Nε-Lysine Acetylation Control Conserved in All Three Life Domains

The relative simplicity of studying microbes could prove critical for understanding this posttranslational modification system

Jorge C. Escalante-Semerena

Working with microbes provides opportunities to better understand fundamental cellular processes, including some that, although discovered in eukaryotes, have their origins in Bacteria or Archaea. One such example, reversible Nε-lysine (Nε-Lys) acetylation of proteins, appears crucial for all organisms. With that universality in mind, the relative simplicity of bacteria and archaea could prove critical for advancing our understanding of how this posttranslational modification system influences cell survival.

Nε-Lys acetylation could well prove to be a fundamental process for cells from all three domains of life. If true, this control of protein function likely arose early during evolution.

Chemical changes can control the activities of both small and macromolecules in cells, rapidly halting or enhancing a wide range of physiological processes. These control mechanisms are particularly important to microorganisms, especially those in habitats subject to sudden changes in composition, temperature, and pH. These modification mechanisms can prove pivotal for microbial stress response systems that enable microbes to adapt to and survive in such environments.

How Chemical Change-Based Cellular Control Systems Work

Many biochemical change systems typically depend on transferase enzymes to modify target molecules. Such modifications include alkylation, acylation, phosphorylation, thiolation, or nitrosylation of specific chemical groups, including hydroxyls, sulfhydryls, amines, or carbon skeletons. Those changes can exert positive or negative effects on the biological functions of the target molecules. In some cases, the changes lead to turnover of the modified target molecule.

Such enzymes modify all kinds of molecules in cells. Some of them, for example, account for particular kinds of antibiotic resistance. Thus, in some cases, transferases of bacterial pathogens modify the core structure of an antibiotic—e.g., acetylating chloramphenicol or phosphorylating kanamycin—rendering them inactive.

Other enzymes modify DNA, RNA, tRNA, or proteins. One type of those protein modifications—the reversible acetylation of the ε-amino

Summary

• Reversible Nε-lysine (Nε-Lys) acetylation of proteins appears crucial throughout three domains of biology.
• Enzymes that acetylate or deacetylate histone tails help to control gene expression in eukaryotes and archaea; similar processes may also affect gene expression in some bacteria.
• Altogether about 90 proteins in E. coli are acetylated, and more than 50% of them appear to be involved in protein synthesis, catabolism, or energy metabolism.
• The number of Gcn-5 acetyltransferase enzymes—GNATs—in a particular bacterial species appears to reflect its metabolic complexity.
• Bacterial proteins called sirtuins are class III, NAD⁺-consuming histone deacetylases that appear to partake in metabolic stress responses.
group of lysine ($N^\epsilon$-Lys acetylation)—proves crucial for regulating gene expression among eukaryotes.

For instance, the enzymes that acetylate or deacetylate histone tails help to fine tune gene expression in eukaryotes and archaea. Thus, positively charged, hypoacetylated histone tails interact with other chromatin components, silencing genes. $N^\epsilon$-Lys acetylation neutralizes those charges, yielding a less-condensed nucleosome that facilitates transcriptional read-through.

Whether $N^\epsilon$-Lys acetylation affects the DNA-binding activity of bacterial nucleoid-associated proteins is uncertain. However, the catabolism repressor protein (Crp) and the tryptophan biosynthesis repressor (TrpR) from *Escherichia coli* are acetylated in vivo, according to recent reports by Y. Zhao (now at the University of Chicago) and coworkers and collaborators, and by J.-G. Pan and coworkers at the Korea Research Institute of Bioscience and Biotechnology at Daejon. Other transcription factors might also fall under $N^\epsilon$-Lys acetylation control in this or other bacteria and archaea. Similar studies have been performed in *Salmonella enterica*. A group of investigators from several research centers in China and the United States recently reported the potentially broad impact of $N^\epsilon$-Lys acetylation in *Salmonella enterica* physiology.

