UC Santa Cruz

UC Santa Cruz Previously Published Works

Title

Identification of Novel Protein Lysine Acetyltransferases in Escherichia coli.

Permalink

https://escholarship.org/uc/item/3ss6r8q2

Journal

mBio, 9(5)

Authors

Christensen, David

Meyer, Jesse

Baumgartner, Jackson

et al.

Publication Date

2018-10-23

DOI

10.1128/mBio.01905-18

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

RESEARCH ARTICLE
Molecular Biology and Physiology

Identification of Novel Protein Lysine Acetyltransferases in *Escherichia coli*

David G. Christensen,[a] Jesse G. Meyer,[b] Jackson T. Baumgartner,[c] Alexandria K. D'Souza,[b] William C. Nelson,[d] Samuel H. Payne,[d] Misty L. Kuhn,[c] Birgit Schilling,[b] Alan J. Wolfe[a]

[a]Department of Microbiology and Immunology, Stritch School of Medicine, Health Sciences Division, Loyola University Chicago, Maywood, Illinois, USA
[b]Buck Institute for Research on Aging, Novato, California, USA
[c]Department of Chemistry and Biochemistry, San Francisco State University, San Francisco, California, USA
[d]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, USA

ABSTRACT Posttranslational modifications, such as *N*ε-lysine acetylation, regulate protein function. *N*ε-lysine acetylation can occur either nonenzymatically or enzymatically. The nonenzymatic mechanism uses acetyl phosphate (AcP) or acetyl coenzyme A (AcCoA) as acetyl donor to modify an *N*ε-lysine residue of a protein. The enzymatic mechanism uses *N*ε-lysine acetyltransferases (KATs) to specifically transfer an acetyl group from AcCoA to *N*ε-lysine residues on proteins. To date, only one KAT (YfiQ, also known as Pka and PatZ) has been identified in *Escherichia coli*. Here, we demonstrate the existence of 4 additional *E. coli* KATs: RimI, YiaC, YjaB, and PhnO. In a genetic background devoid of all known acetylation mechanisms (most notably AcP and YfiQ) and one deacetylase (CobB), overexpression of these putative KATs elicited unique patterns of protein acetylation. We mutated key active site residues and found that most of them eliminated enzymatic acetylation activity. We used mass spectrometry to identify and quantify the specificity of YfiQ and the four novel KATs. Surprisingly, our analysis revealed a high degree of substrate specificity. The overlap between KAT-dependent and AcP-dependent acetylation was extremely limited, supporting the hypothesis that these two acetylation mechanisms play distinct roles in the posttranslational modification of bacterial proteins. We further showed that these novel KATs are conserved across broad swaths of bacterial phylogeny. Finally, we determined that one of the novel KATs (YiaC) and the known KAT (YfiQ) can negatively regulate bacterial migration. Together, these results emphasize distinct and specific nonenzymatic and enzymatic protein acetylation mechanisms present in bacteria.

IMPORTANCE *N*ε-Lysine acetylation is one of the most abundant and important posttranslational modifications across all domains of life. One of the best-studied effects of acetylation occurs in eukaryotes, where acetylation of histone tails activates gene transcription. Although bacteria do not have true histones, *N*ε-lysine acetylation is prevalent; however, the role of these modifications is mostly unknown. We constructed an *E. coli* strain that lacked both known acetylation mechanisms to identify four new *N*ε-lysine acetyltransferases (RimI, YiaC, YjaB, and PhnO). We used mass spectrometry to determine the substrate specificity of these acetyltransferases. Structural analysis of selected substrate proteins revealed site-specific preferences for enzymatic acetylation that had little overlap with the preferences of the previously reported acetyl-phosphate nonenzymatic acetylation mechanism. Finally, YiaC and YfiQ appear to regulate flagellum-based motility, a phenotype critical for pathogenesis of many organisms. These acetyltransferases are highly conserved and reveal deeper and more complex roles for bacterial posttranslational modification.

Received 31 August 2018 **Accepted** 18 September 2018 **Published** 23 October 2018

Citation Christensen DG, Meyer JG, Baumgartner JT, D'Souza AK, Nelson WC, Payne SH, Kuhn ML, Schilling B, Wolfe AJ. 2018. Identification of novel protein lysine acetyltransferases in *Escherichia coli*. mBio 9:e01905-18. https://doi.org/10.1128/mBio.01905-18.

Editor Nancy E. Freitag, University of Illinois at Chicago

Copyright © 2018 Christensen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Alan J. Wolfe, awolfe@luc.edu.

This article is a direct contribution from a Fellow of the American Academy of Microbiology. Solicited external reviewers: Greg Somerville, University of Nebraska-Lincoln; Larry Reitzer, University of Texas at Dallas.

KEYWORDS acetylation, acetyltransferase, bacteria, GNAT, mass spectrometry, proteomics, protein acetyltransferase, lysine acetyltransferase

During *N*ε-lysine acetylation, an acetyl group is added to the epsilon amino group of a lysine residue of a protein, which neutralizes the positive charge and increases the size of the side chain. In eukaryotes, the effects of protein acetylation have been well described, historically and most fully in the context of histone tail acetylation, which regulates eukaryotic transcription. Over the past decade, it has become clear that many nonhistone proteins are also *N*ε-lysine acetylated and that this posttranslational modification is similarly abundant in archaea and bacteria (1–4).

In *Escherichia coli*, two mechanisms of *N*ε-lysine acetylation are known. The predominant mechanism is a nonenzymatic donation of an acetyl group from the high-energy central metabolite acetyl-phosphate (AcP) onto a susceptible lysine of a protein (5, 6). AcP is the intermediate of the phosphotransferase (Pta)-acetate kinase (AckA) pathway that interconverts acetyl-CoA (AcCoA), inorganic phosphate, and ADP to acetate, CoA, and ATP. Alternatively, an *N*ε-lysine acetyltransferase (KAT) can catalyze acetylation of a specific lysine using AcCoA as the acetyl donor. In *E. coli*, only one KAT has been discovered to date, YfiQ (also called Pka and PatZ) (7–9). This KAT was first discovered in *Salmonella enterica*, where it is called Pat; elegant studies showed that it inactivates acetyl-CoA synthetase (Acs) by acetylation, preventing acetate assimilation (8). It is conserved in many other bacterial species, including *E. coli* (4). Global acetylation profiles of Δ*yfiQ* mutants provide evidence that YfiQ might acetylate targets in addition to Acs (5, 6). Other studies report that YfiQ can acetylate RNA polymerase, RNase R, RNase II, and DnaA on lysines that alter function (10–13). Transcription of *yfiQ* depends on activation by cyclic AMP (cAMP)-bound catabolite activator protein (CAP; also known as the cAMP receptor protein [CRP]) (14). During growth in minimal glucose medium, it is upregulated in stationary phase. However, another report using *E. coli* strain MG1655 in YT medium showed that YfiQ protein levels are reduced in stationary phase (15). Thus, there is more to learn about the regulation of YfiQ and consequence of acetylation.

YfiQ is just one member of the Gcn5-related *N*-acetyltransferase (GNAT) family. GNATs acetylate a broad range of substrates, including antibiotics, polyamines, amino acids, nucleotides, tRNAs, proteins, and peptides (4, 16). The *E. coli* K-12 genome carries 26 genes whose products are annotated as GNATs, of which only about half have annotated functions (4, 17, 18). Therefore, we investigated whether other *E. coli* GNATs beyond YfiQ could function as KATs.

Using an *E. coli* strain that lacked both known acetylation mechanisms (YfiQ and AcP), we found four GNATs that increased relative site-specific acetylation levels on proteins *in vivo* as one would expect for a KAT. By Western immunoblotting and mass spectrometric analyses, we demonstrated that these four GNATs facilitate increased *N*ε-lysine acetylation and identified their cognate substrates. This increased acetylation was lost upon mutation of conserved catalytic or active site residues found in other known GNATs. We conclude that *E. coli* encodes multiple KATs that exhibit substrate specificities that differ from nonenzymatic acetylation by AcP.

RESULTS

Identification of putative uncharacterized KATs. If acetylation in *E. coli* depends solely on the two known mechanisms of protein acetylation, then we would be unable to detect acetylation in a mutant that lacks (i) the only known *E. coli* acetyltransferase, YfiQ, and (ii) the ability to generate AcP, by deleting either Pta or the entire Pta-AckA pathway. However, anti-acetyllysine Western blot analysis of a strain that lacked both of these mechanisms (Δ*ackA pta yfiQ*) revealed that residual acetylation remained (see Fig. S1 in the supplemental material). We therefore sought the mechanism(s) behind this residual AcP- and YfiQ-independent acetylation and hypothesized that this residual acetylation activity could be attributed to uncharacterized KATs. Since YfiQ is a member

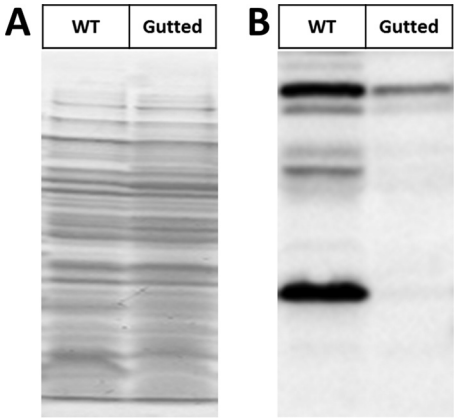


FIG 1 Inactivation of the two known acetylation mechanisms in *E. coli* eliminates the majority of acetylation. Wild-type (WT) *E. coli* (strain BW25113) and an isogenic Δ*pta yfiQ acs cobB* mutant (Gutted) were aerated in TB7 supplemented with 0.4% glucose for 10 h. Whole-cell lysates were analyzed (A) by Coomassie blue-stained SDS-polyacrylamide gel to ensure equivalent loading and (B) by anti-acetyllysine Western blotting.

of the GNAT family of proteins, and *E. coli* contains 25 other members of this family, we tested whether these proteins had KAT activity.

To determine whether these GNATs have KAT activity, we compared acetylation profiles of strains overexpressing each of the GNATs via anti-acetyllysine Western blotting. We used a Δ*pta yfiQ acs cobB* background to enhance the signal-to-noise ratio, which we refer to as the acetylation "gutted" strain. This strain reduces background acetylation levels from AcP and YfiQ (Δ*pta yfiQ*) while ensuring that residual acetylation that occurs would not be reversed by the CobB deacetylase (Δ*cobB*). Acs was also deleted, as it has been reported to acetylate the chemotaxis response regulator CheY (19). Furthermore, YfiQ regulates Acs activity and loss of that control can have a detrimental effect on growth (20). As with the Δ*pta ackA yfiQ* mutant (Fig. S1), the gutted strain (Δ*pta yfiQ acs cobB*) exhibited only limited acetylation (Fig. 1). To validate that this strain behaved as expected and hyperacetylated specific lysine sites with the known KAT YfiQ, we first compared YfiQ overexpression in a gutted strain that expresses the YfiQ substrate Acs (Δ*pta yfiQ cobB*, Acs^+) to that in a gutted strain that does not express Acs (Δ*pta yfiQ cobB*, Acs^-). Indeed, by anti-acetyllysine Western blotting, we observed an acetylated band in the gutted Acs^+ strain that was absent in the gutted Acs^- strain (Fig. 2 and Fig. S2; in Fig. 2, compare lane 1 [positive control] and lane 6 [YfiQ]).

Upon induction of each of the 25 GNAT family members in the gutted strain, overexpression of four GNATs (Aat, ElaA, YiiD, and YafP) inhibited growth. For the 21 strains that did grow, only 8 of the putative GNATs—plus YfiQ—resulted in the appearance of one or more acetylated protein band(s) (Fig. 2). Induction of RimI, YiaC, YjaB, YjgM, and PhnO expression produced a reproducible acetylated protein band(s) (Fig. S3); in contrast, induction of RimJ did not (data not shown). Induction of YncA (17 kDa) and AstA (38.5 kDa) each produced a single acetylated band that migrated consistent with its expected molecular mass, suggesting acetylation of the proteins themselves. We selected RimI, YiaC, YjaB, and PhnO for further assessment of their ability to function as KATs.

Mutation of conserved catalytic amino acids inactivates RimI, YiaC, YjaB, and PhnO. To determine whether these GNATs directly acetylated protein targets, we mutated a few residues that could act as general acids/bases in the reaction or could be important for protein substrate recognition. Acetyltransferases acetylate their substrates using a general acid/base chemical mechanism. Typically, a glutamate (E) or water molecule within a KAT active site acts as the general base by deprotonating the amino group of the substrate. This permits the nitrogen of the amino group to attack

