated Sepharose bead enrichment of N-terminal peptides [80]. This study included an additional enrichment process that included separating peptides by charge using solid cation exchange chromatography (SCX). This study identified 145 N-terminally acetylated proteins, where 59% of the proteins identified possessed Nt acetylation on the iMet, followed by Ser (7%), Ala (7%), and Val (7%) [80]. Just as in the *P. aeruginosa* N-terminome, a large subset of proteins possessing P2 residues with small radius of gyration was acetylated on the iMet when MAP cleavage should have occurred, again suggesting perhaps acetylation of the iMet blocked MAP cleavage of these peptides. Seven peptides identified had subpopulations that were acetylated on the iMet or the free P2 residue, where all but one of these peptides had Asn as the P2 residue. The authors also noted that Asn was the most abundant residue at the P2 position [80].

Another validating note of their dataset was that 9 of the 145 proteins identified to be N-terminally acetylated were already identified as potentially acetylated in other bacteria, those proteins being ribosomal protein L10 [76], ribosomal protein L7/L12 [45], ribosomal protein S18 [36, 45], protein elongation factor EfTs [76], DapD [15], trigger factor [15], preprotein translocase subunit SecB [15, 75], and the beta subunit of nucleoid associated protein HU [75]. The categories of proteins found to be N-terminally acetylated were involved in metabolism, transcription and translation, and carbon catabolism, but a few specific genes that differed were those of the biofilm and adhesion associated protein Bap (A1S_2724), starvation-induced protein CtsA, and the antibiotic efflux pump component AbeM, to name a few [80].

Other investigators performed a quantitative analysis of Nt acetylation in the pathogenic bacterium *Mycobacterium tuberculosis* strain Erdman and the opportunistic pathogen

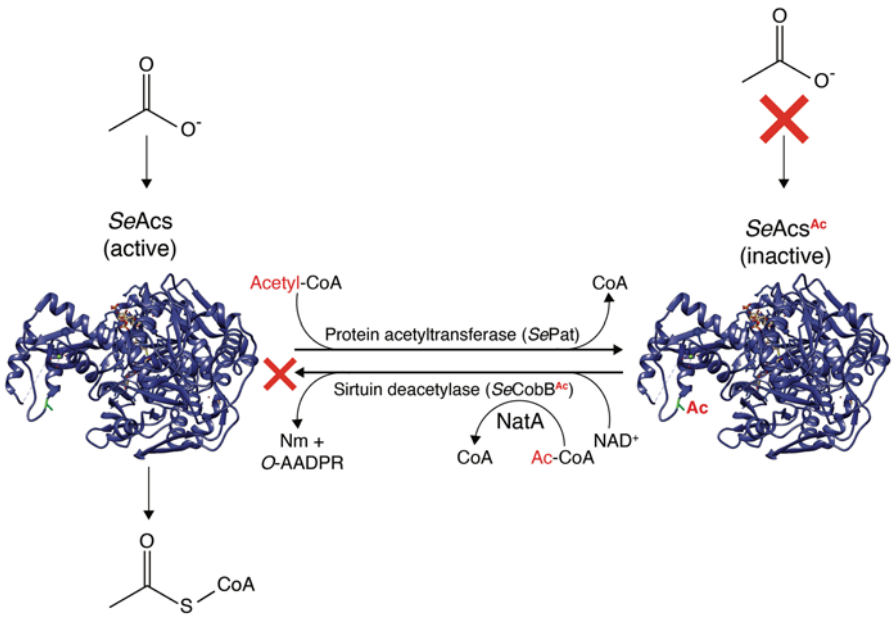

Figure 2. Schematic of the role of NatA in the modulation of acetate metabolism in *Salmonella* Typhimurium. *Se*Acs (acetyl-CoA synthetase) converts acetate into acetyl-CoA *via* an acetyl-adenosine monophosphate (AMP) intermediate. *Se*Acs can be acetylated in active site lysine 609 by the protein lysine acetyltransferase *Se*Pat, which renders *Se*Acs inactive. *Se*Acs activity can be restored *via* deacetylation by *Se*CobB. NatA adds an additional layer of regulation through acetylation of $CobB_L$, which renders $CobB_L$ inactive and prevents flux of acetate through *Se*Acs.

