The Causes and Consequences of Non-Enzymatic Protein Acylation

Andrew M. James[1,*], Cassandra L. Smith[1], Anthony C. Smith[1], Alan J. Robinson[1], Kurt Hoogewijs[1,2] and Michael P. Murphy[1,3,*]

[1]Medical Research Council Mitochondrial Biology Unit, University of Cambridge, Cambridge, CB2 0XY, UK

[2] Department of Organic and Macromolecular Chemistry, Ghent University, Krijgslaan 281-S4, B-9000 Ghent, Belgium

[3]Lead contact

[*]Correspondence: aj@mrc-mbu.cam.ac.uk; mpm@mrc-mbu.cam.ac.uk

Thousands of protein acyl modification sites have now been identified *in vivo*. However, at most sites the acylation stoichiometry is low, making functional enzyme-driven regulation in the majority of cases unlikely. As unmediated acylation can occur on the surface of proteins when acyl-CoA thioesters react with nucleophilic cysteine and lysine residues, slower non-enzymatic processes likely underlie most protein acylation. Here we review how non-enzymatic acylation of nucleophilic lysine and cysteine residues occurs, the factors that enhance acylation at particular sites, and the strategies that have evolved to limit protein acylation. We conclude that protein acylation is an unavoidable consequence of the central role of reactive thioesters in metabolism. Finally, we propose a hypothesis for why low-stoichiometry protein acylation is selected against by evolution and how it might contribute to degenerative processes such as aging.

Introduction

Numerous different acyl-CoA species arise in metabolism (Supplementary Table 1), and many have been shown by mass spectrometry (MS) to generate corresponding acyl-lysine modifications on proteins *in vivo* [1-7]. However, with the exception of those catalyzing histone acetylation, few enzymes that transfer acyl groups to protein have been identified [8]. Recent work has challenged the importance of enzymatic protein acylation at most sites, as the observed median stoichiometry of *N*-acetylation is ~0.1% [9, 10]. The lack of enzymes that transfer acyl groups to protein, the broad range of acyl-lysine modifications, the low stoichiometry of acylation, and the vast number of acylation sites suggest that most lysine *N*-acylation arises non-enzymatically. The mechanism and stoichiometry of protein acylation

are crucial for understanding links between deacylating sirtuin enzymes and degenerative processes such as aging [11-13], as high stoichiometry enzymatic acylation of specific sites may be regulatory, while low stoichiometry non-enzymatic acylation of many sites suggests a stress [14, 15].

Kinetics and thermodynamics of non-enzymatic acylation

To understand the likelihood of non-enzymatic protein acylation, first consider the law of mass action which is composed of two key parts: first the reaction rate is proportional to the product of the reactant concentrations; and the ratio between the concentration of reactants and products is constant at equilibrium. The first part dictates that the rate of non-enzymatic protein acylation depends upon the local acyl-CoA and nucleophile concentrations. The second part is relevant as energy from the continual oxidation of carbohydrates, fats and amino acids causes a displacement from equilibrium of couples such as the acyl-CoA/CoA ratio. The chemical potential energy in these couples creates a directional flux as acyl-CoAs are synthesized by certain reactions and consumed by others. Although metabolic flux largely flows through enzymatically catalyzed pathways, the chemical potential energy of acyl-CoA/CoA couples can also drive slower non-enzymatic side reactions. Thus, the key question is not whether slower non-enzymatic acyl-transfer reactions to protein residues can occur [16, 17], but whether they are fast enough to occur within biological timeframes [18]. For example, carbonic anhydrase enhances the rate of CO_2 hydration a billion-fold, but the half-life of the non-enzymatic process is still only 5 s. In contrast, the non-enzymatic half-life of some biologically important reactions is millions of years making enzymatic catalysis essential [18]. Non-enzymatic acyl-transfer from acyl-CoAs to protein residues falls between these extremes with the *in vivo* half-life of reversible *S*→*S*-transfer of an acyl moiety to another thiol ~10 $M^{-1}s^{-1}$ [16]. This suggests meaningful reversible *S*→*S*-transfer occurs within minutes to hours while irreversible *S*→*N*-transfer of an acyl moiety to an amine takes hours to days [16, 18-20].

CoA and acyl-CoAs

Non-enzymatic protein acylation by acyl-CoAs depends on their cellular concentrations, and these are intimately linked to that of CoA itself, which is synthesized from ATP, cysteine and pantothenate (vitamin B5) by five essential enzymes [21, 22]. Pantothenate biosynthesis is dispensable in animals, which have a pantothenate transporter for uptake from the gut [21,

22]. The first committed step in CoA biosynthesis is the ATP-dependent conversion of pantothenate to cell-impermeant 4-phosphopantothenate, which is catalyzed by four isoforms of pantothenate kinase (PanK) [23]. The external supply of pantothenate, the level of expression of each PanK isoform and feedback inhibition of PanK by CoA and its thioesters collectively control CoA biosynthesis [24]. The size of the combined pool of CoA and its acylated forms varies widely among mammalian tissues, being largest in the liver at ~150 µM [25, 26]. Its subcellular distribution reflects its diverse roles, and ranges from 20-140 µM in the cytosol, 700 µM in peroxisomes and 2.2-5 mM in the mitochondria [27-30]. Non-enzymatic protein acylation is also influenced by how much the acyl-CoA/CoA ratio is displaced from equilibrium. Usually ~10-50% of CoA is *S*-acylated *in vivo* [29, 31], but this will vary depending on cell compartment, tissue and substrate availability.

Acetyl-CoA is the product of several major metabolic pathways including pyruvate, fatty acid and branched-chain amino acid oxidation [32] and these reactions predominantly take place in the mitochondrial matrix. Acetyl-CoA is also generated in the cytosol where ATP is used to make acetyl-CoA from acetate and cleave citrate into oxaloacetate and acetyl-CoA [32]. Many other acyl-CoAs exist in metabolism, but most participate in fewer reactions and thus tend to be compartmentalized within the cell (Supplementary Table 1). This includes succinyl-CoA, an intermediate in the tricarboxylic acid cycle within the mitochondrial matrix; malonyl-CoA, a precursor for fatty acid synthesis in the cytosol; and very-long chain fatty acyl-CoAs, which are shortened in the peroxisome before transfer to mitochondria to be broken down to acetyl-CoA. There are also many minor acyl-CoAs that are intermediates in other metabolic pathways, such as amino acid catabolism (Supplementary Table 1). MS has identified corresponding acyl-lysine modifications for many acyl-CoAs [1-7, 33], but dedicated transferases have remained elusive raising the possibility such enzymes do not exist.