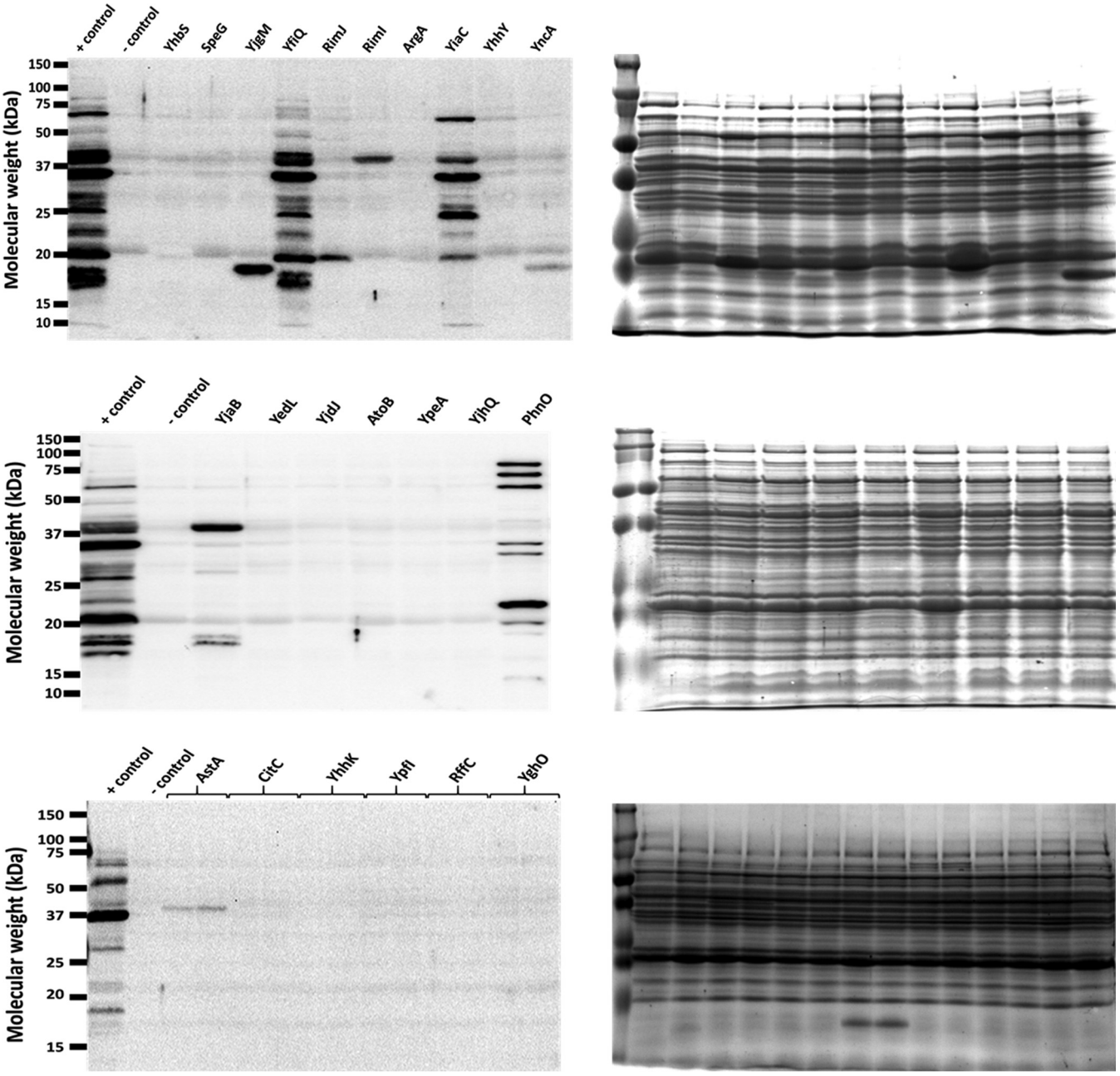


FIG 2 Overexpression of five GNAT family members results in altered lysine acetylation patterns by anti-acetyllysine Western blotting. The gutted strain (BW25113 Δ*pta yfiQ acs cobB*) was transformed with the pCA24n vector control (negative [−] control) or pCA24n containing the indicated genes under an IPTG-inducible promoter (69). As a positive (+) control, an isogenic strain that retained the WT allele of *acs* (Δ*pta yfiQ cobB*) was transformed with pCA24n containing YfiQ. The resulting strains were aerated in TB7 supplemented with 0.4% glucose, 50 μM IPTG, and 25 μg/ml chloramphenicol for 10 h. Whole-cell lysates were analyzed (right panels) by Coomassie blue-stained SDS-polyacrylamide gel electrophoresis to ensure equivalent loading and (left panels) by anti-acetyllysine Western blotting. Note that the band in RimJ was not reproducible. The positive control contains one additional YfiQ-dependent band around 72 kDa, which corresponds to Acs (82). YncA and AstA each produce an acetylated band that can be observed in the Coomassie blue-stained gel at the expected molecular weight of these proteins.

the carbonyl carbon of the acetyl group of AcCoA and results in an acetylated product and CoA anion. An amino acid such as tyrosine (Y) then acts as the general acid to reprotonate the thiolate of CoA (21). To select amino acids for mutagenesis, we compared the putative GNAT sequences using the protein structure prediction tool Phyre2 (22) and then generated the following mutants based on this analysis: YiaC (F70A, Y115A), YjaB (Y117A, Y117F), RimI (Y115A), and PhnO (E78A, Y128A). Plasmids carrying the mutant alleles were introduced into the gutted strain, and the cell lysates

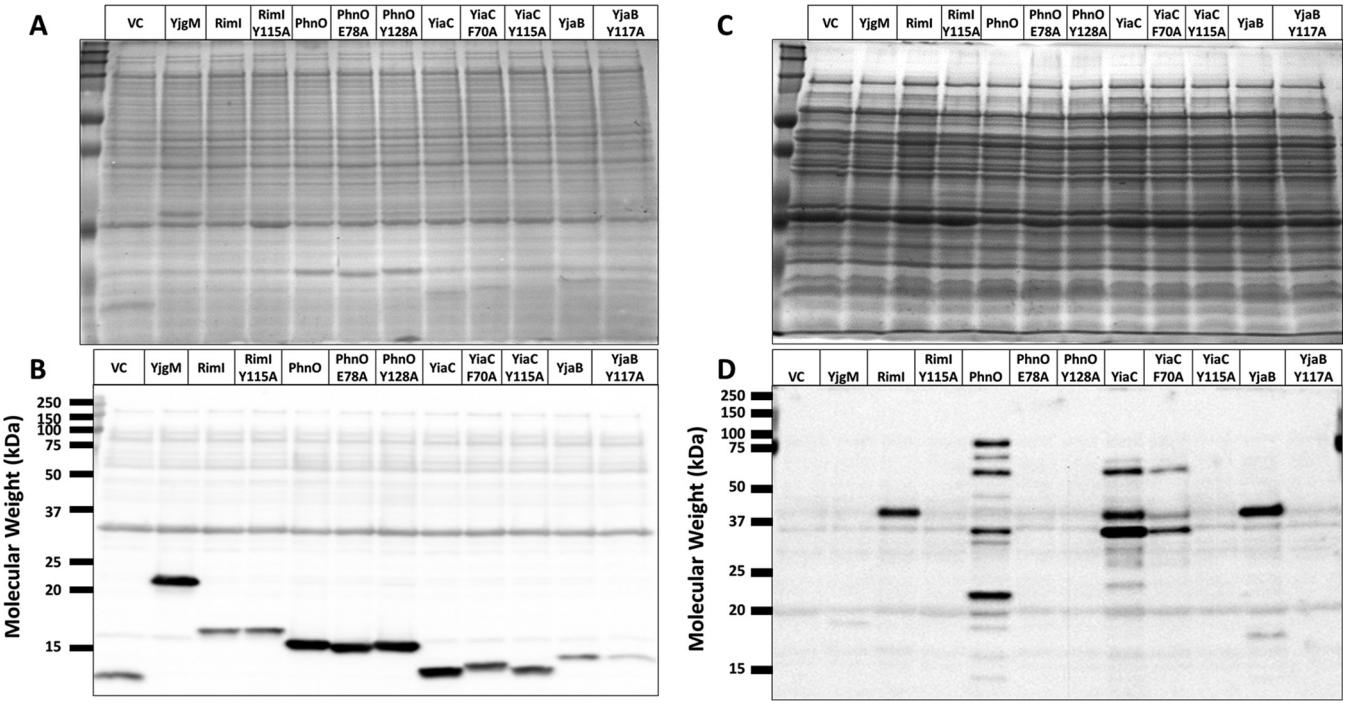
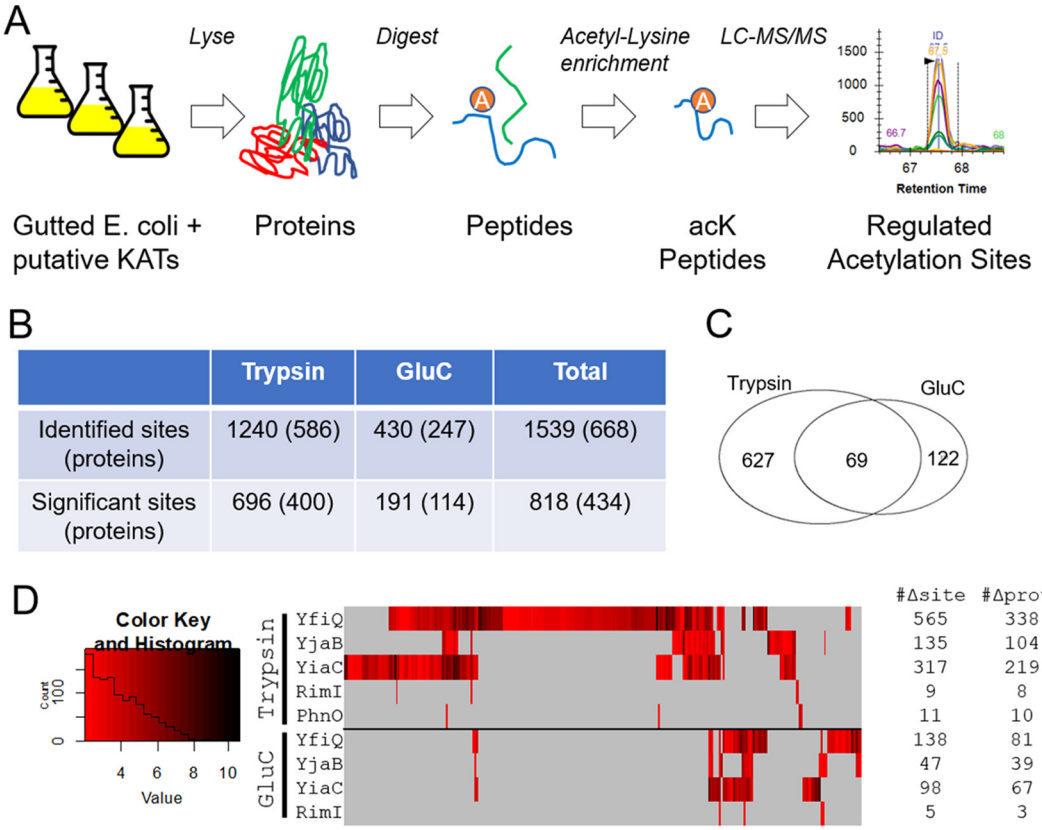


FIG 3 Mutation of conserved catalytic amino acids prevents RimI-, PhnO-, YjaB-, and YiaC-dependent acetylation. The gutted strain (BW25113 Δ*pta yfiQ acs cobB*) was transformed with the pCA24n vector control, pCA24n carrying the wild-type allele for each putative KAT, or mutant alleles for each putative KAT with alanine substitutions of the indicated residues. The resulting strains were grown in TB7 supplemented with 0.4% glucose, 100 μM IPTG, and 25 μg/ml chloramphenicol for 8 h. Crude lysates harvested after 4 h were analyzed for expression of the KAT proteins. Whole-cell lysates harvested after 8 h were analyzed for acetylation. Coomassie blue-stained SDS-PAGE gels (A and C) served as loading controls for anti-His (B) and anti-acetyllysine (D) Western blot analysis.

were analyzed for successful expression of the mutant proteins and for acetylation. All putative KAT variants were detected at comparable levels by anti-His Western blotting, except YjaB Y117A, whose levels were clearly reduced relative to its wild-type isoform (Fig. 3A and B). Overexpression of all tyrosine and glutamate mutants of YiaC, YjaB, RimI, and PhnO eliminated the acetylation signal produced by the wild-type isoforms (Fig. 3C and D). However, the YiaC F70A mutant did not completely lose activity, as it produced the same acetylated bands as the wild-type isoform, but with reduced intensity.

Because the amount of the YjaB Y117A protein was reduced relative to wild-type YjaB in the anti-His Western blot, we mutated this residue to phenylalanine (Y117F) to determine if soluble expression of this mutant improved. This Y-to-F mutation removes the hydroxyl group involved in reprotonation of CoA but retains the phenyl ring. The YjaB Y117F mutant showed similar soluble expression levels compared to wild type (WT) and a decreased acetylation signal similar to that of the other tyrosine mutants (Fig. S4). Overall, these data provided very strong evidence that RimI, YiaC, YjaB, and PhnO function as KATs.

Identification of putative KAT substrate proteins by mass spectrometry. Given the evidence that these four GNATs function as KAT enzymes, we sought to identify their substrate proteins and the amino acids that they acetylate. We used acetyllysine enrichment and mass spectrometry for unbiased identification and quantification of acetylation sites as described previously (Fig. 4A) (5, 23, 24). Proteome samples were isolated from the *E. coli* strains overexpressing RimI, YiaC, YjaB, and PhnO, as well as the known acetyltransferase YfiQ as a positive control and empty vector as a negative control. Using the standard workflow with trypsin digestion, we identified 1240 unique acetylation sites on 586 unique proteins (Fig. 4B and Table S1A). To increase the protein sequence coverage and therefore quantifiable acetylation sites, we performed the same experiments in parallel but substituted trypsin for a complementary protease, GluC



	Trypsin	GluC	Total
Identified sites (proteins)	1240 (586)	430 (247)	1539 (668)
Significant sites (proteins)	696 (400)	191 (114)	818 (434)

FIG 4 Identification of site-specific regulation of acetylation sites by KATs. (A) Cartoon showing workflow used to identify KAT target sites. (B) Overview of the significant sites and proteins regulated by at least one putative KAT. Significance defined as FDR of <0.01 within-set and $\log_2$(FC) of >2. (C) Venn diagram showing the complementary nature of trypsin and GluC digestion in terms of significant acetylation sites. (D) Heat map of all significant changes; acetylation sites are grouped by unsupervised hierarchical clustering.

(25–27), which expanded the total number of identifications by nearly 25% to 1,539 unique acetylation sites on 668 proteins (Fig. 4B and Table S1A).

To determine the set of acetylation sites regulated by these novel KATs and YfiQ, we applied stringent filters to the quantitative comparisons between the overexpression samples and controls (q-value < 0.01 and $\log_2$ [FC] ≥ 2, which is a $\geq$4-fold increase), resulting in a total of 818 acetylation sites on 434 proteins whose acetylation increased with overexpression of at least one KAT (Fig. 4B). These altered acetylation site levels were not driven by proteome remodeling, as only a few proteins were altered due to overexpression of any KAT (Table S1B). Again, the additional data from GluC digestion proved complementary, revealing 122 additional significantly increased acetylation sites (Fig. 4C). The acetylation sites, their fold increase, and the overlap between these putative KATs and YfiQ are shown as a heat map in Fig. 4D. As expected, the known acetyltransferase, YfiQ, acetylated the most lysines, with a total of 649 sites with significantly enhanced acetylation on 364 proteins (Table 1). YiaC and YjaB overexpression resulted in lower, yet substantial, numbers of significantly increased acetylation of

TABLE 1 Number of proteins and lysine residues with significantly increased acetylation upon overexpression of KATs

KAT	No. of unique proteins acetylated	No. of unique lysines acetylated
YfiQ	364	649
YjaB	128	171
YiaC	251	391
RimI	11	14
PhnO	10	11

sites/proteins (391/251 and 171/128, respectively). Overexpression of RimI and PhnO elicited the fewest changes, each acetylating fewer than 20 sites. It should be noted that we observed many more acetylated proteins by mass spectrometry than the number of bands we obtained via Western blot analysis (Fig. 1B). Mass spectrometry will detect site-specific acetylated peptides with greater sensitivity than Western blotting as previously shown (5, 23). Additionally, different acetylated proteins may migrate together on a gel and result in the appearance of only one band on a Western blot.

To further explore the specificity of these KATs, we compared the sites acetylated by KAT overexpression with sites that we previously found to be sensitive to deletion of *ackA*, which causes accumulation of the highly reactive acetyl donor AcP and therefore results in nonenzymatic protein acetylation (Fig. S5A and Table S1C) (5). Remarkably, of the 592 AckA-regulated sites, only 29 overlapped the 818 sites acetylated by PhnO, RimI, YiaC, YjaB, or YfiQ, further reinforcing their specificity and thus likely distinct functions. We also analyzed the primary amino acid sequences surrounding lysines that were acetylated by these novel KATs and found no specific neighboring residue preference (Fig. S5B and C). This suggests that substrate specificity cannot be determined by primary sequence alone and that three-dimensional analysis of protein structures should be taken into account.