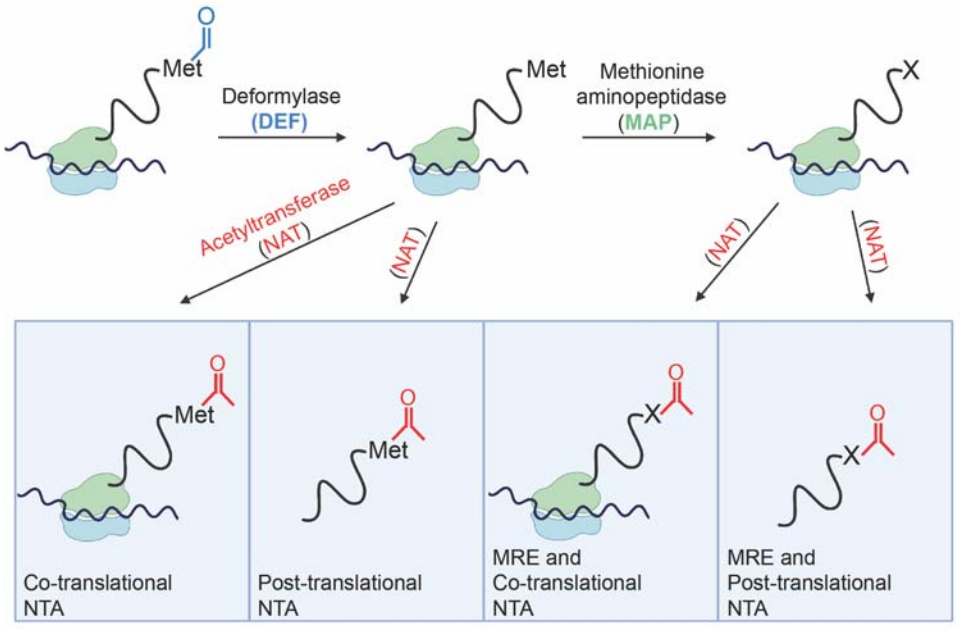

Figure 3. Working model for sub-populations of protein N-termini in bacterial species where N-terminomics has been studied. Deformylase (DEF) removes the formyl group from newly synthesized peptides at the ribosome (formyl group in blue). The N-terminus of the peptide can either be acetylated co- or post-translationally by N-terminal acetytransferases (NAT, red), or the initiator methionine (Met) can be excised by methionine aminopeptidase (MetAP, green). The remaining N-terminal residue may be acetylated co-or post-translationally by NATs (red).

Mycobacterium marinum. These studies were performed by combining several N-terminal peptide enrichment protocols followed by nano-UHPLC-MS/MS [81]. These enrichment protocols included chemical acetylation of free amines using N-acetoxy-D_3-succinamide, so that physiologically relevant acetylation could be distinguished from chemical acetylation. For protein digestion, trypsin and GluC proteases were used to increase the sequence coverage of N-terminal peptides that have Arg or Lys-rich N-termini. Finally, the use of both strong cation exchange (SCX) and strong anion exchange (SAX) chromatography enabled greater enrichment of N-terminal peptides from internal peptides. The resulting study identified 211 proteins as N-terminally acetylated out of 1920 total protein N-termini identified in *M. tuberculosis*, and 347 N-terminally acetylated peptides out of 838 identified N-terminal peptides in *M. marinum*. Of the N-terminally acetylated peptides identified, the majority of peptides that were acetylated on the retained iMet had P2 residues that were Ala, followed by Lys, Gln, and Phe [81]. Of N-termini that had been acetylated on the remaining P2 residue, almost all were acetylated on Thr residues, followed by a few peptides possessing acetyl-Ser. This finding was striking, and in stark contrast to the preferential acetylation of Ser or Ala residues seen in *E. coli* [15].

A large proportion of proteins identified as N-terminally acetylated were categorized as being functionally involved in intermediary metabolism and respiration in these organisms. The next largest category included proteins that are considered 'conserved hypothetical' proteins, and the next largest subset belonged to cell wall and cell processes, information pathways, and lipid metabolism [81].

N-terminomics of the obligate photoautotrophic anaerobic green sulfur bacterium *Chlorobaculum tepidum* was conducted using the Combined Fractional Diagonal Chromatography (COFRADIC) method [82] on sub-fractionated proteins. The resulting analysis only found 4 proteins to be N-terminally acetylated, those proteins being a subunit of a transketolase (locus tag acc #: CT0804) acetylated on an N-terminal Gly, a glutathionine S-transferase (locus tag acc #: CT0249) acetylated on the iMet residue with a P2 residue of Asn; a putative response regulator (locus tag acc #: CT2059) acetylated on the iMet residue with a Lys at P2; and finally, a putative uncharacterized protein (locus acc #: CT2140) acetylated on the iMet residue with a P2 Asn residue [83].