The chemistry of non-enzymatic lysine *N*-acylation

Protein acylation, whether it is enzymatic or non-enzymatic, is a nucleophilic acyl substitution reaction where a nucleophile attacks the carbonyl of a thioester such as an acyl-CoA. This implies non-enzymatic protein acylation is inevitable *in vivo*, because of the ubiquity of nucleophiles on protein surfaces. However, the rate of non-enzymatic protein lysine *N*-acylation at individual sites will be relatively slow as only ~0.1% of solvent-exposed lysine residues are in the nucleophilic, deprotonated form (pK_a ~10.5) in the cytosol. Acyltransferases can circumvent this constraint using active site ionizable residues to keep

lysines deprotonated (Figure 1A). While individual surface lysine residues are less reactive than corresponding active site nucleophiles, as a group they are more abundant. The total concentration of surface-exposed protein lysine residues in a tissue is ~100 mM giving a nucleophilic amine concentration of ~100 μM (see Supplemental Calculations). As the concentration of metabolic enzymes, and thereby their stabilized active site nucleophiles, is generally below 100 μM [34], it is realistic that surface lysine residues as a group can compete with metabolic pathways for acyl-CoAs. This is especially true in the mitochondrial matrix where the higher pH and elevated acyl-CoA and protein concentrations will combine to further favor non-enzymatic acylation [9, 17, 27-29, 33-35].

Abundance and stoichiometry of lysine *N*-acylation

Sites of protein lysine *N*-acylation are identified through trypsinization, followed by enrichment of *N*-acylated peptides with an antibody against the acyl-lysine residue and MS analysis (Box 1) [1-7]. These studies typically generate long lists of lysine residues that are modified *in vivo* and with recent improvements in MS sensitivity several thousand sites can be identified [36]. However, MS still preferentially observes easily detectable peptides on abundant proteins, suggesting current observations underestimate protein *N*-acetylation sites (there are ~700,000 unique lysine residues encoded by the mouse transcriptome and typically ~95% of lysines are solvent exposed) [5, 37, 38]. While there are many such lists of protein acylation sites, identifying those sites used for metabolic regulation is difficult because stoichiometric analysis of the extent of acylation of each site remains rare [36]. Recent analysis of ~4000 lysine *N*-acetylation sites in mouse liver observed only one site *N*-acetylated by >10%, with median *N*-acetylation at each site of just ~0.1% [9]. This low stoichiometry is also observed in *Escherichia coli* and is consistent with calculated values obtained by calibrating *in vivo* acetylation against proteins labelled *in vitro* with ^{14}C acetyl-CoA [10, 20]. The low stoichiometry of lysine *N*-acetylation suggests that at most sites *N*-acylation is stochastic and that regulatory acetylation is the exception rather than the rule.

Cysteine and other nucleophilic protein residues can also be acylated

Protein acylation could occur on other nucleophilic residues, such as serine, threonine, histidine and cysteine. Serine and threonine residues often act as the nucleophiles within enzyme active sites, but are weak nucleophiles in the absence of ionizable groups to aid deprotonation. Even so, the *S*→*O*-transfer of acetyl groups from a thioester to a serine

residue generating an *O*-linked acetyl-serine modification was observed on peptides [20], so esterified serine and threonine acyl modifications might occur *in vivo*. While histidine residues are predominantly in their reactive, deprotonated form (pK_a ~6) within the cell and can in principle be *N*-acylated, their importance *in vivo* may be limited as *N*-acyl-histidine residues are unstable [39].

In contrast, cysteine *S*-acylation may be important as the pK_a (~8.5) of the cysteine thiol means ~10% is in the reactive thiolate form in the cytosol, and the thioester bond of *S*-acylcysteine is moderately stable. Consistent with their relative pK_a values, acetyl-CoA acetylated cysteine residues ~100-fold more rapidly than lysine residues at pH 7.8 *in vitro* [20]. This reactivity may help explain why cysteines residues are sparse on protein surfaces [35, 38, 40] with a tissue concentration of exposed protein cysteine residues of ~20 mM. However, despite the low frequency of cysteine residues the nucleophilic thiolate concentration (~2 mM) is actually greater than for lysine (see Supplemental Calculations). Thus, in the absence of mitigation, non-enzymatic *S*-acylation of cysteine residues will be a significant modification *in vivo*, particularly within mitochondria.

Why then is the literature on cysteine *S*-acylation so limited compared to lysine *N*-acylation? The reason is that non-enzymatic protein *S*-acylation is difficult to explore by MS *in vivo* because the thioester bond can hydrolyze or further react with nucleophiles during overnight trypsinization prior to MS analysis (Box 1). Thus, protein *S*-acylation is observed indirectly through blocking of free thiols, selective removal of the acyl group with hydroxylamine, followed by derivatization of the previously *S*-acylated free thiol [41]. While such chemical deacylation followed by derivatization is suitable for high stoichiometry enzyme-catalyzed sites, low stoichiometry protein *S*-acylation may be difficult to differentiate from chemical artifacts [20, 35]. Consequently, non-enzymatic protein *S*-acylation has largely been explored *in vitro* on peptides where proteolytic cleavage is not required and *S*-acylation can be viewed directly [16, 20, 42]. This suggests significant S→S-transfer of acyl groups between acyl-CoA and protein thiols could be expected *in vivo* over minutes to hours.

Another key difference between non-enzymatic protein *S*-acylation and *N*-acylation is reversibility: protein lysine *N*-acylation forms an amide bond resistant to further nucleophilic attack, while protein *S*-acylation generates another thioester on a surface cysteine residue that another nucleophile can remove. The reaction is reversed if the nucleophile is CoA, but this is unlikely as >90% of cellular thiols are on glutathione or protein cysteine residues [43].

Furthermore, enzymatic processes will rapidly remake acyl-CoA from CoA generated by off-target acyl-transfer reactions to protein. This constant flux of acyl-CoA generated by metabolism would be expected to cause other thiols to become as *S*-acylated as the CoA pool. However, this is not observed suggesting that *S*-acyl groups are removed from non-CoA thiols, probably via enzyme-mediated hydrolysis [20]. Acyl-CoAs are then resynthesized using ATP, suggesting that some metabolic inefficiency is tolerated to minimize off-target protein *S*-acylation, presumably because this is more efficient than using thousands of ATP molecules to degrade and remake the damaged protein.

Factors that enhance protein *N*-acylation at some sites

One mechanism by which enzymes achieve rate enhancement is through binding, thereby raising the local concentrations of their substrates (Figures 1A and 1C). As thioesters form on protein surfaces via *S*→*S*-transfer of an acyl moiety to a cysteine residue it is likely *S*-acyl modifications cause widespread enhanced *N*-acetylation of nearby lysine residues. Several enzymatic and non-enzymatic examples of an initial *S*→*S*-acyl-transfer reaction being used to enhance a slower localized *S*→*N*-acyl-transfer to a nearby lysine residue illustrate the feasibility of this proximity-enhanced reaction (Figure 1). This two-step process was also shown using the peptide RYA**K**G**C**A**S**R, which was readily *S*-acetylated by acetyl-CoA on its cysteine residue. This *S*-acetylation caused nearby lysine and serine residues to be acetylated 23-fold faster than on the peptide RYA**K**G**S**A**S**R that lacks a thiol [20]. The same proximity effect enhances lysine *N*-acetylation *in vivo* if cysteine and lysine residues are closer than ~10 Å on the protein surface [35]. This reaction distance of ~10 Å is consistent with the combined distance of eight bond lengths from the α-carbon to the respective amine and carbonyls atoms of lysine and *S*-acyl-cysteine. This distance and their flexible side-chains imply they interact within a volume of ~10^3 Å^3, thus an amine within range of an *S*-acylcysteine is effectively exposed to a local thioester concentration of ~1 M (see Supplemental Calculations). For comparison, a physiological acetyl-CoA concentration of 0.01-1 mM equates to one thioester per ~10^6-10^8 Å^3. A similar proximity effect is predicted to occur if acyl-CoAs associate with nucleotide binding sites on proteins, because CoA, like NADH, NADPH, FADH and ATP, contains an ADP backbone that is recognized by proteins with a Rossman fold [44].