KAT-dependent acetylation of proteins involved in translation and glycolysis. A large number of KAT substrate proteins were found to be involved in the GO Biological Process term translation (Benjamini-corrected *P* value 1.1E−22, determined by DAVID functional enrichment tool) (28, 29). Almost all ribosomal protein subunits were acetylated (51 of 55 proteins); some were acetylated by AcP only (8/55) and some were acetylated by one or more KATs but not AcP (11/55), but most were acetylated by at least one KAT and AcP (32/55) (Table S2A). Very few ribosomal lysines were acetylated by both a KAT and AcP (only 9 of 184 sites on the 55 subunits). In contrast, one-quarter (46/184) of the observed acetylated lysines were targeted by more than one KAT, with as many as 3 KATs acetylating the same lysine. Most of the amino acid-tRNA ligases (16/23 proteins) were acetylated; some were acetylated by AcP only (6/23), some by KATs only (5/23), and some by both (5/23). Again, lysines that were acetylated by both a KAT and AcP were rare (2/41 sites on 23 proteins). Only a few lysines were acetylated by multiple KATs (4/41). Three of the 7 elongation factors were acetylated; these acetylations were largely dependent on AcP (13/15). All of the initiation factors were acetylated, and these acetylations were almost entirely KAT dependent (7/8). These results are consistent with distinct roles for KAT-dependent and AcP-dependent acetylations.

Regarding the interplay between nonenzymatic and enzymatic acetylation, central metabolism was particularly interesting (Fig. 5). Twenty-seven proteins comprise the 3 glycolytic pathways in *E. coli* (Embden-Meyerhof-Parnas [EMP], Entner-Doudoroff [ED], and pentose phosphate [PP]). Of these 27 proteins, 20 were detected as acetylated: 2 strictly by KAT(s), 7 by KAT(s) and AcP, and 11 by AcP alone. A total of 97 lysines were acetylated: 9 by KAT(s) alone, 86 by AcP alone, and only 2 by both AcP and a KAT. Intriguingly, the majority of KAT-dependent acetylations (7/11) were found on proteins responsible for either the early or late steps of glycolysis, i.e., prior to the formation of glyceraldehyde 3-phosphate (GAP) or on enzymes responsible for aerobic AcCoA synthesis. In contrast, the majority of AcP-dependent acetylations (66/88) were found on proteins that all glycolytic pathways share. In support of the concept that KAT-dependent acetylation helps direct flux, 3 other proteins relevant to glycolysis are exclusively acetylated by KAT(s). YfiQ and YiaC acetylated the transcription factor GntR, which controls expression of the enzymes (Eda [2-keto-4-hydroxyglutarate aldolase] and Edd [phosphogluconate dehydratase]) that comprise the ED pathway (30). LipA synthesizes lipoate, whereas LipB transfers a lipoyl group onto a lysine in the E2 subunit (AceF) of the pyruvate dehydrogenase complex (PDHC). The 3 subunits of PDHC, whose activity requires lipoylation, are highly acetylated, but almost entirely by AcP. In contrast, LipA and LipB are entirely acetylated by KATs (7 lysines on LipA by YfiQ, YiaC, and YjaB and 1 lysine on LipB by YfiQ). These observations are consistent with the

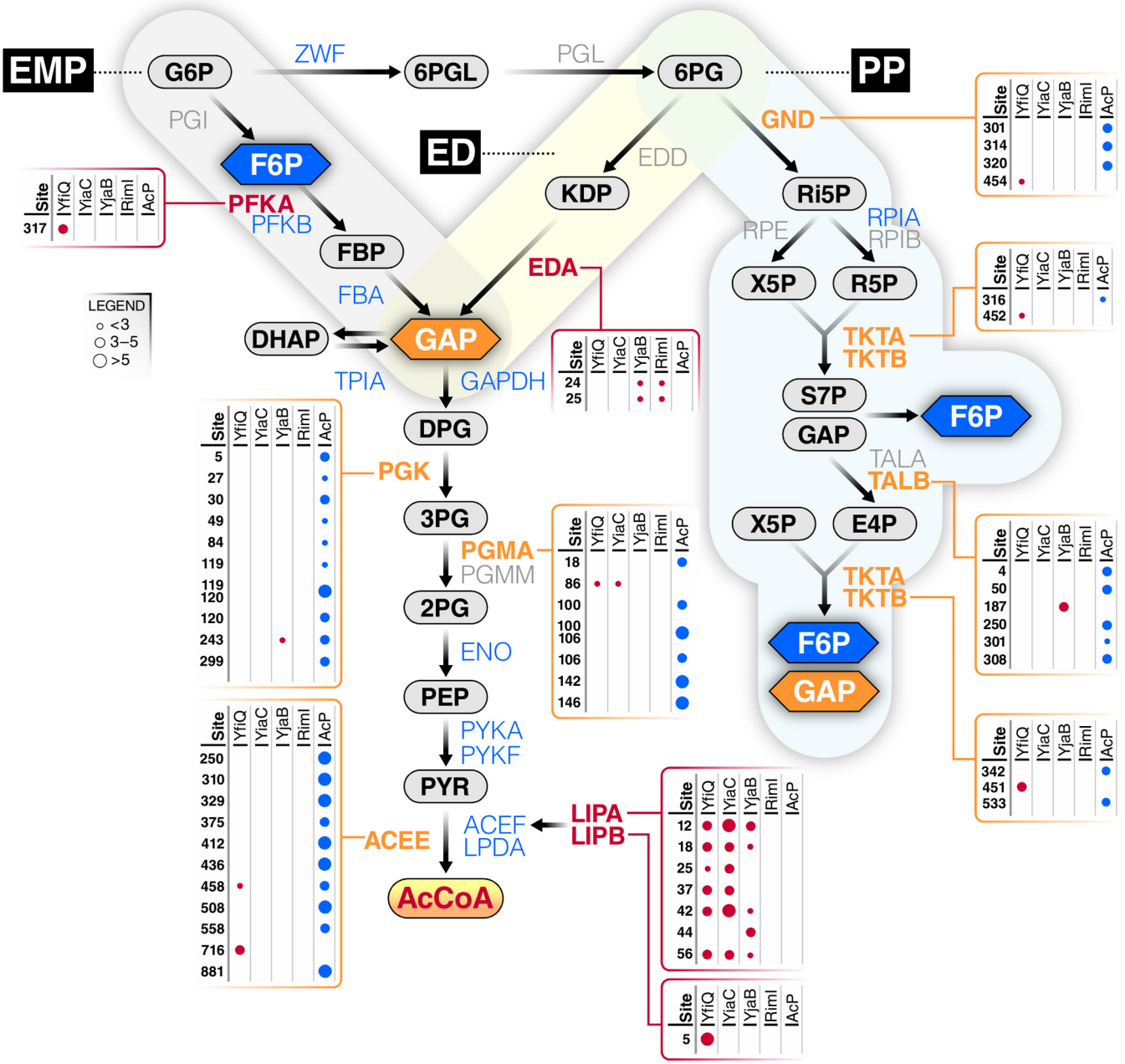


FIG 5 Most of central metabolism is differentially acetylated by acetyl-P and/or KATs. The three glycolytic pathways, Embden-Parnas-Meyerhof (EMP), Entner-Dourdoroff (ED), and pentose phosphate (PP), are shown with metabolites and enzymes indicated. Some enzymes are not acetylated (gray), while others are acetylated by acetyl-P alone (blue), KATs alone (red), or both (orange). Enzymes with boxes were modified by at least one KAT (as indicated); some were also acetylated by acetyl-P (AcP). The size of the dot indicates the fold upregulation for each lysine by either a KAT or AcP.

hypothesis that KAT-dependent acetylation helps direct flux through the 3 different glycolytic pathways and regulates the transition from glycolysis to AcCoA-dependent pathways, such as the TCA cycle, fatty acid biosynthesis, and different forms of fermentation.

Structural analysis of KAT and AcP-dependent acetylation sites. Previously, we analyzed the location of lysine residues on several glycolytic enzymes that are nonenzymatically acetylated by AcP (5). Here, we expanded our structural analysis to include selected enzymes in the EMP, ED, and PP pathways. We specifically investigated the KAT-dependent and/or AcP-dependent acetylated lysines on available *E. coli* protein structures (Fig. 6). Excluding proteins modified by AcP alone, we evaluated acetylated

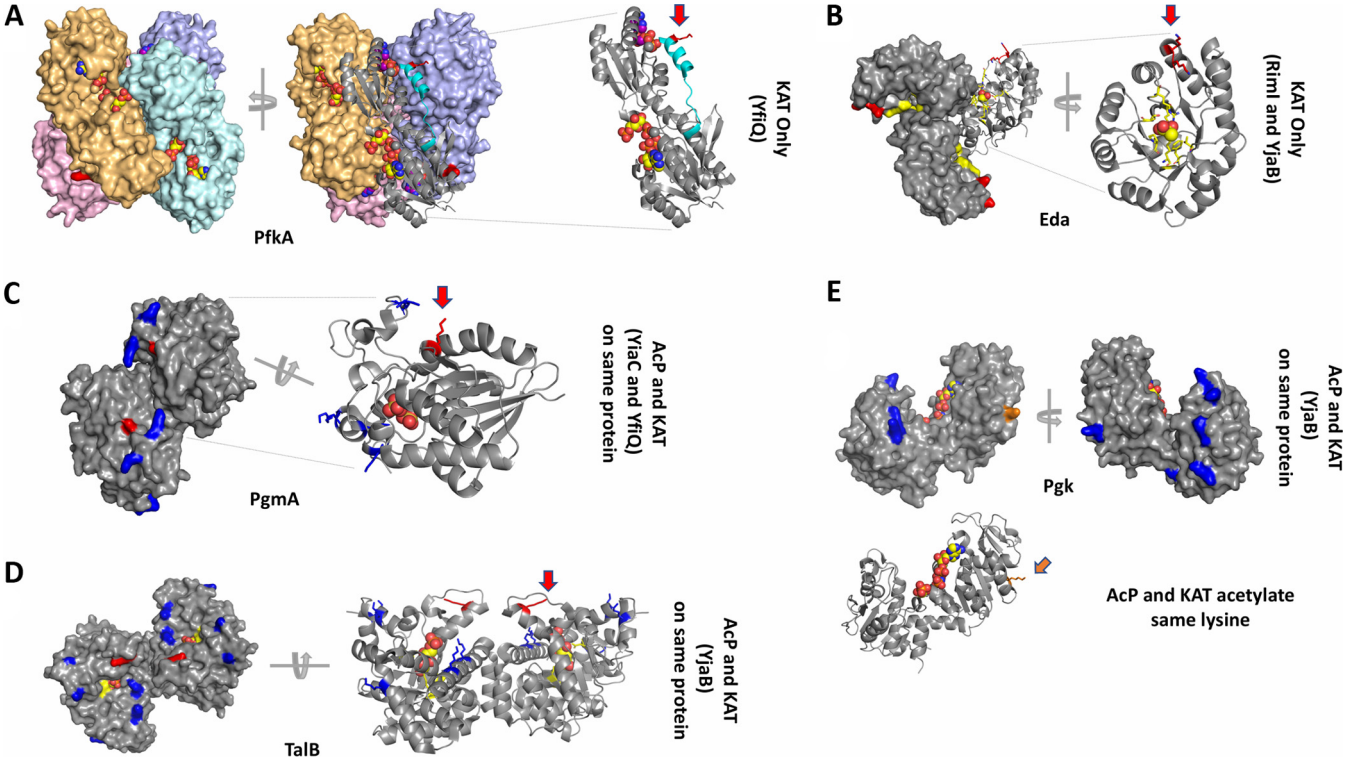


FIG 6 Structural analysis of selected proteins modified by KATs and/or AcP. (A) PfkA (PDB ID 1pfk) structure. Each monomer of the tetramer is tinted in pink, orange, cyan, and violet and shown as a surface representation. The ligand ADP is bound to the allosteric site and the ligand fructose 1,6-bisphosphate is bound to the active site; both are shown as spheres. One monomer is also shown as a ribbon representation. A red arrow indicates the location of K317. The C terminus that is disordered in the 2pfk structure is shown in cyan. (B) Eda (PDB ID 1eua) structure. A surface representation of the trimer is shown in gray. One monomer of the trimer is also shown as a ribbon representation, and a red arrow indicates the two adjacent sites of acetylation. Pyruvate is shown as spheres. The active site residues are colored in yellow, and K24 and K25 are colored in red. (C) PgmA (PDB ID 1e58) structure. The dimer is shown as a surface representation, and one monomer of the dimer is shown as a ribbon representation. Sulfate is shown as spheres in the active site. K86, which is acetylated by YfiQ and YiaC, is shown in red and indicated by a red arrow. K18, 100, 106, 142, and 146 are acetylated by AcP and shown in blue. (D) TalB (PDB ID 4s2c) structure. The dimer is shown in both surface and ribbon representations. K187 is acetylated by YjaB and shown in red with a red arrow. K4, 50, 250, 301, and 308 are acetylated by AcP and shown in blue. Fructose 6-phosphate is shown as spheres in the active site, and surrounding residues are shown as yellow sticks. (E) Pgk (PDB ID 1zmr) structure. The monomer is shown as a surface and ribbon representation where K243, which is acetylated by both YjaB and AcP, is shown in orange and an orange arrow points to its location. K5, 27, 30, 49, 84, 119, 120, and 299 are acetylated by AcP and shown in blue. Phosphoaminophosphonic acid-adenylate ester and 3-phosphoglycerate are shown as spheres in the active site and modeled from the 1vpe structure.

proteins from three main groups: acetylated by a KAT only, acetylated by a KAT and AcP on different lysines of the same protein, and acetylated by a KAT and AcP on the same lysine of the same protein. Examples of proteins that were acetylated by a KAT only included PfkA and Eda, those modified by either a KAT or AcP on different residues included PgmA and TalB, and those modified by a KAT and AcP on the same lysine residue included Pgk. One representative protein from these pathways that was modified by each individual KAT was selected to evaluate substrate lysine locations in three dimensions (3D) (Fig. 5 and 6A to D): PfkA (YfiQ; EMP), Eda (RimI; ED), PgmA (YiaC; EMP), and TalB (YjaB; PP). Note that some of these proteins are modified by multiple KATs, a scenario that we will discuss below.

(i) Comparison of KAT-only acetylated lysines on selected substrate proteins. Phosphofructokinase A (PfkA) is an allosterically regulated tetrameric protein. We found that the acetylated lysine (K317) of PfkA was located at the end of an α helix at the C terminus of the protein and lies in a pocket formed by a second monomer of the tetramer (Fig. 6A). Therefore, K317 is found at the interface between monomers of the tetramer and lies outside the active site and allosteric site of the protein. K317 forms a salt bridge with D273 of an adjacent monomer, and Paricharttanakul (31) previously found that D273 is likely important for stabilizing the tetramer and affects the allosteric activation and inhibition network. A disruption of this salt bridge via acetylation could possibly alter allosteric properties of the protein. Furthermore, the C terminus is

important for stability of the oligomer (32), and when the allosteric effector ADP is not bound, this region becomes disordered (33). The fact that this region is disordered in the absence of the allosteric effector in the crystal structure suggests that this portion of the protein is mobile and therefore may be accessible to YfiQ for acetylation.