Methods and limitations in studying N-terminal acetylation

The field of Nt acetylation has several challenges and limitations. Quantitative N-terminomics studies suggest that the ratio of acetylated *vs* unacetylated N-termini is very low, creating difficulties in detection of acetylated N-termini [75]. Another caveat is that in contrast to lysine acetylation where anti-acetyllysine antibodies exist to enable detection of acetylated proteins, no antibodies exist for the detection and enrichment of N-terminally acetylated proteins. This is in part due to the variability of amino acid sequence of the N-terminus.

Mass spectrometry-based approaches

Mass spectrometry is the most frequently used method for the identification of novel, *in vivo* N-terminally acetylated proteins. Several methods have been developed to enrich for N-terminal peptides and exclude internal peptides, which typically are more abundant and crowd N-terminal peptide signals. In-depth descriptions of methodologies used in N-terminomics studies have been reviewed [84]. In general, an N-terminomics study includes i) chemical acetylation of free amines with isotope-labeled acetyl-groups, ii) protease digestion, iii) fractionation of N-terminal peptides with SCX and/or SAX, iv) liquid chromatography followed by mass spectrometry, and v) bioinformatics analysis. For reviews concerning each of the different methods, refer to these references: COFRADIC method [83, 85-90]; SILProNAQ [15, 77]; Cyanogen bromide-activated Sepharose [76, 89]; SCX [91-92]; SAX [81]. N-terminal peptides that retained their iMet residues can be separated from N-terminally cleaved peptides by oxidation of methionine to methionine sulfoxide using a reagent such as hydrogen peroxide. Peptides containing methionine sulfoxide elute earlier during reverse-phase-HPLC separation than those without iMet residues [92].

In vitro assays

4-Chloro-7-nitrobenzofurazan (NBD-Cl) assay

An affordable, quick method for the *in vitro* detection of N-terminally acetylated proteins compared to unmodified N-termini with detection in the micromolar range was developed utilizing fluorogenic derivatization with 4-chloro-7-nitrobenzofurazan (NBD-Cl) [93]. At pH 7, NBD-Cl reacts with free N-terminal amines of proteins, giving off a fluorescent signal at wavelength 535 nm. Proteins that possess a blocked N-terminal amino group (*i.e.*, by an acetyl group) do not react with NBD-Cl and no fluorescence is seen. This method is specific to N-terminal amino groups, where free amino groups on the side chain of lysyl residues do not react with NBD-Cl. A great aspect of this method is that proteins do not need to be modified for the assay to be conducted; however, the pH must be maintained near 7 for the NBD-Cl dye to function properly.

Radiolabeled Ac-CoA transfer assay

N-terminal acetylation and lysine acetylation of proteins can be visualized *in vitro* through the utilization of ^{3}H- or ^{14}C-radiolabeled acetyl-CoA, where a carbon or hydrogen associated with acetyl group is a radioisotope. Radiolabeled acetyl-CoA can be incubated with cell lysates that contain acetyltransferase proteins and target substrates, then are resolved by SDS-PAGE, exposed to a phosphor screen, and imaged [1]. The appearance of bands in the phosphor image that correlate with the molecular mass of target proteins in conjunction with control reactions shows whether acetylation has occurred. A major consideration when conducting *in vitro* acetylation assays is to perform the assays near pH 7. At pH values 8 or greater, a higher percentage of lysyl residues are protonated and the thioester bond of acetyl-CoA is more labile to hydroxyl anions. These factors favor non-enzymatic acetylation to occur, which may cause false positives to occur since radiolabeled acetylation assays do not differentiate epsilon acetylation from alpha acetylation [6].

An alternative method that allows *in vivo* incorporation of radiolabeled acetyl groups using ^{3}H-sodium acetate has been developed called cell-based acetylation [94-96]. Cells of interest in culture media are incubated with ^{3}H-sodium acetate. Protein targets of interest are purified out of cell extracts and then resolved by SDS-PAGE and imaged using a phosphor imager. A major caveat with this assay is non-enzymatic acetylation *via* acetyl-phosphate and acetyl-CoA. It has been shown that in stationary phase, lysyl residues of proteins are used as 'sinks' for storage of excess acetyl groups, owing to the value that acetyl-phosphate and acetyl-CoA hold as high energy molecules in the cell. Non-enzymatic incorporation of radiolabeled acetyl groups may occur to proteins and result in false positives [6, 7, 56].