These proximity effects mean the distinction between enzymatic and non-enzymatic reactions is somewhat arbitrary and protein acylation will often be auto-catalyzed by nearby cysteine residues or nucleotide binding sites (Figure 2, Key Figure).

Low stoichiometry protein acylation is a stress with detrimental consequences

The above analysis indicates non-enzymatic acylation of exposed protein residues by acyl-CoAs occurs readily. Consequently, protein acylation in the absence of sufficient deacylation may represent a 'carbon stress' in much the same way that oxidative stress occurs when reactive oxygen species (ROS) generation overwhelms cellular antioxidant systems [14, 15]. This hypothesis is attractive as acyl-CoAs, unlike ROS, are central metabolites present at relatively high concentrations that cannot be aggressively degraded. It also fits with work showing *E. coli* CobB deacetylase depresses already low stoichiometry lysine *N*-acetylation yet further [10]. However, it remained unclear whether low stoichiometry *N*-acylation at a myriad of sites had detrimental consequences [14]. This was recently addressed by analyzing lysine residues with a proximal cysteine residue closer than ~10 Å, as these lysine residues exhibit a higher degree of *N*-acetylation *in vivo* [35]. If lysine acylation is detrimental such sites should experience greater selective pressure and mutation of either the lysine or cysteine residue should prevent any deleterious effect [35]. It was shown that mutation of these low stoichiometry (<1%) sites of lysine auto-acetylation in mouse liver occurred almost twice as frequently in 66 vertebrate genomes than non-acetylated sites from the same proteins [35]. Linking low stoichiometry lysine *N*-acylation with negative selection at hundreds of sites strongly supports the carbon stress hypothesis. Intriguingly, the species with the longest maximum lifespan exhibited lower conservation of pairs of nearby cysteine and *N*-acetylated lysine residues [35]. This strong correlation with lifespan ($p < 0.0001$) only held for cytosolic proteins despite increased *N*-acetylation of proteins in the mitochondrial matrix [35]. Thus, low stoichiometry lysine *N*-acylation, not associated with any particular acylation site, may be linked to the wide variation in maximum lifespan between species [45, 46].

Evolved strategies to limit or reverse non-specific protein acylation by thioesters

Given that low stoichiometry lysine *N*-acylation appears to be a stress [10, 14, 15, 35], it is expected that mechanisms have evolved to limit off-target protein acylation (Figure 3). The simplest way to do this is to limit the free thioester concentration by binding acyl-CoAs to carrier proteins, of which humans have three [47, 48]. Acyl groups can also transit through the cytosol to mitochondria as less reactive *O*-linked acyl-carnitines that do not cause protein

lysine *N*-acetylation [20]. Supporting this protective role, skeletal muscle from mice deficient in carnitine acetyltransferase (CrAT) had increased tissue acetyl-CoA levels and susceptibility to diet-induced *N*-acetylation of a range of mitochondrial proteins [49]. Another mechanism for limiting acyl-CoA concentration is to hydrolyze it to its cognate carboxylic acid and CoA. Despite this process being inefficient, with reconstitution requiring ATP, there are 15 acyl-CoA thioesterase (ACoT) genes in mouse and 12 in humans [50]. While this suggests the acyl-CoA pool is tightly regulated, the role of ACoTs is still poorly understood, and their impact on acylation has yet to be assessed.

Non-enzymatic S→S-transfer of acyl groups from acyl-CoA to protein cysteine thiols generates *S*-acyl cysteine residues on proteins [20]. Glutathione is present at millimolar concentrations within tissues and can equilibrate with *S*-acyl groups on proteins via an S→S-transfer reaction, recovering the protein thiol and forming *S*-acyl-glutathione [16, 20]. The thioesterase, hydroxyacyl glutathione hydrolase (HAGH, also called glyoxalase II), can then rapidly hydrolyze *S*-acyl-glutathione [20]. This is necessary as *S*-acetyl-glutathione also causes unmediated *N*-acetylation of lysine residues *in vitro* [20, 51]. By keeping the *S*-acyl-glutathione concentration low, *S*-acyl-groups can be drained from *S*-acyl-cysteine residues and *in vitro* HAGH limits lysine *N*-acetylation by acetyl-CoA in the presence of glutathione [20]. As the acyl group is bound to glutathione and HAGH recognizes only glutathione, one enzyme can indirectly limit a broad range of *S*- and *N*-acyl modifications of disparate targets [20].

HAGH is ineffective once an *N*-acyl lysine modification forms on a protein surface, and instead several acyl-specific sirtuin deacetylases (Sirt) remove *N*-acyl lysine modifications from proteins and this has been extensively reviewed [52, 53]. In humans Sirt1, Sirt2 and Sirt3 all have efficient deacetylase activity that extends to propionyl-, butyryl- and myristoyl-lysine residues. Sirt4 and Sirt5 shows increased specificity towards negatively-charged succinyl-, malonyl-, glutaryl- and (3-methyl)glutaryl lysine modifications [54, 55]. Sirt6 shows selectivity for myristoyl- and palmitoyl-lysine residues [56], while Sirt7 has been reported as a desuccinylase [57]. While much work has focussed on the sirtuin family, a further 11 Zn-dependent deacetylases can also remove acyl-modifications [58].

Hypothesis: Low stoichiometry acylation prevents effective protein turnover

While strategies have evolved to limit lysine *N*-acylation and the majority of it appears detrimental [20, 35, 49, 53], it is unclear how low stoichiometry lysine *N*-acylation of

proteins at a myriad of sites damages the organism. As lysine *N*-acylation changes a positively-charged side-chain into one that is hydrophobic or negatively-charged, high stoichiometry enzymatic acylation of a single lysine residue can inactivate enzymes, alter interacting partners or change protein localization. Transferase-catalyzed examples of this mode of regulation include histone acylation which limits their binding to negatively-charged DNA and thereby increases transcription [8], and long-chain acylation of a single cysteine or lysine residue to anchor a protein to a membrane [59]. However, these regulatory changes can only occur because a high proportion of the protein is modified. Negative selection against hundreds of sites of low stoichiometry protein acylation is not consistent with regulation and instead implies that there is a common problem introduced by *N*-acylation of proteins [14, 15, 35]. A plausible hypothesis is that lysine acylation increases hydrophobicity favoring protein misfolding and aggregation (Figure 4A) and several observations support this possibility. Increased protein *N*-acetylation in *E. coli* generated larger inclusion bodies that were processed more slowly, while decreased *N*-acetylation led to smaller inclusion bodies and their faster processing [60]. *N*-acetylation of lysine residues triggers aggregation of Tau protein leading to neurodegenerative disease [61], and acetylation is decreased by mutation of a nearby auto-catalytic cysteine residue [62] (Figure 1D). Tracts of glutamine, a similar but less hydrophobic mimetic of *N*-acetyllysine, lead to insoluble protein aggregates in PolyQ disorders such as Huntington's Disease [63]. Histone deacetylase inhibitors increase protein aggregation and bromodomains that bind acetylated protein are enriched in these amyloid aggregates [58, 64]. In addition to increased hydrophobicity, *N*-acylation has another effect that might decrease protein turnover and cause aggregation. Degradation of cytosolic proteins by the proteasome predominantly occurs by ubiquitination of lysine residues on target proteins (Figure 1C). If a lysine is *N*-acylated it cannot be ubiquitinated, thus lysine *N*-acylation could prevent a modified protein from being degraded by the proteasome [65]. That mitochondrial matrix proteins are turned over at the level of the organelle rather than individually could explain why it is overexpression of the cytosolic/nucleic sirtuins, Sirt1 and Sirt 6, that extends lifespan in mice [12, 66]. It is also consistent with knockdown of ATP-citrate lyase, which generates acetyl-CoA for fatty acid synthesis in the cytosol, extending lifespan in flies [67], and the observation maximum lifespan correlates with cytosolic, but not mitochondrial matrix *N*-acylation [35, 52, 68].