KHG/KDPG aldolase (Eda) is a trimeric protein and was found to be acetylated on two adjacent lysines (K24 and K25) by both RimI and YjaB (Fig. 6B). These amino acids are found at the end of a surface-accessible α helix, which is outside the active site and is not at the interface of monomers of the trimer. The only interaction observed for either of these amino acids was a salt bridge between K25 and E193 on a neighboring α helix. For this reason, the function of lysine acetylation on this protein is unclear.

(ii) Comparison of KAT and AcP acetylation sites on different lysines of the same protein. Phosphoglycerate mutase (PgmA) contains six lysines that were acetylated by either AcP or a KAT (Fig. 6C). K86 was the only site of enzymatic acetylation (YiaC and YfiQ), whereas K18, 100, 106, 142, and 146 were all nonenzymatically acetylated by AcP. While this enzyme was not previously considered to be allosteric, it was recently proposed to function as an allosteric enzyme, whereby dimer stabilization acts as the allosteric signal that is transmitted rather than the more typical binding of a specific effector to an allosteric site. In this case, the ordering and stabilization of the region that contains the lysines acetylated by AcP act as the transmission signal (34, 35). While all AcP acetylations occurred on this highly flexible domain of the protein, the KAT acetylation site (K86) was found outside the active site on a small 3_{10}-helix near but not directly interacting with the opposite monomer at the dimer interface. This lysine coordinates a water molecule between itself and E166 on a neighboring α helix. However, if stabilization of the dimer is truly acting as the allosteric signal, then this lysine (K86) is only indirectly involved in the allosteric site. Investigation of the AcP-modified lysines showed that only K100 was found to be in the active site. The function of all other lysines that were acetylated by AcP are currently unclear.

Transaldolase B (TalB) is a dimeric protein that is modified through both enzymatic and nonenzymatic acetylation (Fig. 6D). Similarly to PgmA, TalB is also acetylated on one lysine by a KAT (YjaB; K187) and several lysines by AcP (K4, 50, 250, 301, and 308). All these acetylated lysines are found on α helices on the same face of the C-terminal side of the beta-barrel. The α helices that surround the beta-barrel are known to be mobile (36). The enzymatic acetylation site occurs near the end of an α helix outside the protein active site and is not found at the interface between monomers of the dimer. There appears to be some specificity of acetylation in this location of the protein because two additional lysines are directly downstream (K192 and 193) on a loop and are not acetylated by either a KAT or AcP. Lysines acetylated nonenzymatically by AcP are also surface accessible. Two of these lysines (K301 and 308) are found on a long α helix that creates the dimer interface, but neither participates in interfacial interactions. K301 is near the active site where sugar phosphates bind, but K308 is further down the helix. This C-terminal helix is preceded by an extremely long loop (residues 254 to 277) that connects it with two additional helices that contain AcP-modified lysines K4 and K250. K4 forms polar interactions between its ε amino group and the backbone carbonyl oxygens of both S255 and E256, while K250 forms a salt bridge with E254. The ε amino group of K50 also forms polar contacts with the backbone carbonyl oxygen of E46 but is not linked to the long loop or helix where other AcP-modified lysines are found. It is unclear what effect these acetylated lysines have on protein function or oligomerization.

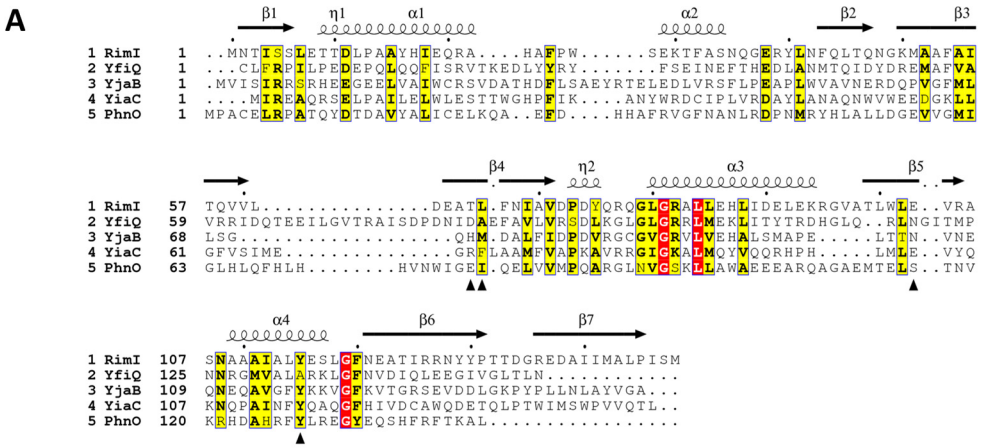
(iii) KAT and AcP acetylation of the same lysine residue. Only select lysines within each substrate protein are acetylated. Most often, the method by which these lysines become acetylated is either exclusively by a KAT or by AcP. In rare cases, the two mechanisms compete for the same lysine. One example where this occurs is in phosphoglycerate kinase (Pgk), whereby both AcP and a KAT (YjaB) acetylate K243 (Fig. 6E). Pgk is monomeric, and the *E. coli* protein has been crystallized in the open conformation. In other homologs, the structures of the partially closed and fully closed

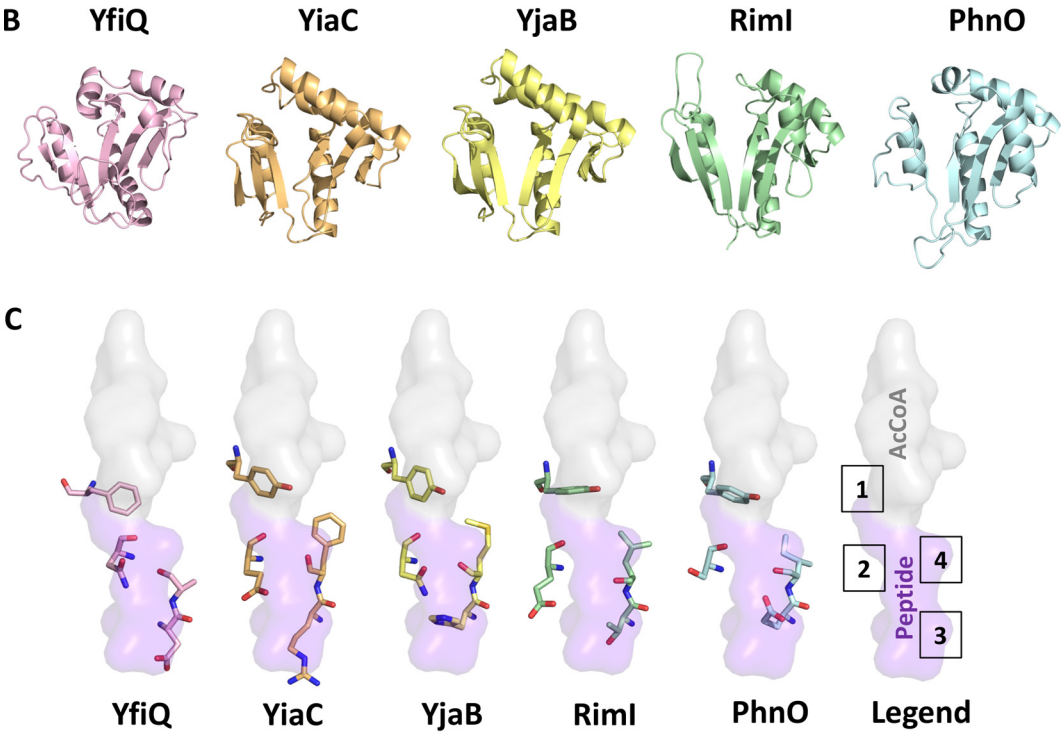
forms of the protein have also been determined (37). K243 is found at the end of an α helix and has no specific interactions with other amino acids of the protein. AcP also acetylated lysines 5, 27, 30, 49, 84, 119, 120, and 299. All KAT- and/or AcP-acetylated lysines are found on the surface of the protein and do not interact directly with the active site. Two lysines acetylated by AcP (K27 and 30) are on a loop that moves upon closure of the protein. Nearly all lysines acetylated by AcP are found within the N-terminal domain, whereas K243 and K299 are located in the C-terminal domain. It is unclear why both KAT and AcP acetylate K243 and why the other lysines are preferred sites for acetylation by AcP.

Structural and active site residue comparison of KATs. Initially, we used Phyre2 to predict amino acids that may be involved in catalysis of KATs, and these predictions informed our mutagenesis trials. Here, we chose to perform a more thorough structural analysis to determine whether these suggested amino acids were present in locations known to be important for activity in homologs. The sequence identity between KATs is low (<30%), but since GNATs share the same structural fold, we performed a structural comparison of these KATs in order to identify the location of active site residues in 3D. The *E. coli* crystal structure of RimI (5isv) and NMR structure of YjaB (2kcw) have been deposited into the Protein Data Bank (PDB); however, no structures have been determined for the other *E. coli* KATs (YfiQ, YiaC, and PhnO). Therefore, we built homology models of these three proteins.

Based on our models and available structures, we found that all KATs adopted the standard GNAT fold with a characteristic V-like splay. The structures and models also informed our manual refinement of our sequence alignment of KATs (Fig. 7A and B). There is significant sequence and structural variability in the α1-α2 and β6-β7 regions of each KAT. However, all of their active sites, with the exception of YfiQ, contained a conserved tyrosine known to act as a general acid in other GNAT homologs (21, 38). Upon further analysis, we found that the identity or location of the amino acid that tends to act as a general base may not be as conserved across KATs as the amino acid that acts as the general acid. For instance, E103 coordinates a water molecule to act as a general base in RimI from *Salmonella enterica* serovar Typhimurium LT2 (21) is in the same location in 3D as the corresponding amino acids (both E103) in *E. coli* RimI and YiaC (Fig. 7C). In contrast, N105 and S116 are in the same location in YjaB and PhnO, respectively. In theory, these amino acids can coordinate a water molecule, but to our knowledge the effect of substituting these amino acids for glutamate has not been evaluated. Our mutagenesis of E78 in PhnO significantly decreased its acetylation activity in the "gutted" strain (Fig. 3), indicating that this amino acid is critical for catalysis. Thus, the location of the amino acid that either coordinates a water molecule that functions as the general base in the reaction or the amino acid that directly participates in this function may be in a different location in 3D on different KATs. Mutation of F70 in YiaC decreased acetylation activity (Fig. 3), but not as substantially as the tyrosine mutation. In 3D, the equivalent amino acid in all KATs is hydrophobic (Fig. 7C), which may be important for substrate recognition. Regardless, this amino acid does not directly participate in the chemical reaction.

Newly identified KATs are conserved. Having identified these KATs in *E. coli*, we next asked whether orthologs of these KATs were present across bacterial phylogeny. To find these orthologs, a Hidden Markov Model (HMM) was built for each gene of interest and searched against 5,589 representative reference genomes from RefSeq (39, 40). Cutoffs for the HMM match score were derived manually by identifying a score such that genomes contained at most only one match. This cutoff was chosen to draw the line between orthologs and paralogs, i.e., when a genome has multiple copies of similar sequences but only one contains the biological function of the query sequence. As a highly conserved gene, *rimI* was identified in 4,459 genomes (Table S3A). The *yiaC* and *yjaB* genes were all broadly distributed across bacterial taxa and found in 421 and 692 genomes, respectively (Table S3B). However, *phnO* was found to have a very limited distribution, identified in only 22 genomes, and appears to belong exclusively to the





F138	Y115	Y117	Y115	Y128	1
N119	E103	N105	E103	S116	2
D82	R69	H72	T65	E78	3
A83	F70	M73	L66	I79	4

FIG 7 Sequence and structural comparison of KAT proteins and their key catalytic residues. (A) Sequence alignment of all five *E. coli* KATs. Only the GNAT portion of the sequence for YfiQ is shown. The structural elements above the sequences are based on the 5isv RimI structure. Red highlighting represents 100% identity, whereas yellow highlighting shows a global score of 70% identity based on ESPript 3.0 parameters. Black arrowheads beneath the sequences indicate the residues selected for structural comparison in panel C. (B) Comparison of overall structures and homology models of *E. coli* YfiQ (pink), YiaC (orange), YjaB (yellow), RimI (green; full C terminus not shown in the figure), and PhnO (blue) proteins in a ribbon representation. 3D structures of YjaB and RimI were determined previously (PDB IDs 2kcw and 5isv, respectively). We built homology models of the remaining KATs using the following structures as the templates: 4nxy for YfiQ, 2kcw for YiaC, and 1z4e for PhnO. Only the GNAT portion of the YfiQ protein sequence was used for the homology model. Further details regarding parameters for building and selecting representative homology models for these proteins are described in Materials and Methods. (C) Comparison of select active site residues potentially important for substrate recognition and catalysis in GNATs. The crystal structure of RimI (5isv) has the C terminus of one monomer bound in the active site of the second monomer. A surface

(Continued on next page)

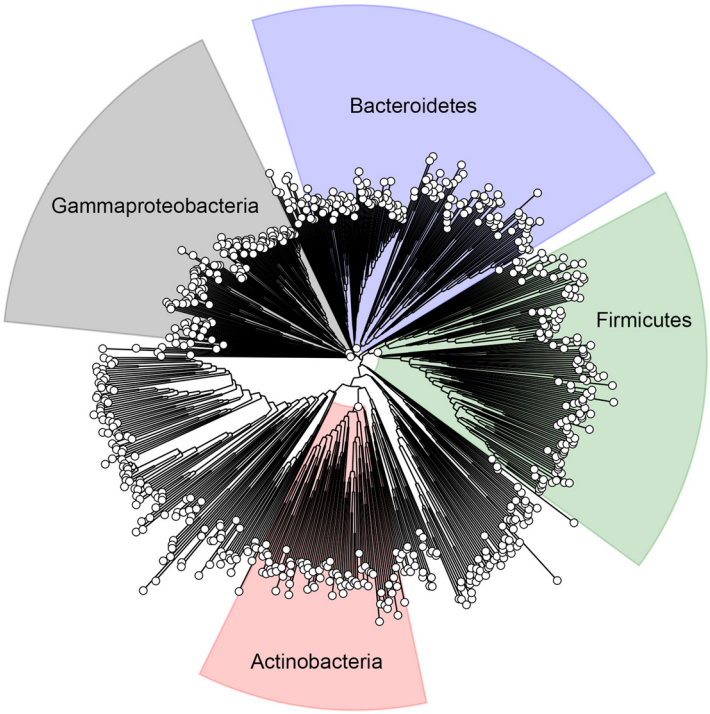


FIG 8 YjaB is highly conserved across bacteria. A representative phylogenetic tree showing bacterial species that contain a gene homologous to *yjaB* from *E. coli*. The trees for *rimI* and *yiaC* are comparable.

gammaproteobacteria. A representative phylogenetic tree shows the broad distribution of *yjaB* across the bacterial domain (Fig. 8).