Ellman's reagent assay

5,5-Dithio-*bis*-(2-nitrobenzoic acid) (DTNB) is commonly used for the detection and quantification of free sulfhydryl groups, and can be used to detect the free thiol generated from CoA after an acetyltransferase adds the acetyl-group from acetyl-CoA to an amino group of a protein [97, 98]. The reaction of DTNB with the free thiols can be followed spectrophotometrically at 412 nm [99]. This *in vitro* assay is suitable for both epsilon and alpha-acetylation, since free CoA is the product for both reactions, and allows for the determination of kinetic parameters of enzymes. This method, however, is better suited for purified proteins and is sensitive to reducing agents that may be present in protein preparations.

HPLC assays and N-terminal peptide libraries

Reverse-phase high-performance liquid chromatography (HPLC) can be used to quantify and kinetically (with radiolabeled peptides) define *in vitro* acetylation products of CoA and N-acetylated protein substrates. Alternatively, acetylated peptides, being more hydrophobic, can be separated from unacetylated counterparts. Synthetic peptide libraries have been used in order to elucidate N-terminal amino acid sequence preferences, but this method of screening peptide libraries is a more biased approach and does not consider any interactions that occur between acetyltransferases and target proteins that involve secondary or tertiary structural arrangements.

A less biased method using peptide libraries was developed that utilizes proteome-derived peptides [100]. Briefly, total proteins are isolated,

alkylated, chemically acetylated, then digested with trypsin. Since lysyl residues were chemically acetylated, trypsin only cleaves C-terminal to arginyl residues. N-terminally acetylated peptides are then separated from unacetylated N-terminal peptides with low pH using SCX [86], and peptides that contain free N-terminal amines are kept and used as substrate for purified acetyltransferase enzyme and radiolabeled acetyl-CoA. Acetylated peptides can then be separated from unacetylated peptides by SCX, and then identified by LC-MS/MS [87, 100].

Chloroacetyl-CoA assay

Chloroacetyl-CoA was used in several *in vitro* experimental designs to identify protein substrates [93]. Chloroacetyl-CoA was synthesized by combining acidic free CoA and chloroacetic anhydride, then purified by HPLC. The synthesized chloroacetyl-CoA was used as a substrate for several acetyltransferases, including the eukaryotic Hat1 acetyltransferase that acetylates histone H4, as well as the *S. Typhimurium* RimL acetyltransferase that acetylates ribosomal protein L12. An *in vitro* acetylation assay was developed where target protein(s) were incubated with chloroacetyl-CoA and the purified acetyltransferase protein. Cysteamine-TAMRA is generated *in situ* by combining the fluorophore TAMRA [5(6)-carboxytetramethylrhodamine] with DIPEA [N,N-diisopropylethylamine] and the peptide-coupling agents HOBt/PyBOP. The thiol group of cysteamine-TAMRA displaces the chloride of the chloroacetyl moiety on the N-terminal amine of the target protein, thus generating a TAMRA-labeled fluorescent protein. The modified protein can then be detected by fluorescent imaging of proteins resolved by SDS-PAGE. To isolate acetylated targets of an acetyltransferase, a fluorescently labeled peptide called Fluorescein-His8GGC was generated that acted in a similar manner to the cysteamine-TAMRA to capture the chloroacetylated protein, but with the ability to isolate the chloroacetylated protein from cell extracts utilizing nitrilotriacetic acid (NTA) chromatography and visualization with the fluorescein fluorophore [93]. This method does not differentiate epsilon acetylation from alpha acetylation, as well as preparation of substrates may be technically challenging and laborious for some labs. However, like anti-acetyllysine antibodies, it can allow for isolation of acetylated proteins, whether it be on N-termini or lysyl residues.

Concluding remarks and outstanding questions

The field of prokaryotic N-terminal acetylation has just begun to make advancements. There are many outstanding questions and challenges to address in this field of study. Due to the low sequence homology of GCN5-acetyltransferases, identifying functional homologues of acetyltransferases based on sequence alignments is difficult. Another difficulty in the field is the ability to discern lysine acetylation of protein targets from N-terminal acetylation. Non-enzymatic acetylation of lysyl residues raises the background 'noise' that can only be mitigated by the identification of an acetyltransferase specific for the protein substrate of interest. Only then, the acetylation site can be established as a *bona fide* modification target. Obviously, the foremost outstanding question is the physiological reason for N-terminal acetylation; however, other outstanding questions in the field are as follows:

Why are different N-terminal populations seen?