While detrimental pairs of proximal cysteine and lysine residues were identified from a protein *N*-acetylation dataset [9, 35], other non-enzymatic lysine modifications could cause negative selection at these sites and thereby underlie the correlation with maximum lifespan

[69]. The most problematic could be longer-chain acyl-CoA species as these would generate very hydrophobic acyl-lysine modifications at the identified sites through proximity-based acyl-transfer reactions (Figure 4A). This is consistent with a similar negative correlation of maximum lifespan with plasma long-chain free fatty acid concentrations [70] and lifespan extension with overexpression of the long-chain deacylase, Sirt6 [12, 13].

This hypothesis fits with the long history of dietary restriction extending lifespan [71], that this may require autophagy [72], and that autophagy is linked to cytosolic acetyl-CoA levels [73, 74]. While limiting non-enzymatic acylation by acyl-CoAs could link dietary restriction and lifespan, many questions remain (see Outstanding Questions).

Conclusions

Given that any thioester can non-enzymatically acylate a primary amine such as the side-chain of a lysine residue and all living things use thioesters in metabolism, it would appear non-enzymatic protein acylation is a very ancient process. This implies the vast majority of lysine *N*-acylation predates regulation and is simply a consequence of cellular protein being bathed in reactive thioesters displaced from equilibrium. This non-enzymatic acylation would provide a need for general deacylation mechanisms which in turn could facilitate the evolution of enzymatic acylation at certain sites as a regulatory modification (Figure 4B). We believe that protein acylation should be viewed within this context - many deacetylases will be enzymes that keep the surface of other proteins clean and genuine regulatory acetylation should be considered the exception rather than the rule.

Acknowledgements

This work was supported by the Medical Research Council UK (MC_U105663142) and by a Wellcome Trust Investigator award (110159/Z/15/Z) to MPM. K.H. has received funding from the FWO and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665501, and Fonds Wetenschappelijk Onderzoek-Vlaanderen research grant 1.5.193.18N.

Figure 1. *S*→*S*-acyl-transfer followed by a proximity-enhanced *S*→*N*-acyl-transfer are features of both non-enzymatic and enzymatic reactions. A, enzymatic histone acetylation by ESA1 occurs through an active site *S*-acetyl-cysteine intermediate [75]. B, non-enzymatic

native chemical ligation of two peptides [76]. An N-terminal cysteine thiol on peptide 1 (green) attacks a thioester on the C-terminus of peptide 2 (gray) forming a thioester bond between peptides 1 and 2. The proximity of the N-terminal amine favors an *S*→*N*-acyl transfer and formation of a stable amide bond between peptides 1 and 2. C, enzymatic ubiquitination of proteins for degradation [77]. Ubiquitin (Ub) is first conjugated to E1 via a thioester before *S*→*S*-acyl transfer to E2. E3 generates a proximity effect by localizing *S*-ubiquinated-E2 near a target protein. An *S*→*N*-acyl transfer forms a stable amide bond between ubiquitin and the target protein. D, Tau protein is auto-acetylated. Acetyl-CoA acetylates C291 and C322 and these auto-acetylate K280 [62]. Acetylation of K280 leads to Tau aggregation [61].

Key Figure 2. There is no clear distinction between non-enzymatic and enzymatic acylation. Acylation at many sites will be enhanced by auto-catalysis and similar reactions could be considered enzymatic or non-enzymatic simply because the *N*-acylated lysine residue resides on a different protein.

Figure 3. Non-enzymatic acylation. Reactive *S*-linked CoA-thioesters (red boxes), such as acetyl-CoA, are generated by metabolism. These are transformed into stable *O*-linked carnitines by CrAT or hydrolysed to free acids by ACoTs to prevent them react non-enzymatically with other nucleophiles. If reactive *S*-linked thioesters form on protein cysteine residues they can exchange with glutathione and be hydrolyzed to the free acid by HAGH. If the acetyl group irreversibly transfers to a protein lysine residue then Sirt deacetylases can remove it at the expense of NAD.

Figure 4. A, a hypothetical mechanism for low stoichiometry non-enzymatic acylation to affect lifespan. B, possible evolutionary sequence of acylation and deacylation. Reactive thioesters would have caused slow non-specific non-enzymatic acylation of protein nucleophiles. Positive selection generates general deacylation pathways for repair that are more efficient that protein degradation and resynthesis. General deacylation pathways in turn allow evolution of regulatory enzymatic acylation at certain sites.

Box 1. Measuring acylation

***N*-Acylation –** *Method* - Usually overnight trypsinization followed by enrichment of acylated peptides with an anti-acyllysine antibody. *Advantages* – *N*-acylation is stable so can be

observed directly. *Limitations* – Only detects one modification at a time. No suitable antibodies for long-chain acyl groups. Hard to measure all acylation at a site.

***S*-Acylation –** *Method* - Usually alkylation of free thiols, followed by deacylation with hydroxylamine and alkylation of previously acylated thiols with a tag. This is followed by overnight trypsinization and enrichment tagged peptides with an antibody. *Advantages* - Can measure all acylation at a site. *Limitations* – *S*-acylation is not stable so can only be observed indirectly. Hydroxylamine breaks all thioester bonds so no information about the type of acyl modification. For low stoichiometry acylation, thiol alkylation must be complete and hydroxylamine must not react with other thiol modifications.