### $N^\epsilon$-Lys Acetylation Could Control Many Cellular Processes

Altogether about 90 proteins in *E. coli* are acetylated, and more than 50% of them appear to be involved either in protein synthesis, including ribosomal proteins; catabolism, including proteins that are part of the pentose phosphate pathway, glycolysis, and the tricarboxylic acid (TCA) cycle; or energy metabolism, including proteins involved in synthesizing menaquinone and ubiquinone. The acetylation state of these proteins varies with growth stage, suggesting that these systems respond to physiological signals.

The reactivity of acetyl-CoA, which can acetylate proteins on its own, complicates efforts to measure enzyme-dependent, reversible $N^\epsilon$-Lys acetylations in cells. For example, we find that many proteins from *E. coli* and *Salmonella enterica* can undergo autoacetylation, even under conditions that minimize the lability of the thioester bond in acetyl-CoA. Because we measure these enzyme-independent events under chemically defined conditions with homogeneous proteins, we need to be cautious when ascertaining whether any specific protein is a substrate for a particular acetyltransferase. The role of protein autoacetylation is not known.

### What We Know about Acetyltransferase Systems in Microbes

Eukaryotes produce different classes of histone acetyltransferases (HATs) and deacetylases (HDACs). Bacterial homologues of the yeast Gcn-5 acetyltransferase (yGNAT)—called GNATs—are common. Although the catalytic core of GNATs is conserved, their numbers vary in bacteria. For example, the *E. coli* K-12 genome encodes 23 putative GNATS, while the *Rhodopseudomonas palustris* genome encodes about 40, the *Bacillus subtilis* genome about 50, and the *Streptomyces coelicolor* genome about 100. Apparently, bacteria acylate many target molecules to control their availability or reactivity. Thus the number of GNATs in a bacterial species appears to reflect its metabolic complexity.

Our understanding of the contributions of GNATs to bacterial and archaeal cell physiol-
ogy is limited. For example, in *E. coli*, there are experiments to describe the functions of only 10 GNAT-encoding genes, namely *speG*, *rimI*, *rimJ*, *rimL*, *yfiQ*, *wecD*, *phnO*, *argA*, *aat*, and *tmcA*, leaving the other 13 GNATs wholly enigmatic. In contrast, the numbers of homologues of deacetylases in *E. coli* and *Salmonella enterica* are much smaller, and we know more about them. These bacteria use Zn(II)-dependent deacetylases in breaking down amino sugars and for producing ornithine, lipid A, and macromolecules.

However, except in the case of the sirtuins, deacetylases are not implicated as part of protein modification systems in *E. coli* or *S. enterica*. *Bacillus subtilis* is different. For instance, a class II, acetate-forming deacetylase is part of a system in *B. subtilis* that modulates the activity of acetyl-CoA synthetase.

### Bacterial, Sirtuin-Dependent Protein Acetylation System

Bacterial proteins called sirtuins are class III histone deacetylases that are homologues of the yeast Sir2 protein (ySir2p). For instance, the *cobB* gene in both *E. coli* and *S. enterica* encodes a sirtuin, CobB sirtuin, that is involved in short-chain fatty acid catabolism. The cognate GNAT that works in concert with CobB in *S. enterica* is protein acetyltransferase (Pat), encoded by *pat* (formerly, *yfiQ*).

CobB and Pat are components of the sirtuin-dependent protein acylation-deacylation system (SDPADS), which controls acetyl-CoA synthetase (Acs) by reversible acetylation and propionyl-CoA synthetase (PrpE) by reversible propionylation (Fig. 1). SDPADS appears to respond to metabolic stress when CoA becomes imbalanced and redox homeostasis is disrupted.

Pat deactivates both Acs and PrpE by acylating conserved lysine residues in the active sites of these enzymes, presumably in response to a decrease in the CoA:acyl-CoA ratio. Acyl-CoA could build up as its use for biosynthesis or generating energy decreases (Fig. 1). Under such conditions, the Pat enzyme of SDPADS would slow down, preventing further imbalances in CoA. Acyl-CoA would revert to physiological levels when it is consumed or if thioesterases hydrolyze its thioester bond. Once CoA homeostasis is reestablished, the cell must reactivate acylated acyl-CoA synthetases to build acyl-CoA levels for growth.