Nineteen genomes contained all four new KATs. Unsurprisingly, many of these genomes correspond to *E. coli* strains or the closely related species *Salmonella enterica*. One hundred forty-eight genomes contained three KATs, and 782 genomes contained two KATs. Study of these KATs expressed heterologously in *E. coli* could help us to understand the potential role of acetylation in these other bacteria.

YiaC and YfiQ can inhibit migration in soft agar. We sought to determine the physiological relevance of acetylation by these KATs. Based on the *E. coli* gene expression database (https://genexpdb.okstate.edu/databases/genexpdb/), we found conditions under which these KATs may be expressed and, thus, when they may be relevant. Expression of each of the KATs appeared to be upregulated in stationary phase and/or under biofilm conditions. Thus, we tested overexpression constructs of the four novel KATs and YfiQ in a mucoidy assay, but we did not observe any difference relative to wild-type cells. We then tested these strains for motility. We found that overexpression of YiaC and YfiQ consistently reduced migration in a soft agar motility assay (Fig. 9A). The inhibition of migration was not due to a reduction in growth rate as the overexpression strains grew as well as their vector controls (data not shown). To determine whether this reduction required the acetyltransferase activity of YiaC, we tested overexpression of YiaC F70A, which had reduced acetyltransferase activity, and YiaC Y115A, which lost activity (Fig. 9B). Overexpression of YiaC YF70A inhibited migration similarly to overexpression of wild-type YiaC. In contrast, YiaC Y115A was unable to inhibit migration. To ensure this was not a strain-specific phenomenon, we recapitulated these data for YiaC in another strain background, MG1655 (Fig. S6). However, the overex-

FIG 7 Legend (Continued)
representation of this portion of the protein that encompasses the AcCoA donor (gray) and peptide acceptor (purple) site is shown. Each of the KAT homology models and structures was aligned using TopMatch and PyMOL. Four active site residues are shown. A table beneath the structures shows the specific residue numbers for each KAT. Residues that were mutated are shown in red.

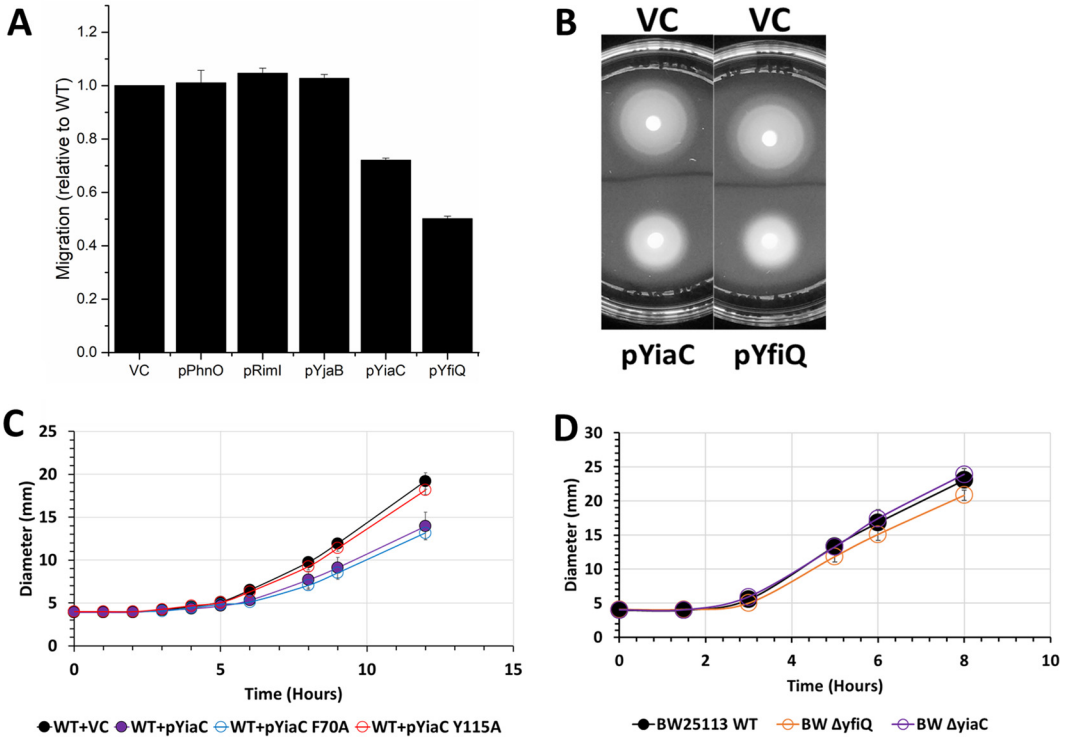


FIG 9 YiaC and YfiQ inhibit migration. Cultures were grown overnight in TB medium supplemented with chloramphenicol and 50 μM IPTG. Five microliters of each normalized culture was spotted on low-percentage TB plates supplemented with chloramphenicol and 50 μM IPTG. The diameter of the cell spot was measured hourly. (A) Final diameter relative to vector control (VC) after 12 h is shown for wild-type *E. coli* BW25113 strains carrying the indicated plasmids. (B) Representative motility plates of BW25113 carrying VC, pYiaC, or pYfiQ. (C) Hourly migration of wild-type *E. coli* BW25113 strains carrying pCA24n encoding YiaC, YiaC mutants, or vector control (VC). (D) Hourly migration of wild-type *E. coli* strain BW25113 or isogenic deletion mutants on low-percentage TB plates without supplement.

pression of YfiQ caused a growth inhibition in MG1655 (Fig. S6). If YiaC and YfiQ inhibit motility, then deletion of those genes may increase migration. However, the Δ*yiaC* mutant migrated equivalently to the wild-type parent, while the Δ*yfiQ* mutant had a slight reduction in migration in BW25113 (Fig. 9D).

DISCUSSION

Over the last decade, *N*ε-lysine acetylation has become recognized as an important posttranslational modification of bacterial proteins that regulates physiology. While acetylation of certain lysines may have a clear output, such as inhibition of enzyme activity due to active site *N*ε-lysine acetylation, the functional importance of many acetyllysine modifications is more difficult to discern. To uncover the role of acetylation in these unclear cases, it is helpful to use a model bacterium (e.g., *E. coli*) with a vast knowledge base of pathways, protein structure-function relationships, and physiology. Therefore, we constructed a "gutted" strain that lacked both known acetylation mechanisms (5, 6, 23) to examine if KATs other than YfiQ exist. This approach substantially decreased background acetylation, increased the signal-to-noise ratio, and allowed us to identify four enzymes that possess robust KAT activity: RimI, YiaC, YjaB, and PhnO. We acknowledge that overexpression may produce artifacts. However, knowledge of amino acids required for catalytic activity in homologous enzymes allowed us to construct inactive or minimally active mutant enzymes and determine that they function as KATs.

To define statistically significant KAT lysine target sites, we applied very stringent requirements of >4-fold increase in acetylated lysines in the KAT overexpression strains relative to the vector control and a *q* value of less than 0.01. Using these strict criteria, we identified 818 acetylated lysines on 434 proteins. Most of these modified lysines

were acetylated by a single KAT. While the overlap of lysines acetylated by the 5 KATs is relatively minor, the overlap between all the KAT-dependent acetylations and AcP-dependent acetylations is even smaller. The specificity of each KAT suggests that *E. coli* has evolved distinct regulatory modalities, perhaps reflecting the need to remodel the proteome in certain environments. This concept is supported by the patterns of acetylation of translation-associated proteins and of the glycolytic pathway proteins (Fig. 5). To further emphasize this concept, we determined that YfiQ and YiaC, but not the other three KATs, inhibit motility. This suggests that acetylations catalyzed by YfiQ and YiaC have distinct outcomes and thus are specific. Determining conditions under which these KATs are active could reveal the advantage of minimal redundancy among targets.

The possibility that GNAT family members other than YfiQ might possess KAT activity was examined in two studies by Venkat and coworkers, where the ability of GNATs to *in vitro* acetylate malate dehydrogenase and tyrosyl-tRNA synthetase was assessed (41, 42). Neither protein was enzymatically acetylated by GNATs, at least not above the acetylation level achieved by incubating with AcCoA alone. Importantly, our mass spectrometry data corroborate those data, as neither malate dehydrogenase nor tyrosyl-tRNA synthetase was modified by the KATs we identified *in vivo*. With similar *in vitro* acetylation assays performed by our group, the level of acetylation is already high on purified target proteins, so the effect of the KATs on acetylation of the targets was small or unobservable. We are currently working on optimizing this protocol to investigate the activity of these KATs *in vitro*.

Out of the hundreds of proteins we identified as acetylated by the newly identified KATs, several are central metabolic proteins and many are components of the translational machinery. From our structural analysis of a limited set of proteins, it appears that KAT-dependent acetylations occur primarily on the ends of α helices near allosteric or active sites, and sometimes at oligomeric interfaces. Most sites tend to be surface accessible and may be intricately involved in allosteric signaling networks and/or mediate protein-protein interactions. On the other hand, AcP-dependent acetylations are mainly located on α helices, loops, and active site amino acids. Additional studies are needed to understand why multiple KATs acetylate the same lysine on the same protein. Future studies will decipher the role of specific target proteins, such as LipA, that are differentially acetylated by multiple different KATs and may be differentially regulated depending on stress, environment, and nutrient availability. Similarly, it will be important to determine if KAT-acetylated lysines on helices have the propensity to unwind compared to other helices in the KAT substrate proteins or to helices containing AcP-acetylated lysines. To examine whether these trends hold for a larger set of KAT substrate proteins, we are currently performing a wider analysis across all identified substrates with structures.

At this point, we can only speculate on the effects these modifications have on many of the target proteins, but for some proteins not included in our structural analysis, critical lysines are acetylated. For example, selenide water dikinase (also known as selenophosphate synthetase) is acetylated by YiaC on K20, an amino acid known to be critical for catalyzing selenophosphate synthesis (43). Both YiaC and YfiQ acetylated FabI on K205, an amino acid known to be important for an essential step in fatty acid biosynthesis: the reduction of an enoyl-acyl carrier protein (44). YfiQ acetylated adenylate kinase (Adk) on K136, which is known to be important in stabilizing the open state of the enzyme by forming a salt bridge with D118. This salt bridge appears to be important for dynamic transitions between different states (45). Finally, cysteine synthase A (CysK) is acetylated by YiaC on K42, which is within the active site of the protein and becomes covalently modified by pyridoxal 5′-phosphate (PLP); the enzymatic activity of CysK depends on Schiff base formation of this lysine with PLP (46). In the case of large protein complexes, such as the ribosome, that are multiply acetylated on several subunits, it is tempting to speculate that acetylation of some seemingly inconsequential lysines may produce significant effects when combined, for example, stabilizing or destabilizing complex formation or altering ribosomal function.

We find proteins in pathways involved in metabolism and translation are particularly heavily acetylated. However, many KAT-dependent acetylations in these processes were distinct from those catalyzed by AcP as previously reported (5, 6) The KATs described here tend to acetylate enzymes that regulate the branch points of central metabolism, while AcP seems to modify many of the central metabolic enzymes (Fig. 5). This suggests the tempting hypothesis that KATs have evolved to specifically regulate key flux points in metabolism, while AcP-dependent acetylation may be a global response to the carbon flux.

A recent report revealed the only known *E. coli* deacetylase, CobB, has lipoamidase (delipoylase) activity (47). Lipoyl groups can be found on subunits of several major central metabolic complexes and contribute to the activity of these complexes. Rowland et al. (47) found that CobB could regulate the activities of the pyruvate dehydrogenase (PDH) and α-ketoglutarate dehydrogenase (KDH) complexes and could delipoylate the AceF and SucB components of PDH and KDH, respectively. As mentioned in the results above, we find that most of these metabolic complexes are multiply acetylated by AcP and/or KATs, including AceF and SucB. Additionally, we found that a protein responsible for generating the lipoyl groups for these lipoylated subunits, LipA, was highly acetylated by KATs, and Rowland et al. found that LipA coimmunoprecipitated with CobB. The tight cooccurrence of lipoylated and acetylated proteins at key nodes of central metabolism with the potential to be regulated by CobB suggests an interesting dynamic between acetylation and lipoylation that warrants further study.

E. coli RimI appears to acetylate lysines on multiple proteins. This is an interesting observation, as RimI from both *E. coli* and *Salmonella* Typhimurium is known to function as an N-terminal alanine acetyltransferase that has but one known target, the ribosomal protein S18 (21, 48, 49). While RimI from *E. coli* and *S.* Typhimurium is characterized by stringent N-terminal alanine specificity, RimI from *Mycobacterium tuberculosis* exhibits relaxed N-α amino acid substrate specificity *in vitro* (50). Intriguingly, we observed that *E. coli* RimI can also acetylate an *N*ε-lysine on a different ribosomal protein, L31. The *N*ε-lysine of L31 is found on a long unstructured region of the protein and may bind to RimI in a similar conformation as the C terminus of the *E. coli* RimI protein in its crystal structure (PDB ID 5isv). The α-amino group of alanine on S18 is also found at the end of a long unstructured region. This insinuates that RimI could exhibit both Nα-amino acid and *N*ε-lysine acetylation activity, but the substrate specificity of this enzyme is still unclear.

Similarly, PhnO is an aminoalkylphosphonate acetyltransferase in both *E. coli* and *S. enterica* (51, 52). PhnO is part of a gene cluster involved in the utilization of phosphonate under inorganic phosphate starvation conditions. While PhnO is not absolutely required for phosphonic acid utilization, it does acetylate (*S*)-1-aminoethylphosphonate and aminomethylphosphonate (52, 53). Of the 10 proteins that PhnO acetylates, one is inorganic triphosphatase, which indicates additional levels of phosphate regulation via this KAT.

Prior to our study, both RimI and PhnO were identified as having functions unrelated to *N*ε-lysine acetylation. It is interesting that these two KATs have significantly fewer internal lysine protein substrates (11 and 10, respectively) than does YfiQ, YiaC, or YjaB. The reasons for this dual character of the two KATs are unknown. While it could simply be that these enzymes have broad substrate specificity, it is tempting to speculate that the *N*ε-lysine acetylation by RimI and PhnO may be part of a more complex cellular regulatory mechanism for bacteria that harbor these KATs.

We found that overexpression of both YiaC and YfiQ inhibits motility. For YiaC, none of the targets that we determined by mass spectrometry provide a simple explanation for this phenotype. For YfiQ, the effect of deletion or overexpression on motility has not been directly assessed, although a previous report supports the idea that YfiQ could inhibit motility, as a Δ*yfiQ* mutant exhibited slightly enhanced transcription of flagellar genes (54). However, there is also evidence that YfiQ can enhance motility. This report and others find that YfiQ can acetylate K180 of RcsB, a response regulator that represses transcription of the master regulator of flagellar biosynthesis, *flhDC*. Acetylation of K180

would thus be expected to prevent repression by inhibiting RcsB from binding to DNA, enhancing migration (55). This is contrary to what we observe, which suggests that YfiQ inhibits migration through a target other than RcsB, at least under the tested conditions. Finally, YfiQ acetylates K67 and K76 of FlgM, an anti-sigma factor for the sigma factor FliA (σ^{28}). FliA is required for initiation of many genes involved in flagellar biosynthesis. If YfiQ was acting through FlgM to inhibit motility, it would suggest that acetylated FlgM would bind FliA more tightly.