References

1. Chen, Y. et al. (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. Mol Cell Proteomics 6, 812-9.
2. Tan, M. et al. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell 146, 1016-28.
3. Peng, C. et al. (2011) The first identification of lysine malonylation substrates and its regulatory enzyme. Mol Cell Proteomics 10, M111.012658.
4. Dai, L. et al. (2014) Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. Nat Chem Biol 10, 365-70.
5. 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, 842-51.
6. Tan, M. et al. (2014) Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. Cell Metab 19, 605-17.
7. Wagner, G.R. et al. (2017) A class of reactive acyl-CoA species reveals the non-enzymatic origins of protein acylation. Cell Metab 25, 823-837 e8.
8. Marmorstein, R. and Zhou, M.M. (2014) Writers and readers of histone acetylation: structure, mechanism, and inhibition. Cold Spring Harb Perspect Biol 6, a018762.
9. Weinert, B.T. et al. (2015) Analysis of acetylation stoichiometry suggests that SIRT3 repairs nonenzymatic acetylation lesions. EMBO J 34, 2620-32.
10. Weinert, B.T. et al. (2017) Accurate quantification of site-specific acetylation stoichiometry reveals the impact of sirtuin deacetylase CobB on the E. coli acetylome. Mol Cell Proteomics 16, 759-769.
11. Lin, S.J. et al. (2000) Requirement of NAD and SIR2 for life-span extension by calorie restriction in Saccharomyces cerevisiae. Science 289, 2126-8.
12. Kanfi, Y. et al. (2012) The sirtuin SIRT6 regulates lifespan in male mice. Nature 483, 218-21.
13. Pan, H. and Finkel, T. (2017) Key proteins and pathways that regulate lifespan. J Biol Chem 292, 6452-6460.
14. Trub, A.G. and Hirschey, M.D. (2018) Reactive acyl-CoA species modify proteins and induce carbon stress. Trends Biochem Sci.

15. Wagner, G.R. and Hirschey, M.D. (2014) Nonenzymatic protein acylation as a carbon stress regulated by sirtuin deacylases. Mol Cell 54, 5-16.
16. Bizzozero, O.A. et al. (2001) Structural determinants influencing the reaction of cysteine-containing peptides with palmitoyl-coenzyme A and other thioesters. Biochim Biophys Acta 1545, 278-88.
17. Wagner, G.R. and Payne, R.M. (2013) Widespread and enzyme-independent N{epsilon}-acetylation and N{epsilon}-succinylation in the chemical conditions of the mitochondrial matrix. J Biol Chem 288, 29036-45.
18. Wolfenden, R. and Snider, M.J. (2001) The depth of chemical time and the power of enzymes as catalysts. Acc Chem Res 34, 938-45.
19. Bracher, P.J. et al. (2011) The relative rates of thiol-thioester exchange and hydrolysis for alkyl and aryl thioalkanoates in water. Orig Life Evol Biosph 41, 399-412.
20. James, A.M. et al. (2017) Non-enzymatic N-acetylation of lysine residues by acetylCoA often occurs via a proximal S-acetylated thiol intermediate sensitive to glyoxalase II. Cell Rep 18, 2105-2112.
21. Gerdes, S.Y. et al. (2002) From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. J Bacteriol 184, 4555-72.
22. Leonardi, R. et al. (2005) Coenzyme A: back in action. Prog Lipid Res 44, 125-53.
23. Rock, C.O. et al. (2002) The murine pantothenate kinase (Pank1) gene encodes two differentially regulated pantothenate kinase isozymes. Gene 291, 35-43.
24. Leonardi, R. and Jackowski, S. (2007) Biosynthesis of pantothenic acid and coenzyme A. EcoSal Plus 2.
25. Shibata, K. et al. (2012) Simultaneous high-performance liquid chromatography determination of coenzyme A, dephospho-coenzyme A, and acetyl-coenzyme A in normal and pantothenic acid-deficient rats. Anal Biochem 430, 151-5.
26. Dansie, L.E. et al. (2014) Physiological roles of the pantothenate kinases. Biochem Soc Trans 42, 1033-6.
27. Idell-Wenger, J.A. et al. (1978) Coenzyme A and carnitine distribution in normal and ischemic hearts. J Biol Chem 253, 4310-8.
28. Horie, S. et al. (1986) Changes in CoA pools in hepatic peroxisomes of the rat under various conditions. J Biochem 99, 1345-52.
29. Garland, P.B. et al. (1965) Steady-state concentrations of coenzyme A, acetyl-coenzyme A and long-chain fatty acyl-coenzyme A in rat-liver mitochondria oxidizing palmitate. Biochem J 97, 587-94.
30. Park, J.O. et al. (2016) Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. Nat Chem Biol 12, 482-9.
31. Tubbs, P.K. and Garland, P.B. (1964) Variations in tissue contents of coenzyme A thioesters and possible metabolic implications. Biochem J 93, 550-7.
32. Pietrocola, F. et al. (2015) Acetyl coenzyme A: a central metabolite and second messenger. Cell Metab 21, 805-21.
33. Carrico, C. et al. (2018) The mitochondrial acylome emerges: Proteomics, regulation by sirtuins, and metabolic and disease implications. Cell Metab 27, 497-512.
34. Srere, P.A. (1967) Enzyme concentrations in tissues. Science 158, 936-7.
35. James, A.M. et al. (2018) Proximal cysteines that enhance lysine N-acetylation of cytosolic proteins in mice are less conserved in longer-lived species. Cell Reports In Press.
36. Baeza, J. et al. (2016) Mechanisms and dynamics of protein acetylation in mitochondria. Trends Biochem Sci 41, 231-44.
37. Menger, K.E. et al. (2015) Fasting, but not aging, dramatically alters the redox status of cysteine residues on proteins in Drosophila melanogaster. Cell Rep 11, 1856-65.

38. Schein, C.H. (1990) Solubility as a function of protein structure and solvent components. BioTechnology 8, 308-317.
39. Jencks, W.P. and Carriuolo, J. (1959) Imidazole catalysis. II. Acyl transfer and the reactions of acetyl imidazole with water and oxygen anions. J Biol Chem 234, 1272-9.
40. Trinquier, G. and Sanejouand, Y.H. (1998) Which effective property of amino acids is best preserved by the genetic code? Protein Eng 11, 153-69.
41. Gould, N.S. et al. (2015) Site-specific proteomic mapping identifies selectively modified regulatory cysteine residues in functionally distinct protein networks. Chem Biol 22, 965-75.
42. Bano, M.C. et al. (1998) Pseudo-enzymatic S-acylation of a myristoylated yes protein tyrosine kinase peptide in vitro may reflect non-enzymatic S-acylation in vivo. Biochem J 330, 723-31.
43. Requejo, R. et al. (2010) Cysteine residues exposed on protein surfaces are the dominant intramitochondrial thiol and may protect against oxidative damage. FEBS J 277, 1465-80.
44. Tsuchiya, Y. et al. (2017) Protein CoAlation: a redox-regulated protein modification by coenzyme A in mammalian cells. Biochem J 474, 2489-2508.
45. Ma, S. and Gladyshev, V.N. (2017) Molecular signatures of longevity: Insights from cross-species comparative studies. Semin Cell Dev Biol 70, 190-203.
46. Tacutu, R. et al. (2013) Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing. Nucleic Acids Res 41, D1027-33.
47. Knudsen, J. et al. (2000) Role of acyl-CoA binding protein in acyl-CoA metabolism and acyl-CoA-mediated cell signaling. J Nutr 130, 294S-298S.
48. Burton, M. et al. (2005) Evolution of the acyl-CoA binding protein (ACBP). Biochem J 392, 299-307.
49. Davies, M.N. et al. (2016) The acetyl group buffering action of carnitine acetyltransferase offsets macronutrient-induced lysine acetylation of mitochondrial proteins. Cell Rep 14, 243-54.
50. Tillander, V. et al. (2017) Deactivating fatty acids: Acyl-CoA thioesterase-mediated control of lipid metabolism. Trends Endocrinol Metab 28, 473-484.
51. Weinert, B.T. et al. (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in E. coli. Mol Cell 51, 265-72.
52. Lin, S. et al. (2018) Sirtuins in mitochondrial stress: Indispensable helpers behind the scenes. Ageing Res Rev 44, 22-32.
53. Bheda, P. et al. (2016) The substrate specificity of sirtuins. Annu Rev Biochem 85, 405-29.
54. Du, J. et al. (2011) Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. Science 334, 806-9.
55. Anderson, K.A. et al. (2017) SIRT4 Is a Lysine Deacylase that Controls Leucine Metabolism and Insulin Secretion. Cell Metab 25, 838-855 e15.
56. Jiang, H. et al. (2013) SIRT6 regulates TNF-alpha secretion through hydrolysis of long-chain fatty acyl lysine. Nature 496, 110-3.
57. Li, L. et al. (2016) SIRT7 is a histone desuccinylase that functionally links to chromatin compaction and genome stability. Nat Commun 7, 12235.
58. Liao, D. (2017) CBP/p300 Bromodomain Mediates Amyloid Formation. Cell Chem Biol 24, 128-129.
59. Jiang, H. et al. (2018) Protein lipidation: Occurrence, mechanisms, biological functions, and enabling technologies. Chem Rev.
60. Kuczynska-Wisnik, D. et al. (2016) The effect of protein acetylation on the formation and processing of inclusion bodies and endogenous protein aggregates in Escherichia coli cells. Microb Cell Fact 15, 189.