From analysis of N-terminomics data that exist already, it seems that different populations of N-termini of the same protein can exist in the cell, ranging from iMet excision from part of a population but not all, to full iMet excision, to no iMet excision (Figure 3). What is the purpose of having different residues on the N-terminus of a protein, as well as what is the purpose of chemical modifications of the N-termini? Why are some N-termini acetylated on the iMet, while others require NME-mediated iMet excision to be acetylated?

Why are sequences recognized by MAP not cleaved by NME? And why does NME cleave sequences that are not predicted to be its substrates?

Current N-terminomics data show populations of proteins with iMet cleavage that are for proteins that do not fit within the canonical set of residues in the P2 position of the peptide chain, such as Gly, Ala, Ser, Pro, Val, Thr, or Cys. This begs the question of whether MAP functions differently

in vivo, or alternative iMet excision enzyme-encoding genes exist within the genome that are active against alternative P2 residues. It has also been demonstrated that lack of cleavage of iMet by NME of proteins that do possess optimal P2 residue sequences also occurs *in vivo*. It is thought that the NME process occurs shortly after a protein is synthesized and deformylated, which begs the question as to why those proteins that are optimal for MAP-mediated iMet cleavage retain their iMet residues. Does acetylation serve a role in preventing iMet cleavage? Can iMet cleavage occur on older proteins?

What role does methionine oxidation/reduction play in NME and NTA, and can N-terminal acetyltransferases acetylate oxidized methionine?

Free methionines and methionines within a peptide chain can be oxidized to methionine sulfoxide by reactive oxygen species (ROS). Methionine sulfoxide can be reduced back to methionine through the function of methionine sulfoxide reductases as discussed earlier. It is known that MAP activity is decreased for methionine sulfoxide *in vitro*. What is the physiological hinderance of N-terminal methionine sulfoxide on MAP activity? Another question around N-terminal Met-SO is can N-terminal acetyltransferases acetylate N-terminal Met-SO?

Is N-terminal acetylation co-translational or post-translational?

In eukaryotes, most N-terminal acetyltransferases have been shown to interact at the ribosome in order to acetylate the N-terminus of newly synthesized polypeptides. Most N-terminal acetyltransferases even possess an auxiliary subunit that enables ribosomal attachment. Are prokaryotic N-terminal acetyltransferases also interacting with the ribosome and performing N-terminal acetylation at the ribosome, just after deformylation, or is N-terminal acetylation occurring post-translationally? How does the interplay between deformylation and N-terminal acetylation occur?

It is clear there is a dearth of knowledge in the field of prokaryotic N-terminal protein acetylation, and that new advancements in techniques and understanding the caveats of studying acetylation are needed to truly understand the complex system that is occurring *in vivo*.

Conflict of interest statement

The authors do not have any conflict of interest to declare.

References

1. Starai, V. J. and Escalante-Semerena, J. C. 2004, J. Mol. Biol., 340, 1005.
2. Escalante-Semerena, J. C. 2010, Microbe, 5, 340.
3. Liu, J., Wang, Q., Jiang, H., Zhao, D., Han, J., Luo, Y. and Xiang, H. 2017, J. Proteome Res., 16, 3229.
4. Allfrey, V. G., Faulkner, R. and Mirsky, A. E. 1964, Proc. Natl. Acad. Sci. USA, 51, 786.
5. Starai, V. J. Celic, I., Cole, R. N., Boeke, J. D. and Escalante-Semerena, J. C. 2002, Science, 298, 2390.
6. VanDrisse, C. M. and Escalante-Semerena, J. C. 2019, Annu. Rev. Microbiol., 73, 111.
7. Hentchel, K. L. and Escalante-Semerena, J. C. 2015, Microbiol. Mol. Biol. Rev., 79, 321.
8. Christensen, D. G., Xie, X., Basisty, N., Byrnes, J., McSweeney, S., Schilling, B., and Wolfe, A. J. 2019, Front. Microbiol., 10, 1604.
9. Shaw, W. V. 1966, Antimicrob. Agents Chemother. (Bethesda), 6, 221.
10. Burckhardt, R. M. and Escalante-Semerena, J. C. 2020, Microbiol. Mol. Biol. Rev., 84, e00090-19.
11. Polevoda, B. and Sherman, F. 2003, J. Mol. Biol., 325, 595.
12. Arnesen, T. 2009, Proc. Natl. Acad. Sci. USA, 106, 8157.
13. Bienvenut, W. V., Sumpton, D., Martinez, A., Lilla, S., Espagne, C., Meinnel, T. and Giglione, C. 2012, Mol. Cell. Proteomics, 11, M111 015131.
14. Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrne, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R. and Heinemann, M. 2016, Nat. Biotechnol., 34, 104.