Contrary to our expectations, both Δ*yfiQ* and Δ*yiaC* mutants migrated at a rate similar to their wild-type parent. There are three possible explanations: (i) YfiQ and YiaC do not regulate motility; (ii) YfiQ and YiaC may not be expressed under the tested conditions, and thus, deletion of these genes would have no effect; and (iii) YfiQ and YiaC may compensate for each other; alternatively, some other KAT or AcP may contribute to compensation.

While we have a phenotype for overexpression of YiaC, the other three new KATs do not yet have clear phenotypes. Analysis of the *E. coli* gene expression database suggested conditions under which these KATs may be expressed. Based on these data, RimI (56–58), PhnO (59, 60), YjaB (61, 62), and YiaC are all upregulated during stationary phase dependent on the stationary-phase sigma factor, RpoS, and under biofilm-forming conditions. RimI (58), PhnO, and YjaB (62, 63) are upregulated under heat shock conditions, while PhnO and YiaC are downregulated during cold shock. Furthermore, YiaC is upregulated during oxidative stress (63). We also searched the genomic context and found YiaC and PhnO are encoded in polycistronic operons, while YjgM and YjaB are monocistronic. The *yiaC* gene is directly downstream and overlaps four nucleotides of the *tag* gene that encodes 3-methyl-adenine DNA glycosylase I. The product of the *tag* gene is important for removing potentially mutagenic alkylation damage from DNA, but it is not induced through the adaptive response. This may suggest a separate promoter for *yiaC* within the *tag* gene that allows it to respond to oxidative stress. As mentioned previously, the *phnO* gene is carried with the other genes necessary for phosphonate utilization. We are currently pursuing phenotypic analyses of these KATs based on this information and the targets that we have identified.

Excitingly, these KATs are well conserved across bacteria. Thus, discoveries about how these KATs affect *E. coli* physiology are likely applicable to other bacteria. For example, many organisms require motility for their pathogenicity, and our data suggest that YiaC and YfiQ may regulate motility in *E. coli*. However, both *Yersinia pestis* and *Klebsiella pneumoniae* encode YiaC, and the fact that both species are nonmotile suggests other roles for these enzymes in these particular bacteria. KAT homologs are also encoded in pathogens such as *Listeria monocytogenes* and *Pseudomonas aeruginosa*, and it would be interesting to determine whether these KATs regulate pathogenesis. A simple method to determine whether these homologs possess KAT activity would be to use our "gutted" approach. By expressing a heterologous putative KAT in our "gutted" strain, one could perform Western blotting or mass spectrometry to determine any changes to the acetylome. Indeed, we have evidence that this can work for at least one protein from *Neisseria gonorrhoeae* (unpublished data). However, it is important to note that *E. coli* may not encode the targets from the native species, so determination of native targets must be validated independently.

In conclusion, we identified four GNAT family members that have KAT activity in addition to the known KAT, YfiQ. These five KATs catalyze acetylation of hundreds of proteins on over 1,500 lysines, and the acetyltransferase activity depends on conserved catalytic tyrosines and/or key glutamates found in many GNAT family members. Furthermore, the conservation of YiaC in certain pathogenic organisms like *Yersinia pestis* warrants consideration as a topic of study. Clearly, our results provide a starting point for further analysis that is sure to yield fruitful mechanistic and regulatory insight into the complex orchestration of acetylation of proteins in bacterial metabolism, transcription, and other processes.

MATERIALS AND METHODS

Chemicals. HPLC-grade acetonitrile and water were obtained from Burdick & Jackson (Muskegon, MI). Reagents for protein chemistry, including iodoacetamide, dithiothreitol (DTT), ammonium bicarbonate, formic acid (FA), and urea, were purchased from Sigma-Aldrich (St. Louis, MO). Sequencing-grade trypsin was purchased from Promega (Madison, WI). HLB Oasis SPE cartridges were purchased from Waters (Milford, MA).

Strains, media, and growth conditions. All strains used in this study are listed in Table 2 (64–66). *E. coli* strains were aerated at 225 rpm in TB7 (10 g/liter tryptone buffered at pH 7.0 with 100 mM potassium phosphate) supplemented with 0.4% glucose with a flask-to-medium ratio of 10:1 at 37°C. Derivatives were constructed by moving the appropriate deletions from the Keio collection (67) by generalized transduction with P1kc, as described previously (68). Kanamycin cassettes were removed, as described previously (67). Plasmids carrying known and putative GNAT family members were isolated from the ASKA collection (69) and transformed into the indicated strains. Mutagenesis of the plasmids was performed via QuikChange II site-directed mutagenesis kit (Agilent Technologies) using primers listed in Table 3. To maintain plasmids, chloramphenicol was added to a final concentration of 25 μg/ml. To induce GNAT expression from the pCA24n plasmid, IPTG (isopropyl-β-D-1-thiogalactopyranoside) was added to a final concentration of 50 μM or 100 μM.

Western blot analysis of protein acetylation and detection of His-tagged proteins. *E. coli* cells were aerated at 37°C in TB7 supplemented with 0.4% glucose for 10 h unless otherwise noted. When necessary, chloramphenicol was added to a final concentration of 25 μg/ml, while IPTG was added to a final concentration of 50 μM. Bacteria were harvested by centrifugation and lysed using BugBuster protein extraction reagent (Novagen, Merck Millipore, Billerica, MA). The amount of cell lysate loaded on the gel was normalized to the total protein concentration, as determined by the bicinchoninic acid (BCA) assay (Thermo Scientific Pierce, Waltham, MA). Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and normalization was verified by Coomassie blue staining. Protein acetylation was determined using a rabbit polyclonal anti-acetyllysine antibody (Cell Signaling, Danvers, MA) at a dilution of 1:1,000, as described previously (5, 23).

His-tagged proteins were detected with crude lysates. One milliliter of 1-OD_{600} culture was harvested, pelleted, and resuspended in 200 μl sample loading buffer. The samples were boiled for 10 min, and loaded directly onto SDS-PAGE gels. His-tagged proteins were detected with mouse anti-His tag (27E8) antibody at a dilution of 1:1,000 (Cell Signaling, Danvers, MA) and an anti-mouse IgG HRP-linked antibody at a dilution of 1:2,000 (Cell Signaling, Danvers, MA).

Cell lysis, proteolytic digestion of protein lysates, and affinity enrichment of acetylated peptides. For mass spectrometric analysis, bacteria were cultivated as described for Western blot analysis. We then processed isolated frozen bacterial pellets from the gutted strains carrying vector control (AJW5426) or one of the 5 KAT candidates (i) RimI (AJW5499), (ii) YiaC (AJW5501), (iii) YjaB (AJW5504), and (iv) PhnO (AJW5513), as well as the known KAT YfiQ (AJW5497). Each of the strains was processed as 3 biological replicates. Cell pellets of the indicated strains were suspended in 6 ml of PBS and centrifuged at 4°C, 15,000 $\times$ *g*, for 20 min. The firm cell pellet was suspended and denatured in a final solution of 6 M urea, 100 mM Tris, 75 mM NaCl, and the deacetylase inhibitors trichostatin A (1 mM) and nicotinamide (3 mM). Samples were sonicated on ice (5$\times$ each for 15 s), cellular debris was removed by centrifugation, and the supernatants were processed for proteolytic digestion. Lysates containing 1.5 mg of protein were reduced with 20 mM DTT (37°C for 1 h) and subsequently alkylated with 40 mM iodoacetamide (30 min at RT in the dark). Samples were diluted 10-fold with 100 mM Tris (pH 8.0) and incubated overnight at 37°C with sequencing-grade trypsin (Promega) added at a 1:50 (wt/wt) enzyme/substrate ratio. In parallel, separate 1.5-mg protein aliquots were digested using endoproteinase Glu-C (Roche, Indianapolis, IN) by adding Glu-C at a 1:50 (wt/wt) protease-to-substrate protein ratio and incubating overnight at 37°C. Subsequently, samples were acidified with formic acid and desalted using HLB Oasis SPE cartridges (Waters) (83). Proteolytic peptides were eluted, concentrated to near dryness by vacuum centrifugation, and suspended in NET buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA). A small aliquot of each protein digestion ($\sim$10 μg) was saved for protein-level identification and quantification. The remaining proteolytic peptide samples were used for affinity purification of acetylated peptides (K^{ac}).

Acetylated peptides were enriched using 1/4 tube of the anti-acetyllysine antibody-bead-conjugated PTMScan Acetyl-Lysine Motif [Ac-K] kit (Cell Signaling Technologies) for each of the 1-mg protein lysate samples according to the manufacturer's instructions. Prior to mass spectrometric analysis, the acetylated peptide enrichment samples were concentrated and desalted using C_{18} Zip-tips (Millipore, Billerica, MA).

Mass spectrometric analysis. Samples were analyzed by reverse-phase HPLC-ESI-MS/MS using the Eksigent Ultra Plus nano-LC 2D HPLC system (Dublin, CA) combined with a cHiPLC system, which was directly connected to a quadrupole time-of-flight Sciex TripleTOF 6600 mass spectrometer (Sciex, Redwood City, CA) (23). After injection, peptide mixtures were transferred onto a C_{18} precolumn chip (200-μm by 6 -mm ChromXP C_{18}-CL chip, 3 μm, 300 Å, Sciex) and washed at 2 μl/min for 10 min with the loading solvent (H_2O/0.1% formic acid) for desalting. Subsequently, peptides were transferred to the 75-μm by 15-cm ChromXP C_{18}-CL chip, 3 μm, 300 Å (Sciex) and eluted at a flow rate of 300 nl/min with a 3-h gradient using aqueous and acetonitrile solvent buffers (23).

(i) Data-dependent acquisitions. To build a spectral library for protein-level quantification, the mass spectrometer was operated in data-dependent acquisition (DDA) mode where the 30 most abundant precursor ions from the survey MS1 scan (250 ms) were isolated at 1-*m/z* resolution for collision-induced dissociation tandem mass spectrometry (CID-MS/MS, 100 ms per MS/MS, high-sensitivity product ion

TABLE 2 Bacterial strains and plasmids

Strain	Description	Reference or source
BW25113	F$^-$ λ$^-$ Δ(*araD-araB*)*567* Δ(*rhaD-rhaB*)*568* Δ*lacZ4787 rrnB3 rph-1 hsdR514*	64
MG1655	λ*-rph-1*	A. Ninfa (University of Michigan)
AJW678	*thi-1 thr-1*(Am) *leuB6 metF159*(Am) *rpsL136 lacX74*	65
AJW2013	AJW678 Δ(*ackA pta hisJ hisP dhu*) *zej223*-Tn*10*	66
AJW5070	AJW678 Δ(*ackA pta hisJ hisP dhu*) *zej223*-Tn*10* Δ*yfiQ*::*frt kn*	P1: JW2568 × AJW2013
AJW2922	BW25113 Δ*pta*::*frt kn*	5
AJW5318	BW25113 Δ*pta*::*frt*	Kn cassette flipped from AJW2922
AJW5339	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt kn*	P1: JW2568 × AJW5318
AJW5359	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt*	Kn cassette flipped from AJW5339
AJW5373	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acs*::*frt kn*	P1: JW4030 × AJW5359
AJW5374	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*cobB*::*frt kn*	P1: JW1106 x AJW5359
AJW5420	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acs*::*frt*	Kn cassette flipped from AJW5373
AJW5426	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acs*::*frt* Δ*cobB*::*frt kn*	P1: JW1106 × AJW5420
AJW5493	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yiiD	Transformed pCA24n-yiiD into 5426
AJW5494	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yhbS	Transformed pCA24n-yhbS into 5426
AJW5495	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-speG	Transformed pCA24n-speG into 5426
AJW5496	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yjgM	Transformed pCA24n-yjgM into 5426
AJW5497	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yfiQ	Transformed pCA24n-yfiQ into 5426
AJW5498	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-rimJ	Transformed pCA24n-rimJ into 5426
AJW5499	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-rimI	Transformed pCA24n-rimI into 5426
AJW5500	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-argA	Transformed pCA24n-argA into 5426
AJW5501	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yiaC	Transformed pCA24n-yiaC into 5426
AJW5502	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yhhY	Transformed pCA24n-yhhY into 5426
AJW5503	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yncA	Transformed pCA24n-yncA into 5426
AJW5504	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yjaB	Transformed pCA24n-yjaB into 5426
AJW5505	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yedL	Transformed pCA24n-yedL into 5426
AJW5506	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-aat	Transformed pCA24n-aat into 5426
AJW5507	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-elaA	Transformed pCA24n-elaA into 5426
AJW5508	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yjdJ	Transformed pCA24n-yjdJ into 5426
AJW5509	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-atoB	Transformed pCA24n-atoB into 5426
AJW5510	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yafP	Transformed pCA24n-yafP into 5426
AJW5511	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-ypeA	Transformed pCA24n-ypeA into 5426
AJW5512	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yjhQ	Transformed pCA24n-yjhQ into 5426
AJW5513	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-phnO	Transformed pCA24n-phnO into 5426
AJW6163	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-ypfI	Transformed pCA24n-ypfI into 5426
AJW6164	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-astA	Transformed pCA24n-astA into 5426
AJW6165	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-rffC	Transformed pCA24n-rffC into 5426
AJW6166	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yhhK	Transformed pCA24n-yhhK into 5426
AJW6167	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yghO	Transformed pCA24n-yghO into 5426
AJW6168	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-citC	Transformed pCA24n-citC into 5426
AJW5537	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*cobB*::*frt kn* + pCA24n	Transformed pCA24n into 5374
AJW5538	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*cobB*::*frt kn* + pCA24n-yfiQ	Transformed pCA24n-yfiQ into 5374
AJW5539	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n	Transformed pCA24n into 5426
AJW6126	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yjaB Y117F	Transformed pCA24n-yjaB Y117F into 5426
AJW6130	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-phnO E78A	Transformed pCA24n-phnO E78A into 5426
AJW6131	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-phnO Y128A	Transformed pCA24n-phnO Y128A into 5426
AJW6132	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-rimI Y115A	Transformed pCA24n-rimI Y115A into 5426
AJW6136	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yiaC F70A	Transformed pCA24n-yiaC F70A into 5426
AJW6137	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yjaB Y117A	Transformed pCA24n-yjaB Y117A into 5426
AJW6138	BW25113 Δ*pta*::*frt* Δ*yfiQ*::*frt* Δ*acsA*::*frt* Δ*cobB*::*frt kn* + pCA24n-yiaC Y115A	Transformed pCA24n-yiaC Y115A into 5426
AJW5874	BW25113 + pCA24n-rimI	Transformed pCA24n-rimI into BW25113
AJW5875	BW25113 + pCA24n-phnO	Transformed pCA24n-phnO into BW25113
AJW5877	BW25113 + pCA24n-yjaB	Transformed pCA24n-yjaB into BW25113
AJW5878	BW25113 + pCA24n-yiaC	Transformed pCA24n-yiaC into BW25113
AJW5994	BW25113 + pCA24n	Transformed pCA24n into BW25113
AJW6067	BW25113 + pCA24n-yfiQ	Transformed pCA24n-yfiQ into BW25113
AJW6145	BW25113 + pCA24n-yiaC F70A	Transformed pCA24n-yiaC F70A into BW25113
AJW6146	BW25113 + pCA24n-yiaC Y115A	Transformed pCA24n-yiaC Y115A into BW25113
AJW6147	MG1655 + pCA24n	Transformed pCA24n into MG1655
AJW6151	MG1655 + pCA24n-phnO	Transformed pCA24n-phnO into MG1655
AJW6152	MG1655 + pCA24n-rimI	Transformed pCA24n-rimI into MG1655
AJW6154	MG1655 + pCA24n-yjaB	Transformed pCA24n-yjaB into MG1655
AJW6148	MG1655 + pCA24n-yiaC	Transformed pCA24n-yiaC into MG1655
AJW6149	MG1655 + pCA24n-yiaC F70A	Transformed pCA24n-yiaC F70A into MG1655
AJW6150	MG1655 + pCA24n-yiaC Y115A	Transformed pCA24n-yiaC Y115A into MG1655
AJW6155	MG1655 + pCA24n-yfiQ	Transformed pCA24n-yfiQ into MG1655