61. Cohen, T.J. et al. (2011) The acetylation of tau inhibits its function and promotes pathological tau aggregation. Nat Commun 2, 252.
62. Cohen, T.J. et al. (2013) The microtubule-associated tau protein has intrinsic acetyltransferase activity. Nat Struct Mol Biol 20, 756-62.
63. Ross, C.A. and Poirier, M.A. (2004) Protein aggregation and neurodegenerative disease. Nat Med 10 Suppl, S10-7.
64. Olzscha, H. et al. (2017) CBP/p300 bromodomains regulate amyloid-like protein aggregation upon aberrant lysine acetylation. Cell Chem Biol 24, 9-23.
65. Valdez-Taubas, J. and Pelham, H. (2005) Swf1-dependent palmitoylation of the SNARE Tlg1 prevents its ubiquitination and degradation. EMBO J 24, 2524-32.
66. Satoh, A. et al. (2013) Sirt1 extends life span and delays aging in mice through the regulation of Nk2 homeobox 1 in the DMH and LH. Cell Metab 18, 416-30.
67. Peleg, S. et al. (2016) Life span extension by targeting a link between metabolism and histone acetylation in Drosophila. EMBO Rep 17, 455-69.
68. Pickles, S. et al. (2018) Mitophagy and quality control mechanisms in mitochondrial maintenance. Curr Biol 28, R170-R185.
69. Moellering, R.E. and Cravatt, B.F. (2013) Functional lysine modification by an intrinsically reactive primary glycolytic metabolite. Science 341, 549-53.
70. Jove, M. et al. (2013) Plasma long-chain free fatty acids predict mammalian longevity. Sci Rep 3, 3346.
71. Piper, M.D.W. and Partridge, L. (2017) Drosophila as a model for ageing. Biochim Biophys Acta In press, 10.1016/j.bbadis.2017.09.016.
72. Hansen, M. et al. (2008) A role for autophagy in the extension of lifespan by dietary restriction in C. elegans. PLoS Genet 4, e24.
73. Marino, G. et al. (2014) Regulation of autophagy by cytosolic acetyl-coenzyme A. Mol Cell 53, 710-25.
74. Eisenberg, T. et al. (2014) Nucleocytosolic depletion of the energy metabolite acetyl-coenzyme a stimulates autophagy and prolongs lifespan. Cell Metab 19, 431-44.
75. Yan, Y. et al. (2002) The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. Nat Struct Biol 9, 862-9.
76. Thapa, P. et al. (2014) Native chemical ligation: a boon to peptide chemistry. Molecules 19, 14461-83.
77. Stewart, M.D. et al. (2016) E2 enzymes: more than just middle men. Cell Res 26, 423-40.

Outstanding Questions

-Non-enzymatic protein acylation depends on the concentrations of a large range of acyl-CoA species in various tissues and compartments. However, much of the scant available data is more than 50 years old. How do the absolute concentration of various acyl-CoA species respond to interventions, such as dietary restriction, in different tissues?

-Many degenerative phenotypes affect post-mitotic tissues such as muscle and brain, yet acetylation has generally been studied in regenerative tissues such as liver. Is the low acetylation stoichiometry observed in liver also observed in post-mitotic tissues?

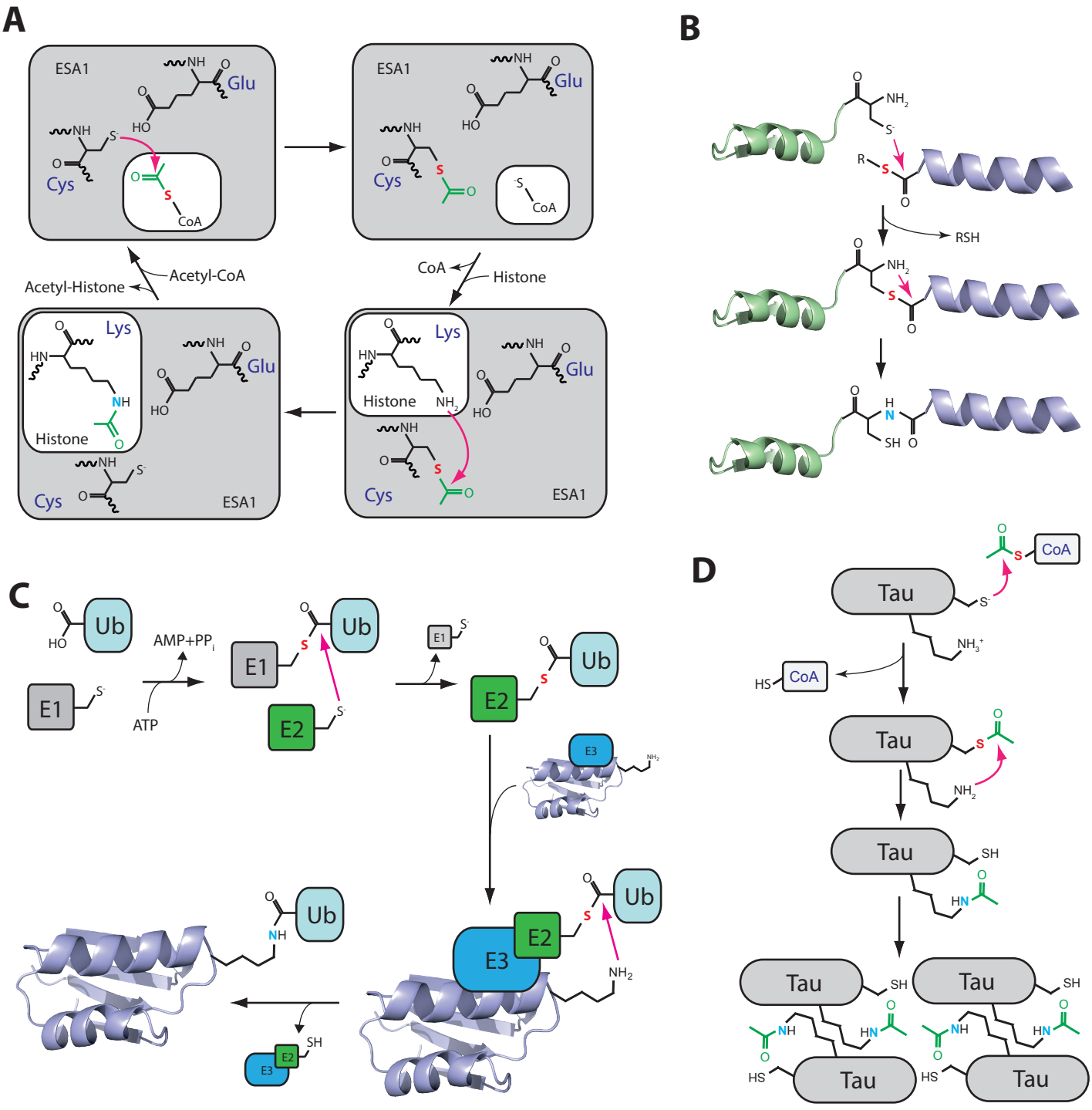
-While the literature contains many lists of acylation sites, acylation stoichiometry at these sites is generally not known. How prevalent are other non-enzymatic acyl modifications such as succinylation, malonylation, and long-chain fatty acylation? Is their stoichiometry high enough to regulate cellular function?