15. Bienvenut, W. V., Giglione, C. and Meinnel, T. 2015, Proteomics, 15, 2503.
16. Jörnvall, H. 1975, J. Theor. Biol., 55, 1.
17. Persson, B., Flinta, C., Vonheijne, G. and Jornvall, H. 1985, Eur. J. Biochem., 152, 523.
18. Hwang, C. S., Shemorry, A. and Varshavsky, A. 2010, Science, 327, 973.
19. Kim, J. M. and Hwang, C. S. 2014, Cell Cycle, 13, 1366.
20. Kim, H. K., Kim, R. R., Oh, J. H., Cho, H., Varshavsky, A. and Hwang, C. S. 2014, Cell, 156, 158.
21. Clark, B. F. and Marcker, K. A. 1966, J. Mol. Biol., 17, 394.
22. Marcker, K. and Sanger, F. 1964, J. Mol. Biol., 8, 835.
23. Pine, M. J. 1969, Biochim. Biophys. Acta, 174, 359.
24. Meinnel, T. and Blanquet, S. 1994, J. Bacteriol., 176, 7387.
25. Solbiati, J., Chapman-Smith, A., Miller, J. L., Miller, C. G. and Cronan, J. E. Jr. 1999, J. Mol. Biol., 290, 607.
26. Mazel, D., Pochet, S. and Marliere, P. 1994, EMBO J., 13, 914.
27. Meinnel, T. and Giglione, C. 2008, Proteomics, 8, 626-649.
28. Ragusa, S., Mouchet, P., Lazennec, C., Dive, V. and Meinnel, T. 1999, J. Mol. Biol., 289, 1445.
29. Frottin, F., Martinnez, A., Peynot, P., Mitra, S., Holz, R. C., Giglione, C. and Meinnel, T. 2006, Mol. Cell. Proteomics, 5, 2336.
30. Sherman, F., Stewart, J. W. and Tsunasawa, S. 1985, Bioessays, 3, 27.
31. Huang, S., Elliot, R. C., Liu, P. S., Koduri, R. K., Weickmann, J. L., Lee, J. H., Blair, L. C., Ghosh-Dastidar, P., Bradshaw, R. A., Bryan, K. M., Einarson, B., Kendall, R. L., Kolacz, K. H. and Saito, K. 1987, Biochemistry, 26, 8242.
32. Nguyen, K. T., Lee, C. S., Mun, S. H., Truong, N. T., Park, S. K. and Hwang, C. S. 2019, J. Biol. Chem., 294, 4464.
33. Tobias, J. W., Shrader, T. E., Rocap, G. and Varshavsky, A. 1991, Science, 254, 1374.
34. Bachmair, A., Finley, D. and Varshavsky, A. 1986, Science, 234, 179.
35. Varshavsky, A. 2011, Protein Sci., 20, 1298.
36. Bonissone, S., Gupta, N., Romine, M., Bradshaw, R. A. and Pevzner, P. A. 2013, Mol. Cell. Proteomics, 12, 14.
37. Arfin, S. M. and Bradshaw, R. A. 1988, Biochemistry, 27, 7979.
38. Varshavsky, A. 1996, Proc. Natl. Acad. Sci. USA, 93, 12142.
39. Aksnes, H., Ree, R. and Arnesen, T. 2019, Mol. Cell, 73, 1097.
40. De Haan, P. G., Hoekstra, W. P., Verhoef, C. and Felix, H. S. 1969, Mutat. Res., 8, 505.
41. Isono, S. and Isono, K. 1978, Mol. Gen. Genet., 165, 15.
42. Isono, K. and Isono, S. 1980, Mol. Gen. Genet., 177, 645.
43. Isono, S. and Isono, K. 1981, Mol. Gen. Genet., 183, 473.
44. Cumberlidge, A. G. and Isono, K. 1979, J. Mol. Biol., 131, 169.
45. Yoshikawa, A., Isono, S., Sheback, A. and Isono, K. 1987, Mol. Gen. Genet., 209, 481.
46. Janda, I., Kitakawa, M. and Isono, K. 1985, Mol. Gen. Genet., 201, 433.
47. Roy-Chaudhuri, B., Kirthi, N., Kelley, T. and Culver, G. M. 2008, Mol. Microbiol., 68, 1547.
48. White-Ziegler, C. A. and Low, D. A. 1992, J. Bacteriol., 174, 7003.
49. White-Ziegler, C. A., Black, A. M., Eliades, S. H., Young, S. and Porter, K. 2002, J. Bacteriol., 184, 4334.
50. Bernal-Perez, L. F., Prokai, L. and Ryu, Y. 2012, Mol. Biosyst., 8, 1128.
51. Fang, H., Zhang, X., Shen, L., Si, X., Ren, Y., Dai, H., Li, S., Zhou, C. and Chen, H. 2009, Appl. Microbiol. Biotechnol., 84, 99.
52. Chen, J., Li, H., Wang, T., Sun, S., Liu, J. and Chen, J. 2017, Oncotarget, 8, 95247.
53. Yu, M., de Carvalho, L. P., Sun, G. and Blanchard, J. 2006, J. Am. Chem. Soc., 128, 15356.
54. Nesterchuk, M. V., Sergiev, P. V. and Dontsova, O. A. 2011, Acta Naturae, 3, 22.