(Continued on next page)

TABLE 2 (Continued)

Strain	Description	Reference or source
AJW5868	BW25113 Δ*yiaC*::*frt kn*	P1: JW3519 × BW25113
AJW5882	BW25113 Δ*yfiQ*::*frt kn*	P1: JW2568 × BW25113
JW3519	KEIO Δ*yiaC*::*frt kn*	67
JW2568	KEIO Δ*yfiQ*::*frt kn*	67
JW4030	KEIO Δ*acs*::*frt kn*	67
JW1106	KEIO Δ*cobB*::*frt kn*	67
JW1908	KEIO Δ*fliC*::*frt kn*	67
JW3859	ASKA pCA24n-yiiD	69
JW3125	ASKA pCA24n-yhbS	69
JW1576	ASKA pCA24n-speG	69
JW5758	ASKA pCA24n-yjgM	69
JW2568	ASKA pCA24n-yfiQ	69
JW1053	ASKA pCA24n-rimJ	69
JW4335	ASKA pCA24n-rimI	69
JW2786	ASKA pCA24n-argA	69
JW3519	ASKA pCA24n-yiaC	69
JW3405	ASKA pCA24n-yhhY	69
JW5233	ASKA pCA24n-yncA	69
JW3972	ASKA pCA24n-yjaB	69
JW1917	ASKA pCA24n-yedL	69
JW0868	ASKA pCA24n-aat	69
JW2262	ASKA pCA24n-elaA	69
JW4088	ASKA pCA24n-yjdJ	69
JW2218	ASKA pCA24n-atoB	69
JW0224	ASKA pCA24n-yafP	69
JW2427	ASKA pCA24n-ypeA	69
JW4269	ASKA pCA24n-yjhQ	69
JW4054	ASKA pCA24n-phnO	69
JW1736	ASKA pCA24n-astA	69
JW0610	ASKA pCA24n-citC	69
JW3424	ASKA pCA24n-yhhK	69
JW2459	ASKA pCA24n-ypfI	69
JW5597	ASKA pCA24n-rffC	69
JW5848	ASKA pCA24n-yghO	69

scan mode) using the Analyst 1.7 (build 96) software with a total cycle time of 3.3 s as previously described (5, 23, 70).

(ii) Data-independent acquisitions. For quantification, all peptide samples were analyzed by data-independent acquisition (DIA, e.g., SWATH) (71), using 64 variable-width isolation windows (5, 72, 73). The SWATH cycle time of 3.2 s included a 250-ms precursor ion scan followed by 45-ms accumulation time for each of the 64 variable SWATH segments.

Mass spectrometric data processing and bioinformatics. Data-independent acquisitions (DIA) from acetyl-peptide enrichments were analyzed using the PTM Identification and Quantification from Exclusively DIA (PIQED) workflow and software tool (24). PIQED uses multiple open source tools to accomplish automated PTM analysis, including Trans-Proteomic Pipeline (74), MS-GF+ (75), DIA-Umpire (76), mapDIA (77), and Skyline (78). Relative quantification of acetylation levels from putative KATs versus VC in biological triplicates was used to determine fold-changes (see Table S1A

TABLE 3 Primers used in this study

Primer	Sequence
PhnO E78A F	5′-GTCAACTGGATCGGCGCAATTCAGGAGTTGGTG-3′
PhnO E78A R	5′-CACCAACTCCTGAATTGCGCCGATCCAGTTGAC-3′
PhnO Y128A F	5′-GACGCGCACCGTTTCGCTCTGCGCGAAGGCTA-3′
PhnO Y128A R	5′-TAGCCTTCGCGCAGAGCGAAACGGTGCGCGTC-3′
RimI Y115A F	5′-GCTGCCGCCATTGCCCTGGCCGAAAGTTTAGGCTTTAA-3′
RimI Y115A R	5′-TTAAAGCCTAAACTTTCGGCCAGGGCAATGGCGGCAGC-3′
YiaC F70A F	5′-CAGCATTATGGAAGGCCGAGCTCTGGCAGCGATGTTTG-3′
YiaC F70A R	5′-CAAACATCGCTGCCAGAGCTCGGCCTTCCATAATGCTG-3′
YiaC Y115A F	5′-GTTTATCAAAAAAATCAACCGGCGATAAATTTTGCCCAGGCACAGGGTTTTC-3′
YiaC Y115A R	5′-GAAAACCCTGTGCCTGGGCAAAATTTATCGCCGGTTGATTTTTTTGATAAAC-3′
YjaB Y117A F	5′-TGAGCAGGCGGTTGGGTTCGCTAAGAAGGTGGGTTTTAAG-3′
YjaB Y117A R	5′-CTTAAAACCCACCTTCTTAGCGAACCCAACCGCCTGCTCA-3′
YjaB Y117F F	5′-GCAGGCGGTTGGGTTCTTTAAGAAGGTGGGTTTTA-3′
YjaB Y117F R	5′-TAAAACCCACCTTCTTAAAGAACCCAACCGCCTGC-3′

in the supplemental material). Sites were called regulated when FDR was <0.01 and fold change was >4. Tables S4A and B contain details of all acetylated peptide identifications from the trypsin digestion and GluC digestion experiments, respectively. Tables S4C and D contain the unfiltered quantification results of all acetylation sites quantified from trypsin digestion and GluC digestion experiments, respectively. Table S4E contains details of all proteins identified from ProteinPilot used for spectral library building and protein-level quantification. Table S1D gives the protein-level changes due to YfiQ overexpression. Table S2B shows the site-level acetylation changes in proteins related to glycolysis shown in Fig. 5.

KAT sequence alignment and homology modeling. A multiple sequence alignment containing each *E. coli* KAT sequence (YfiQ, UniProt ID P76594; YiaC, UniProt ID P37664; YjaB, UniProt ID P09163; RimI, UniProt ID P0A944; and PhnO, UniProt ID P16691) was generated using the multiple alignment Clustal W function and manually modified in BioEdit (79). Only the GNAT domain (residues 726 to 881) of YfiQ was used in the sequence alignment. The final alignment figure was prepared using ESPript 3.0 (http://espript.ibcp.fr) (80). Homology models for YiaC, PhnO, and the GNAT domain of YfiQ were constructed using the ModWeb server (https://modbase.compbio.ucsf.edu/modweb/) (81) with the slow restraint selected for model generation. The models with the highest ModPipe Quality Score (MPQS), the lowest Discrete Optimized Protein Energy (zDOPE) value, a GA341 model score that was closest to 1, and the highest sequence identity were chosen for further analysis. The templates used for each of the final homology models were PDB ID 2kcw for YiaC, PDB ID 1z4e for PhnO, and PDB ID 4nxy for the GNAT domain of YfiQ.

Motility-related assays. Cultures were grown in TB (10 g/liter tryptone, 5 g/liter NaCl) at 37°C to exponential phase (0.3 to 0.5 OD_{600}) and were normalized to 0.3 OD_{600}, and a 5-μl aliquot was spotted onto the surface of a tryptone agar plate (10 g/liter tryptone, 5 g/liter NaCl, 2 g/liter agar). The diameter of the spot was measured hourly. For strains harboring plasmids, IPTG and chloramphenicol were added to both the growth medium and agar plates at a final concentration of 50 μM and 25 μg/ml, respectively.

Data availability. All raw mass spectrometry data files are available from public repositories (MassIVE ID number MSV000082411 and password kitkats and ProteomeXchange PXD009940). The MassIVE repository also includes supplemental tables and details of proteins and peptides that were identified and quantified by mass spectrometric analysis. Skyline files containing spectral libraries and chromatograms of raw data quantification are available on Panorama (https://panoramaweb.org/KAT.url, email login panorama+schilling@proteinms.net and password ^x3GfCJh).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mBio.01905-18.

FIG S1, TIF file, 0.3 MB.
FIG S2, TIF file, 1.9 MB.
FIG S3, TIF file, 5.1 MB.
FIG S4, TIF file, 0.3 MB.
FIG S5, TIF file, 0.1 MB.
FIG S6, TIF file, 0.3 MB.
TABLE S1, XLSX file, 3.5 MB.
TABLE S2, XLSX file, 0.05 MB.
TABLE S3, XLSX file, 0.4 MB.
TABLE S4, XLSX file, 7.3 MB.

ACKNOWLEDGMENTS

We sincerely thank George Gassner, Wayne Anderson, and Ekaterina Filippova for their helpful discussions.

This project was funded by R01 GM066130 (NIGMS to A.J.W.), R01 AI108255 (NIAID subcontract to A.J.W.), R01 AI108255 (NIAID to B.S.), a Center for Computing for Life Sciences (CCLS) grant, and Startup funds (from SFSU to M.L.K.), and J.G.M. was supported by T32 AG000266, NIH, to Campisi.

REFERENCES

1. Drazic A, Myklebust LM, Ree R, Arnesen T. 2016. The world of protein acetylation. Biochim Biophys Acta 1864:1372–1401. https://doi.org/10.1016/j.bbapap.2016.06.007.
2. Soppa J. 2010. Protein acetylation in archaea, bacteria, and eukaryotes. Archaea 2010:1. https://doi.org/10.1155/2010/820681.
3. Wolfe AJ. 2016. Bacterial protein acetylation: new discoveries unanswered questions. Curr Genet 62:335–341. https://doi.org/10.1007/s00294-015-0552-4.
4. Hentchel KL, Escalante-Semerena JC. 2015. Acylation of biomolecules in prokaryotes: a widespread strategy for the control of biological function and metabolic stress. Microbiol Mol Biol Rev 79:321–346. https://doi.org/10.1128/MMBR.00020-15.
5. Kuhn ML, Zemaitaitis B, Hu LI, Sahu A, Sorensen D, Minasov G, Lima BP, Scholle M, Mrksich M, Anderson WF, Gibson BW, Schilling B, Wolfe AJ. 2014. Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation. PLoS One 9:e94816. https://doi.org/10.1371/journal.pone.0094816.
6. Weinert BT, Iesmantavicius V, Wagner SA, Schölz C, Gummesson B, Beli