Like other deacetylases, sirtuins require Zn(II) ions. However, unlike other deacetylases, sirtuins...
use NAD$^+$ as a substrate, not as a cofactor, meaning that sirtuins consume NAD$^+$. The biochemical function of sirtuins was identified in 2000 when investigators learned that these enzymes use NAD$^+$ to deacetylate acetylated proteins. Free acetate is not a product of the sirtuin reaction. Instead, sirtuins synthesize $2\text{-O-}Ac\text{AADPr}$, whose role in bacterial and archaeal cell physiology remains an open question.

The link of sirtuin function to eukaryotic cell longevity, cancer, and other human diseases makes it important not only for the scientific community but also the general public. Studying sirtuins in microbes will likely reveal basic principles that may well apply to cells in other domains of life.

**Roles of N$^\epsilon$-Lys Acetylation in Bacteria and Archaea**

We still know relatively little about reversible N$^\epsilon$-Lys acetylation in bacteria such as *S. enterica* serovar Typhimurium LT2, *E. coli* K-12, *B. subtilis* SMY, *Rhodopseudomonas palustris*, and archaea such as *Halofex volcanii*, *Sulfobus sulfataricus* P2, and *Archaeglobus fulgidus*.

Reversible N$^\epsilon$-Lys acetylation was linked to bacterial metabolism by observing how sirtuin-deficient *S. enterica* strains grow when fed short-chain fatty acids. For instance, such strains fail to use acetate or propionate as carbon and energy source, even though such cells carry enzymes needed for catabolizing these short-chain fatty acids. Apparently, when a Lys residue in the active site of acyl-CoA synthetases is acetylated, the enzyme is no longer catalytic, thus blocking use of those fatty acids. This modification interferes with the first step of the reaction, which ordinarily consumes ATP to form an acyl-adenylate intermediate, releasing pyrophosphate.

Critical to this step is a conserved Lys residue that orients the acid substrate in the active site. Binding of CoA to the active site in acetyl-CoA synthetase (Acs) rotates its C-terminal domain, removing that Lys and exposing it to the medium, thus making it available to the acetyltransferase, according to structural studies (Fig. 2). Although the acylation site is part of a motif found in this class of enzymes, we do not understand how GNATs recognize residues within or flanking that motif. More-
Expanding Our Understanding of GNAT Functions in Cells

GNATs continue to attract interest because of their involvement in several key processes, including microbial antibiotic resistance, compacting eukaryotic DNA, and controlling gene expression. However, GNATs play other roles, including in biochemical metabolism, but the details of these other functions are mostly unknown.

In elucidating gene functions, valuable clues come from databases such as http://genexpdb.ou.edu/index.php and http://smd.stanford.edu/resources/databases.shtml. Genome context is also useful, especially if a GNAT-encoding gene might be co-expressed with others of known function, or if global or local regulators regulate its expression. Another way to approach GNAT functions involves exploring diverse stress conditions that microbes may face in particular environments where posttranslational modification systems confer on microorganisms the capacity to turn genes on or off to respond to changing environmental stimuli.

Phenotypic analyses can be extremely useful, providing an in vivo context for gene functions. The search for a phenotype typically involves testing strains that either carry a null allele of the gene of interest or a plasmid encoding that gene to look at what happens when gene dosage increases. These approaches disrupt optimal levels of a protein, with the goal of revealing a substantial negative or positive effect in the behavior of the strain under specific conditions.

In the absence of any in vivo information, high-throughput in vitro approaches also offer opportunities to study GNAT functions. This approach involves constructing proteome chips similar to those used to study *Saccharomyces cerevisiae*. One approach would be to probe for putative protein substrates of GNATs in the presence of radiolabeled acetyl-CoA. One advantage of this approach is that it probes as many of the expressed and modified proteins as possible. An important consideration for the use of this approach is that proteins may be attached to surfaces in ways that block interactions with GNATs, leading to false negatives.

ACKNOWLEDGEMENTS

This work was supported by NIH grant R01 GM62203 to J. C. E.-S.

SUGGESTED READING