55. Vetting, M. W., Bareich, D. C., Yu, M. and Blanchard, J. 2008, Protein Sci., 17, 1781.
56. 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, PLoS One, 9, e94816.
57. Chan, C. H., Garrity, J., Crosby, H. A. and Escalante-Semerena, J. C. 2011, Mol. Microbiol., 80, 168.
58. Christensen, D. G., Meyer, J. G., Baumgartner, J. T., D'Souza, A. K., Nelson, W. C., Payne, S. H., Kuhn, M. L., Schilling, B. and Wolfe, A. J. 2018, MBio, 9, e01905-18.
59. Pathak, D., Bhat, A. H., Sapehia, V., Rai, J. and Rao, A. 2016, Sci. Rep., 6, 28892.
60. Hou, M., Zhuang, J., Fan, S., Wang, H., Guo, C., Yao, H., Lin, D. and Liao, X. 2019, Acta Biochim. Biophys. Sin. (Shanghai), 51, 960.
61. Tanaka, S., Matsushita, Y., Yoshikawa, A. and Isono, K. 1989, Mol. Gen. Genet., 217, 289.
62. Terhorst, C., Moller, W., Laursen, R. and Wittmann-Liebold, B. 1973, Eur. J. Biochem., 34, 138.
63. Miao, L., Fang, H., Li, Y. and Chen, H. 2007, Biochem. Biophys. Res. Commun., 357, 641.
64. Vetting, M. W., de Carvalho, L. P., Roderick, S. L. and Blanchard, J. S. 2008, J. Biol. Chem., 280, 22108.
65. Ramagopal, S. and Subramanian, A. R. 1974, Proc. Natl. Acad. Sci. USA, 71, 2136.
66. Gordiyenko, Y., Deroo, S., Zhou, M., Videler, H. and Robinson, C. V. 2008, J. Mol. Biol., 380, 404.
67. Griaznova, O. and Traut, R. R. 2000, Biochemistry, 39, 4075.
68. Parks, A. R. and Escalante-Semerena, J. C. 2020, Proc. Natl. Acad. Sci. USA, 117, 15895.
69. Arai, K., Clark, B. F., Duffy, L., Jones, M. D., Kaziro, Y., Laursen, R. A., L'Italien, J., Miller, D. L., Nagarkatti, S., Nakamura, S., Nielsen, K. M., Petersen, T. E., Takahashi, K. and Wade, M. 1980, Proc. Natl. Acad. Sci. USA, 77, 1326.
70. Champion, P. A., Champion, M. M., Manzanillo, P. and Cox, J. S. 2009, Mol. Microbiol., 73, 950.
71. Li, Y., Champion, M. M., Sun, L., Champion, P. A., Wojcik, R. and Dovichi, N. J. 2012, Anal. Chem., 84, 1617.
72. Mba Medie, F., Champion, M. M., Williams, E. A. and Champion, P. A. 2014, Infect. Immun., 82, 4572.
73. Okkels, L. M., Muller, E. C., Schmid, M., Rosenkrands, I., Kaufmann, S. H., Andersen, P. and Jungblut, P. R. 2004, Proteomics, 4, 2954.
74. Dedieu, A., Sahinovic, E., Guerin, P., Blanchard, L., Fochesato, S., Meunier, B., de Groot, A. and Armengaud, J. 2013, Proteome Sci., 11, 3.
75. Ouidir, T., Jarnier, F., Cosette, P., Jouenne, T. and Hardouin, J. 2015, J. Proteomics, 114, 214.
76. Zhang, X., Ye, J., Engholm-Keller, K. and Hojrup, P. 2011, Proteomics, 11, 81-93.
77. Bienvenut, W. V., Giglione, C. and Meinnel, T. 2017, Methods Mol. Biol., 1574, 17.
78. Chenau, J., Michelland, S., Sidibe, J. and Seve, M. 2008, Proteome Sci., 6, 9.
79. Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W. and Selbach, M. 2011, Nature, 473, 337.
80. Kentache, T., Jouenne, T., De, E. and Hardouin, J. 2016, J. Proteomics, 144, 148.
81. Thompson, C. R., Champio, M. M. and Champio, P. A. 2018, J. Proteome Res., 17, 3246.
82. Gevaert, K., Goethals, M., Martens, L., Van Damme, J., Staes, A., Thomas, G. R. and Vandekerckhove, J. 2003, Nat. Biotechnol., 21, 566.
83. Kouyianou, K., De Bock, P. J., Colaert, N., Nikolaki, A., Aktoudianaki, A., Gevaert, K. and Tsiotis, G. 2012, Proteomics, 12, 63.
84. Ree, R., Varland, S. and Arnesen, T. 2018, Exp. Mol. Med., 50, 90.
85. Staes, A., Impens, F., Van Damme, P., Ruttens, B., Goethals, M., Demol, H., Timmerman, E., Vandekerckhove, J. and Gevaert, K. 2011, Nat. Protoc., 6, 1130.