P, Nyström T, Choudhary C. 2013. Acetyl-phosphate is a critical determinant of lysine acetylation in E. coli. Mol Cell 51:265–272. https://doi.org/10.1016/j.molcel.2013.06.003.
7. de Diego Puente T, Gallego-Jara J, Castaño-Cerezo S, Bernal Sánchez V, Fernández Espín V, García de la Torre J, Manjón Rubio A, Cánovas Díaz M. 2015. The protein acetyltransferase PatZ from Escherichia coli is regulated by autoacetylation-induced oligomerization. J Biol Chem 290:23077–23093. https://doi.org/10.1074/jbc.M115.649806.
8. Starai VJ, Escalante-Semerena JC. 2004. Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in Salmonella enterica. J Mol Biol 340:1005–1012. https://doi.org/10.1016/j.jmb.2004.05.010.
9. Hu LI, Chi BK, Kuhn ML, Filippova EV, Walker-Peddakotla AJ, Bäsell K, Becher D, Anderson WF, Antelmann H, Wolfe AJ. 2013. Acetylation of the response regulator RcsB controls transcription from a small RNA promoter. J Bacteriol 195:4174–4186. https://doi.org/10.1128/JB.00383-13.
10. Lima BP, Antelmann H, Gronau K, Chi BK, Becher D, Brinsmade SR, Wolfe AJ. 2011. Involvement of protein acetylation in glucose-induced transcription of a stress-responsive promoter. Mol Microbiol 81:1190–1204. https://doi.org/10.1111/j.1365-2958.2011.07742.x.
11. Liang W, Malhotra A, Deutscher MP. 2011. Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. Mol Cell 44:160–166. https://doi.org/10.1016/j.molcel.2011.06.037.
12. Song L, Wang G, Malhotra A, Deutscher MP, Liang W. 2016. Reversible acetylation on Lys501 regulates the activity of RNase II. Nucleic Acids Res 44:1979–1988. https://doi.org/10.1093/nar/gkw053.
13. Zhang Q, Zhou A, Li S, Ni J, Tao J, Lu J, Wan B, Li S, Zhang J, Zhao S, Zhao G-P, Shao F, Yao Y-F. 2016. Reversible lysine acetylation is involved in DNA replication initiation by regulating activities of initiator DnaA in Escherichia coli. Sci Rep 6:30837. https://doi.org/10.1038/srep30837.
14. Castaño-Cerezo S, Bernal V, Blanco-Catalá J, Iborra JL, Cánovas M. 2011. cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in Escherichia coli. Mol Microbiol 82:1110–1128. https://doi.org/10.1111/j.1365-2958.2011.07873.x.
15. Liang W, Deutscher MP. 2012. Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). RNA 18:37–41. https://doi.org/10.1261/rna.030213.111.
16. Favrot L, Blanchard JS, Vergnolle O. 2016. Bacterial GCN5-related N-acetyltransferases: from resistance to regulation. Biochemistry 55:989–1002. https://doi.org/10.1021/acs.biochem.5b01269.
17. Escalante-Semerena JC. 2010. Nε-lysine acetylation control conserved in all three life domains. Microbe 5:340–344. https://doi.org/10.1128/microbe.5.340.1.
18. Vetting MW, S de Carvalho LP, Yu M, Hegde SS, Magnet S, Roderick SL, Blanchard JS. 2005. Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys 433:212–226. https://doi.org/10.1016/j.abb.2004.09.003.
19. Barak R, Welch M, Yanovsky A, Oosawa K, Eisenbach M. 1992. Acetyladenylate or its derivative acetylates the chemotaxis protein CheY in vitro and increases its activity at the flagellar switch. Biochemistry 31:10099–10107. https://doi.org/10.1021/bi00156a033.
20. Chan CH, Garrity J, Crosby HA, Escalante-Semerena JC. 2011. In Salmonella enterica, the sirtuin-dependent protein acylation/deacylation system (SDPADS) maintains energy homeostasis during growth on low concentrations of acetate. Mol Microbiol 80:168–183. https://doi.org/10.1111/j.1365-2958.2011.07566.x.
21. Vetting MW, Park CH, Hegde SS, Jacoby GA, Hooper DC, Blanchard JS. 2008. Mechanistic and structural analysis of aminoglycoside N-acetyltransferase AAC(6′)-Ib and its bifunctional, fluoroquinolone-active AAC(6′)-Ib-cr variant. Biochemistry 47:9825–9835. https://doi.org/10.1021/bi800664x.
22. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053.
23. Schilling B, Christensen D, Davis R, Sahu AK, Hu LI, Walker-Peddakotla A, Sorensen DJ, Zemaitaitis B, Gibson BW, Wolfe AJ. 2015. Protein acetylation dynamics in response to carbon overflow in Escherichia coli. Mol Microbiol 98:847–863. https://doi.org/10.1111/mmi.13161.
24. Meyer JG, Mukkamalla S, Steen H, Nesvizhskii AI, Gibson BW, Schilling B. 2017. PIQED: automated identification and quantification of protein modifications from DIA-MS data. Nat Methods 14:646–647. https://doi.org/10.1038/nmeth.4334.
25. Schilling B, Gafni J, Torcassi C, Cong X, Row RH, LaFevre-Bernt MA, Cusack MP, Ratovitski T, Hirschhorn R, Ross CA, Gibson BW, Ellerby LM. 2006. Huntingtin phosphorylation sites mapped by mass spectrometry. Modulation of cleavage and toxicity. J Biol Chem 281:23686–23697. https://doi.org/10.1074/jbc.M513507200.
26. Swaney DL, Wenger CD, Coon JJ. 2010. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. J Proteome Res 9:1323–1329. https://doi.org/10.1021/pr900863u.
27. Meyer JG, Kim S, Maltby DA, Ghassemian M, Bandeira N, Komives EA. 2014. Expanding proteome coverage with orthogonal-specificity α-lytic proteases. Mol Cell Proteomics 13:823–835. https://doi.org/10.1074/mcp.M113.034710.
28. Huang DW, Sherman BT, Lempicki RA. 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13. https://doi.org/10.1093/nar/gkn923.
29. Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57. https://doi.org/10.1038/nprot.2008.211.
30. Izu H, Adachi O, Yamada M. 1997. Gene organization and transcriptional regulation of the gntRKU operon involved in gluconate uptake and catabolism of Escherichia coli. J Mol Biol 267:778–793. https://doi.org/10.1006/jmbi.1996.0913.
31. Parichartanakul NM. 2004. Pathway to allostery: differential routes for allosteric communication in phosphofructokinase from Escherichia coli. Texas A&M University, College Station, TX.
32. Serre MC, Garel JR. 1990. Role of the C-terminal region in the allosteric properties of Escherichia coli phosphofructokinase-1. Eur J Biochem 189:487–492. https://doi.org/10.1111/j.1432-1033.1990.tb15513.x.
33. Rypniewski WR, Evans PR. 1989. Crystal structure of unliganded phosphofructokinase from Escherichia coli. J Mol Biol 207:805–821. https://doi.org/10.1016/0022-2836(89)90246-5.
34. Gardner NW, Monroe LK, Kihara D, Park C. 2016. Energetic coupling between ligand binding and dimerization in Escherichia coli phosphoglycerate mutase. Biochemistry 55:1711–1723. https://doi.org/10.1021/acs.biochem.5b00980.
35. Gardner NW, McGinness SM, Panchal J, Topp EM, Park C. 2017. A cooperative folding unit as the structural link for energetic coupling within a protein. Biochemistry 56:6555–6564. https://doi.org/10.1021/acs.biochem.7b00850.
36. Tiwari SP, Reuter N. 2016. Similarity in shape dictates signature intrinsic dynamics despite no functional conservation in TIM barrel enzymes. PLoS Comput Biol 12:e1004834. https://doi.org/10.1371/journal.pcbi.1004834.
37. Zerrad L, Merli A, Schröder GF, Varga A, Gráczer É, Pernot P, Round A, Vas M, Bowler MW. 2011. A spring-loaded release mechanism regulates domain movement and catalysis in phosphoglycerate kinase. J Biol Chem 286:14040–14048. https://doi.org/10.1074/jbc.M110.206813.
38. Majorek KA, Osinski T, Tran DT, Revilla A, Anderson WF, Minor W, Kuhn ML. 2017. Insight into the 3D structure and substrate specificity of previously uncharacterized GNAT superfamily acetyltransferases from pathogenic bacteria. Biochim Biophys Acta 1865:55–64. https://doi.org/10.1016/j.bbapap.2016.10.011.
39. Eddy SR. 1998. Profile hidden Markov models. Bioinformatics 14:755–763. https://doi.org/10.1093/bioinformatics/14.9.755.
40. Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I. 2015. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic Acids Res 43:3872. https://doi.org/10.1093/nar/gkv278.
41. Venkat S, Gregory C, Gan Q, Fan C. 2017. Biochemical characterization of the lysine acetylation of tyrosyl-tRNA synthetase in Escherichia coli. Chembiochem 18:1928–1934. https://doi.org/10.1002/cbic.201700343.
42. Venkat S, Gregory C, Sturges J, Gan Q, Fan C. 2017. Studying the lysine acetylation of malate dehydrogenase. J Mol Biol 429:1396–1405. https://doi.org/10.1016/j.jmb.2017.03.027.
43. Kim IY, Veres Z, Stadtman TC. 1993. Biochemical analysis of Escherichia coli selenophosphate synthetase mutants. Lysine 20 is essential for catalytic activity and cysteine 17/19 for 8-azido-ATP derivatization. J Biol Chem 268:27020–27025.
44. Rafi S, Novichenok P, Kolappan S, Stratton CF, Rawat R, Kisker C, Simmerling C, Tonge PJ. 2006. Structure of acyl carrier protein bound to FabI, the FASII enoyl reductase from Escherichia coli. J Biol Chem 281:39285–39293. https://doi.org/10.1074/jbc.M608758200.
45. Gur M, Madura JD, Bahar I. 2013. Global transitions of proteins explored by a multiscale hybrid methodology: application to adenylate kinase. Biophys J 105:1643–1652. https://doi.org/10.1016/j.bpj.2013.07.058.
46. Rege VD, Kredich NM, Tai CH, Karsten WE, Schnackerz KD, Cook PF. 1996.

A change in the internal aldimine lysine (K42) in O-acetylserine sulfhydrylase to alanine indicates its importance in transimination and as a general base catalyst. Biochemistry 35:13485–13493. https://doi.org/10.1021/bi961517j.

47. Rowland EA, Greco TM, Snowden CK, McCabe AL, Silhavy TJ, Cristea IM. 2017. Sirtuin lipoamidase activity is conserved in bacteria as a regulator of metabolic enzyme complexes. mBio 8:e01096-17. https://doi.org/10.1128/mBio.01096-17.
48. Isono K, Isono S. 1980. Ribosomal protein modification in Escherichia coli. II. Studies of a mutant lacking the N-terminal acetylation of protein S18. Mol Gen Genet 177:645–651.
49. Yoshikawa A, Isono S, Sheback A, Isono K. 1987. Cloning and nucleotide sequencing of the genes riml and rimJ which encode enzymes acetylating ribosomal proteins S18 and S5 of Escherichia coli K12. Mol Gen Genet 209:481–488. https://doi.org/10.1007/BF00331153.
50. Pathak D, Bhat AH, Sapehia V, Rai J, Rao A. 2016. Biochemical evidence for relaxed substrate specificity of Nα-acetyltransferase (Rv3420c/riml) of Mycobacterium tuberculosis. Sci Rep 6:28892. https://doi.org/10.1038/srep28892.
51. Errey JC, Blanchard JS. 2006. Functional annotation and kinetic characterization of PhnO from Salmonella enterica. Biochemistry 45:3033–3039. https://doi.org/10.1021/bi052297p.
52. Hove-Jensen B, McSorley FR, Zechel DL. 2012. Catabolism and detoxification of 1-aminoalkylphosphonic acids: N-acetylation by the phnO gene product. PLoS One 7:e46416. https://doi.org/10.1371/journal.pone.0046416.
53. Metcalf WW, Wanner BL. 1993. Mutational analysis of an Escherichia coli fourteen-gene operon for phosphonate degradation, using TnphoA′ elements. J Bacteriol 175:3430–3442. https://doi.org/10.1128/jb.175.11.3430-3442.1993.
54. Castaño-Cerezo S, Bernal V, Post H, Fuhrer T, Cappadona S, Sánchez-Díaz NC, Sauer U, Heck AJ, Altelaar AF, Cánovas M. 2014. Protein acetylation affects acetate metabolism, motility and acid stress response in Escherichia coli. Mol Syst Biol 10:762. https://doi.org/10.15252/msb.20145227.
55. Thao S, Chen CS, Zhu H, Escalante-Semerena JC. 2010. Nε-lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. PLoS One 5:e15123. https://doi.org/10.1371/journal.pone.0015123.
56. Ito A, May T, Kawata K, Okabe S. 2008. Significance of rpoS during maturation of Escherichia coli biofilms. Biotechnol Bioeng 99:1462–1471. https://doi.org/10.1002/bit.21695.
57. Oberto J, Nabti S, Jooste V, Mignot H, Rouviere-Yaniv J. 2009. The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. PLoS One 4:e4367. https://doi.org/10.1371/journal.pone.0004367.
58. Weber MM, French CL, Barnes MB, Siegele DA, McLean RJ. 2010. A previously uncharacterized gene, yjfO (bsmA), influences Escherichia coli biofilm formation and stress response. Microbiology 156:139–147. https://doi.org/10.1099/mic.0.031468-0.
59. Ito A, May T, Taniuchi A, Kawata K, Okabe S. 2009. Localized expression profiles of rpoS in Escherichia coli biofilms. Biotechnol Bioeng 103:975–983. https://doi.org/10.1002/bit.22305.
60. Domka J, Lee J, Bansal T, Wood TK. 2007. Temporal gene-expression in Escherichia coli K-12 biofilms. Environ Microbiol 9:332–346. https://doi.org/10.1111/j.1462-2920.2006.01143.x.
61. Ito A, Taniuchi A, May T, Kawata K, Okabe S. 2009. Increased antibiotic resistance of Escherichia coli in mature biofilms. Appl Environ Microbiol 75:4093–4100. https://doi.org/10.1128/AEM.02949-08.
62. Cho BK, Zengler K, Qiu Y, Park YS, Knight EM, Barrett CL, Gao Y, Palsson B. 2009. The transcription unit architecture of the Escherichia coli genome. Nat Biotechnol 27:1043–1049. https://doi.org/10.1038/nbt.1582.
63. Jozefczuk S, Klie S, Catchpole G, Szymanski J, Cuadros-Inostroza A, Steinhauser D, Selbig J, Willmitzer L. 2010. Metabolomic and transcriptomic stress response of Escherichia coli. Mol Syst Biol 6:364. https://doi.org/10.1038/msb.2010.18.
64. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 97:6640–6645. https://doi.org/10.1073/pnas.120163297.
65. Kumari S, Beatty CM, Browning DF, Busby SJ, Simel EJ, Hovel-Miner G, Wolfe AJ. 2000. Regulation of acetyl coenzyme A synthetase in Escherichia coli. J Bacteriol 182:4173–4179. https://doi.org/10.1128/JB.182.15.4173-4179.2000.
66. Wolfe AJ, Chang DE, Walker JD, Seitz-Partridge JE, Vidaurri MD, Lange CF, Prüss BM, Henk MC, Larkin JC, Conway T. 2003. Evidence that acetyl phosphate functions as a global signal during biofilm development. Mol Microbiol 48:977–988. https://doi.org/10.1046/j.1365-2958.2003.03457.x.
67. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008. https://doi.org/10.1038/msb4100050.
68. Silhavy TJ, Berman M, Enquist L. 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
69. Kitagawa M, Ara T, Arifuzzaman M, Ioka-Nakamichi T, Inamoto E, Toyonaga H, Mori H. 2005. Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. DNA Res 12:291–299. https://doi.org/10.1093/dnares/dsi012.
70. Schilling B, Rardin MJ, MacLean BX, Zawadzka AM, Frewen BE, Cusack MP, Sorensen DJ, Bereman MS, Jing E, Wu CC, Verdin E, Kahn CR, Maccoss MJ, Gibson BW. 2012. Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. Mol Cell Proteomics 11:202–214. https://doi.org/10.1074/mcp.M112.017707.
71. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R. 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics 11:O111.016717. https://doi.org/10.1074/mcp.O111.016717.
72. Rardin MJ, Schilling B, Cheng LY, MacLean BX, Sorensen DJ, Sahu AK, MacCoss MJ, Vitek O, Gibson BW. 2015. MS1 peptide ion intensity chromatograms in MS2 (SWATH) data independent acquisitions. improving post acquisition analysis of proteomic experiments. Mol Cell Proteomics 14:2405–2419. https://doi.org/10.1074/mcp.O115.048181.
73. Schilling B, Gibson BW, Hunter CL. 2017. Generation of high-quality SWATH. Methods Mol Biol 1550:223–233. https://doi.org/10.1007/978-1-4939-6747-6_16.
74. Deutsch EW, Mendoza L, Shteynberg D, Slagel J, Sun Z, Moritz RL. 2015. Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. Proteomics Clin Appl 9:745–754. https://doi.org/10.1002/prca.201400164.
75. Kim S, Pevzner PA. 2014. MS-GF+ makes progress towards a universal database search tool for proteomics. Nat Commun 5:5277. https://doi.org/10.1038/ncomms6277.
76. Tsou CC, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras AC, Nesvizhskii AI. 2015. DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. Nat Methods 12:258–264. https://doi.org/10.1038/nmeth.3255.
77. Teo G, Kim S, Tsou CC, Collins B, Gingras AC, Nesvizhskii AI, Choi H. 2015. mapDIA: preprocessing and statistical analysis of quantitative proteomics data from data independent acquisition mass spectrometry. J Proteomics 129:108–120. https://doi.org/10.1016/j.jprot.2015.09.013.
78. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26:966–968. https://doi.org/10.1093/bioinformatics/btq054.
79. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95–98.
80. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:W320–W324. https://doi.org/10.1093/nar/gku316.
81. Pieper U, Webb BM, Barkan DT, Schneidman-Duhovny D, Schlessinger A, Braberg H, Yang Z, Meng EC, Pettersen EF, Huang CC, Datta RS, Sampathkumar P, Madhusudhan MS, Sjölander K, Ferrin TE, Burley SK, Sali A. 2011. ModBase, a database of annotated comparative protein structure models, and associated resources. Nucleic Acids Res 39:D465–D474. https://doi.org/10.1093/nar/gkq1091.
82. Christensen DG, Orr JS, Rao CV, Wolfe AJ. 2017. Increasing growth yield and decreasing acetylation in Escherichia coli by optimizing the carbon-to-magnesium ratio in peptide-based media. Appl Environ Microbiol 83:e03034-16. https://doi.org/10.1128/AEM.03034-16.
83. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. 2007. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. Mol Cell Proteomics 6:2212–2229. https://doi.org/10.1074/mcp.M700354-MCP200.