-Several classes of enzymes and proteins that could limit non-enzymatic acylation have not been studied in this context. For example, there are 12 acyl-CoA thioesterases, 3 acyl-CoA binding proteins and 11 Zn-dependent deac(et)ylases encoded by the human genome. Are they important in limiting non-enzymatic acylation?

-Most sites of acetylation in mouse are less conserved within other vertebrate genomes. What makes acylated protein deleterious and what acyl modifications are the most deleterious?

Figure 1

James et al. Figure 1



Figure

James et al. Key Figure 2

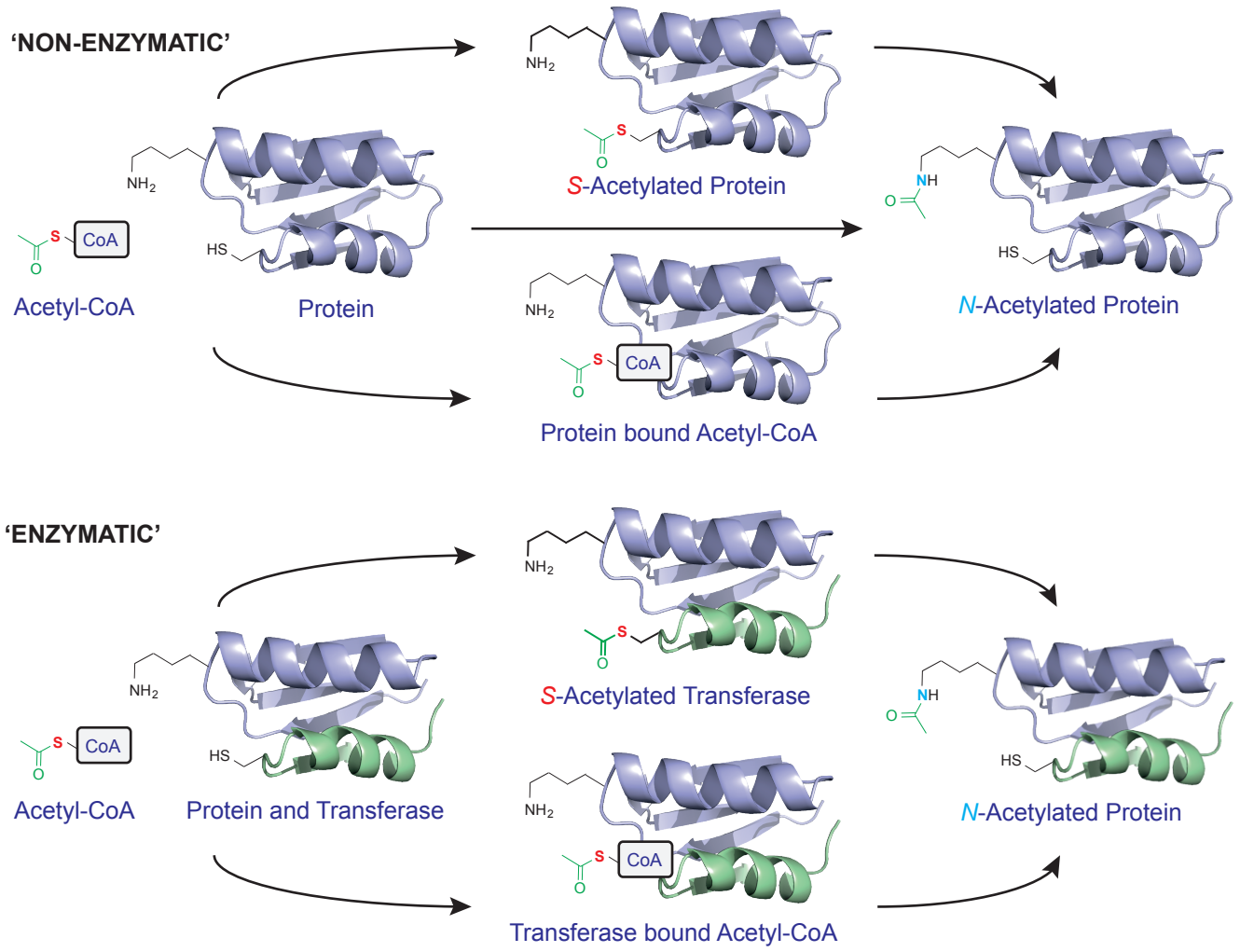


Figure 3

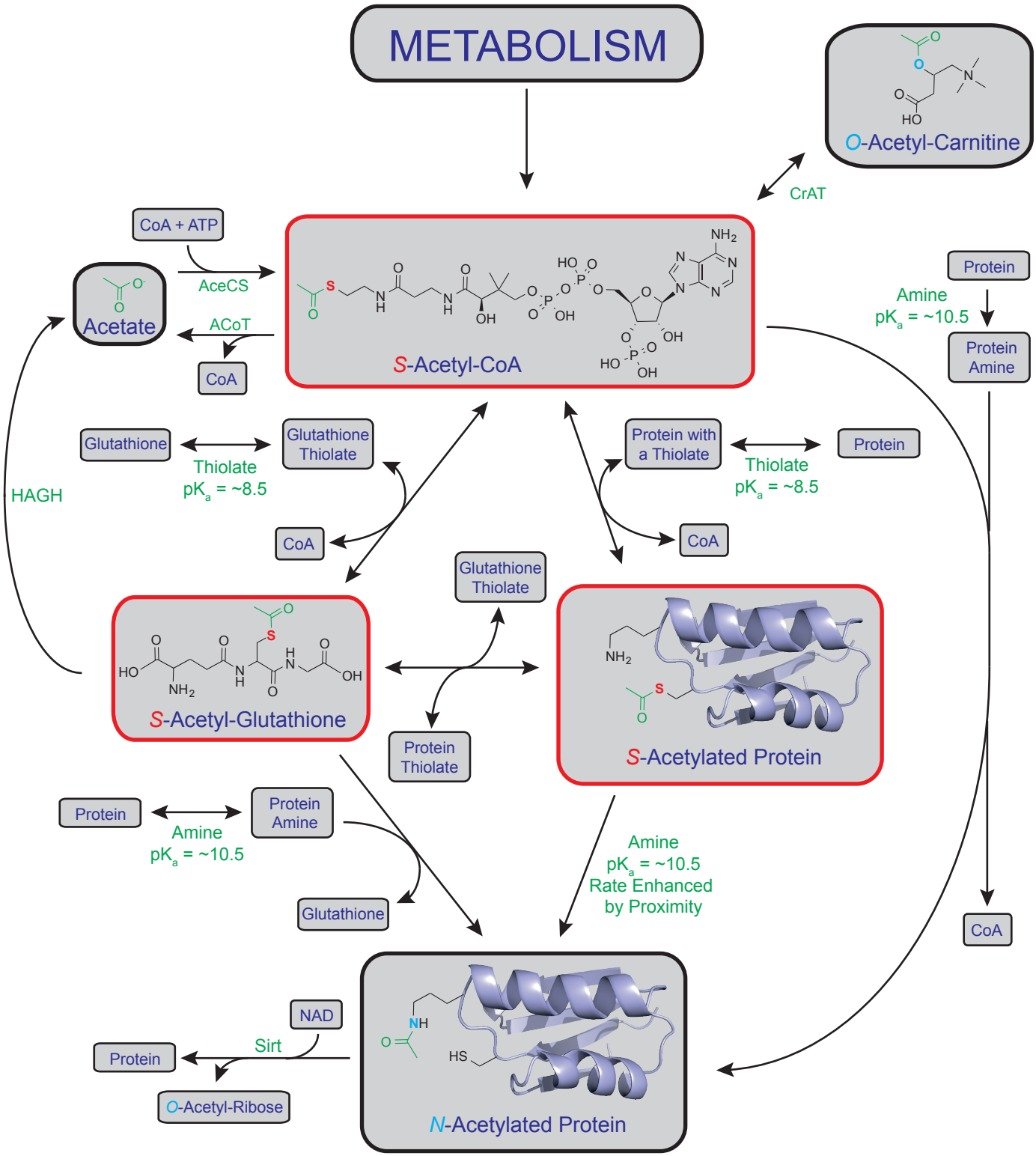
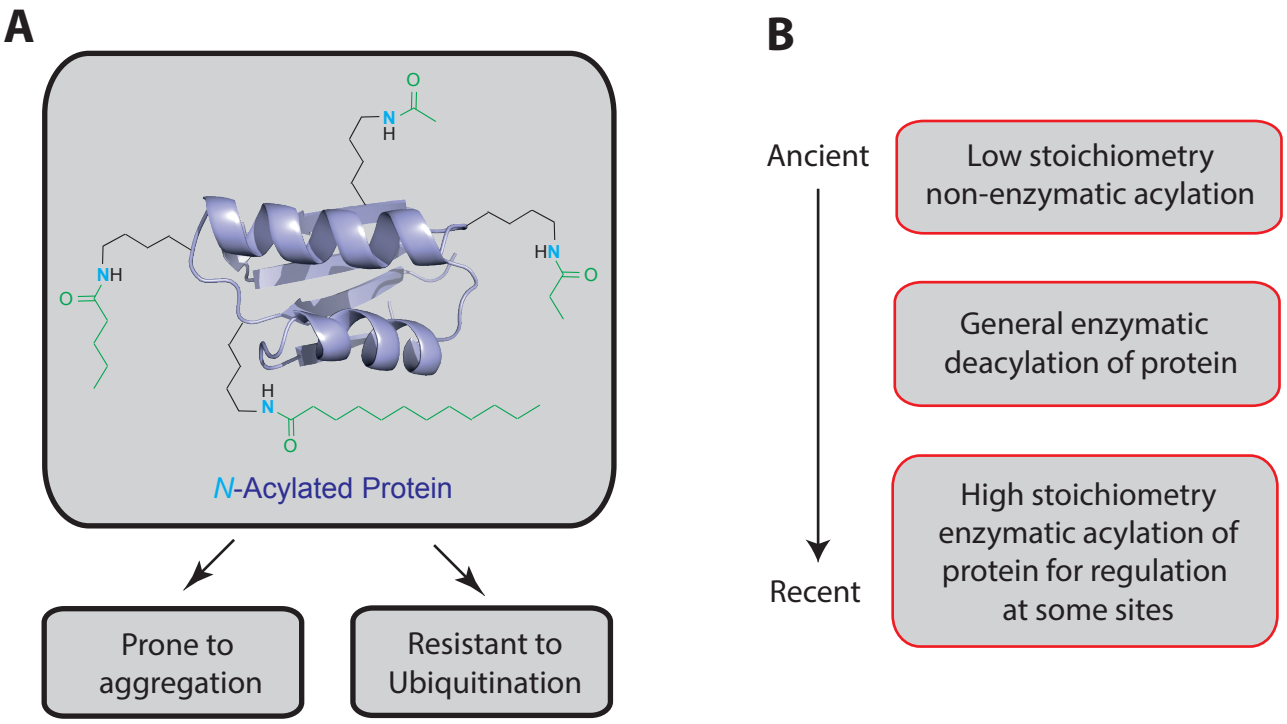

James et al. Figure 3
METABOLISM
O-Acetyl-Carnitine
CrAT
CoA + ATP
AceCS
Acetate
ACoT
CoA
S-Acetyl-CoA
Protein
Amine
pKa = ~10.5
Protein Amine
Glutathione
Thiolate
pKa = ~8.5
Glutathione Thiolate
HAGH
Protein with a Thiolate
Thiolate
pKa = ~8.5
Protein
CoA
CoA
Glutathione Thiolate
S-Acetyl-Glutathione
S-Acetylated Protein
Protein Thiolate
Protein
Amine
pKa = ~10.5
Protein Amine
Glutathione
Amine
pKa = ~10.5
Rate Enhanced
by Proximity
CoA
NAD
Sirt
Protein
O-Acetyl-Ribose
N-Acetylated Protein

Figure 4

James et al. Figure 4



Supplementary Table

Click here to access/download

Author Supplementary Material

SupplementaryTable1.xlsx

Supplementary calculations

Highlights

-Nucleophilic protein residues can react with the thioester carbonyl of acyl-CoAs leading to the continuous transfer acyl groups to protein. Here we discuss the chemistry of this reaction.

-The cell has a range of mitigating strategies that limit the accumulation of acyl modifications on protein. Here we discuss known mechanisms for limiting acyl modifications as well as additional proteins that could have a role in preventing acylation.

-Sites of acylation are less conserved in vertebrates as well as longer-lived mammalian species suggesting low-level protein acylation is a stress. Here we discuss how low stoichiometry acylation could have deleterious consequences.