86. Staes, A., Van Damme, P., Helsens, K., Demol, H., Vandekerckhove, J. and Gevaert, K. 2008, Proteomics, 8, 1362.
87. Van Damme, P., Arnesen, T. and Gevaert, K. 2011, FEBS J., 278, 3822.
88. Van Damme, P., Hole, K., Pimenta-Marques, A., Helsens, K., Vandekerckhove, J., Martinho, R. G., Gevaert, K. and Arnesen, T. 2011, PLoS Genet., 7, e1002169.
89. Zhang, X. and Hojrup, P. 2013, Methods Mol. Biol., 981, 47.
90. Helbig, A. O., Gauci, S., Raijmakers, R., van Breukelen, B., Slijper, M., Mohammed, S. and Heck, A. J. 2010, Mol. Cell. Proteomics, 9, 928.
91. Mohammed, S. and Heck, A. Jr. 2011, Curr. Opin. Biotechnol., 22, 9.
92. Van Damme, P., Van Damme, J., Demol, H., Staes, A., Vandekerckhove, J. and Gevaert, K. 2009, BMC Proc., 3(Suppl. 6), S6.
93. Bernal-Perez, L. F., Prokai, L. and Ryu, Y. 2012, Anal. Biochem., 428, 13.
94. Guo, C., Mi, J., Brautigan, D. L. and Larner, J. M. 2007, Cell Signal, 19, 504.
95. Kraker, A. J., Mizzen, C. A., Hartl, B. G., Miin, J., Allis, C. D. and Merriman, R. L. 2003, Mol. Cancer Ther., 2, 401.
96. Verdin, E., Dequiedt, F., Fischle, W., Frye, R., Marshall, B. and North, B. 2004, Methods Enzymol., 377, 180.
97. Riddles, P. W., Blakeley, R. L. and Zerner, B. 1983, Methods Enzymol., 91, 49.
98. Ellman, G. L., Courtney, K. D., Andres, V. Jr. and Feather-Stone, R. M. 1961, Biochem. Pharmacol., 7, 88.
99. Foyn, H., Thompson, P. R. and Arnesen, T. 2017, Methods Mol. Biol., 1574, 9.
100. Van Damme, P., Evjenth, R., Foyn, H., Demeyer, K., De Bock, P. J., Lillehaug, J. R., Vandekerckhove, J., Arnesen, T. and Gevaert, K. 2011, Mol. Cell. Proteomics, 10, M110 004580.