Insights Into the Specificity of Lysine Acetyltransferases

Alex C. Tucker1,3, Keenan C. Taylor2, Katherine C. Rank2,4, Ivan Rayment2, Jorge C. Escalante-Semerena1 *

1From the Department of Microbiology, University of Georgia, Athens, GA 30602
2Department of Biochemistry, University of Wisconsin, Madison, WI 53706, USA
3Present address: Department of Biological Engineering, MIT, Cambridge, MA 02139
4Present address: Pharmaceutical Product Development, LLC (PPD), Middleton, WI 53562

*Running title: Crystal structure of a protein acetylation complex

To whom correspondence should be addressed: Jorge C. Escalante-Semerena, Department of Microbiology, University of Georgia, 212C Biological Sciences Building, 120 Cedar St, Athens, GA 30602-2605 USA, Tel: (706) 542-2651; E-mail: jcescala@uga.edu

Keywords: Bacterial metabolism, enzyme inactivation, posttranslational modification, acetyl-Coenzyme A, histone acetyltransferase, substrate specificity, substrate recognition, reversible lysine acetylation

Background: Gcn5-related N-acetyltransferases (GNATs) modify proteins in all domains of life. GNATs are known for their role as histone acetyltransferases, but non-histone bacterial protein acetyltransferases have been identified. Only structures of GNAT complexes with short histone peptide substrates are available in databases. Given the biological importance of this modification and the abundance of lysine in polypeptides, how specificity is attained for larger protein substrates is central to understanding acetyl-lysine regulated networks. Here we report the structure of a GNAT in complex with a globular protein substrate solved to 1.9Å. GNAT binds the protein substrate with extensive surface interactions distinct from those reported for GNAT-peptide complexes.

Results: The structure of a GNAT was determined in complex with a protein substrate.

Conclusion: Specificity of the GNAT-protein interaction is dictated by an extensive interaction surface compared to GNAT-peptide structures.

Significance: This is the first structure of a GNAT-protein acetylation complex, and it may enable structure-based identification and engineering of GNAT substrates.

Reversible lysine acetylation by protein acetyltransferases is a conserved regulatory mechanism that controls diverse cellular pathways. Gcn5-related N-acetyltransferases (GNATs), named after their founding member, are found in all domains of life. GNATs are known for their role as histone acetyltransferases, but non-histone bacterial protein acetyltransferases have been identified. Only structures of GNAT complexes with short histone peptide substrates are available in databases. Given the biological importance of this modification and the abundance of lysine in polypeptides, how specificity is attained for larger protein substrates is central to understanding acetyl-lysine regulated networks. Here we report the structure of a GNAT in complex with a globular protein substrate solved to 1.9Å. GNAT binds the protein substrate with extensive surface interactions distinct from those reported for GNAT-peptide complexes.

Our data reveal determinants needed for the recognition of a protein substrate and provide insight into the specificity of GNATs.

Organisms from all domains of life acetylate proteins to dynamically regulate processes such as chromatin maintenance (1), transcriptional regulation (2), protein stability (3), primary metabolism (4), and cell structure (5). The enzymes responsible for this posttranslational modification are classified into five different groups that include the Gcn5-related-N-acetyltransferases (GNATs) (6). The GNAT superfamily is enormously versatile. Although important members are histone acetyltransferases, the domain was first identified in aminoglycoside acetyltransferases and more recently in non-histone bacterial protein acetyltransferase (Pat) enzymes (7). To date, only structures of GNAT complexes with short peptide substrates are available in databases. GNATs are now known to acetylate a
diversity of protein substrates. Thus, a detailed characterization of the GNAT-substrate interface is critical for understanding GNAT specificity for protein targets given the prevalence of lysine in biological proteomes.

In bacteria many members of the multi-domain AMP-forming acyl CoA synthetase family are regulated by Nε-lysine acetylation as was first identified in Salmonella enterica for acetyl-CoA synthetase (SeAcsWT) (8). Acetylation of residue K609 of SeAcsWT by the S. enterica protein acetyltransferase Pat (SePatWT) leads to enzyme inactivation that can be reversed by the CobB sirtuin-type NAD+-dependent deacetylase (Fig. 1A) (9). SePatWT consists of a GNAT catalytic domain and a larger domain of unknown function (Fig. 1B) (10). Studies of SePatWT homologues in other bacteria show that these acetyltransferases can regulate a wide range of acyl-CoA synthetases which all share a conserved PX4GK motif near the C-terminus of these proteins. The lysine in this motif is catalytic and is the target of acetylation. The roles of the proline and glycine residues are unknown and are not critical for acetyl-CoA synthetase activity (11). Interestingly, the presence of the PX4GK motif is necessary but not always sufficient for acetylation of these enzymes suggesting that additional determinants lie outside the signature sequence (12).

To gain insight into how specificity governs recognition and modification in a reversible lysine acetyl-lysine signaling pathway, we characterized the interaction interface between a GNAT and a protein substrate. We determined the structure of the GNAT domain of the Streptomyces lividans PatA (SlPatAWT) enzyme (Figs. 1C, 1D), and used the ClusPro 2.0 server to identify potential interacting surfaces between the GNAT domain of SlPatA (hereafter SlPatAGNAT) and the C-terminal domain of SeAcs (hereafter SeAcsCTD), a substrate of SlPatAWT for which a three-dimensional crystal structure was available (RCSB PDB # 1PG3, 1PG4) (13). We demonstrated that SlPatAGNAT acetylated SeAcsCTD in vivo and in vitro, and that reversing charges near a predicted interface prevented interaction and SeAcsCTD acetylation. To test the biological relevance of the interaction model, we tethered the SlPatAGNAT domain to SeAcsCTD in multiple orientations with cross-linkers that resulted in spacing between the proteins ranging between 2 and 20 Å. We identified a single orientation that resulted in acetylation of SeAcsCTD when SlPatAGNAT was linked to SeAcsCTD by a direct disulfide bond, demonstrating catalysis in an enzyme-substrate complex with limited movement. We report the structure of the catalytic complex comprised of the SlPatAGNAT domain and SeAcsCTD at 1.9Å resolution. The structure revealed a constellation of determinants needed for recognition of a protein substrate by the SlPatAGNAT domain.

Experimental Procedures

Bacterial strains and growth conditions. All strains, plasmids, and oligonucleotides used in this study are listed in Tables 1-3. Escherichia coli and Salmonella enterica strains were grown at 37°C in lysogeny broth (LB, Difco) (14) or no-carbon essential (NCE) minimal medium (15) supplemented with sodium acetate (10 mM), MgSO4 (1 mM), and ampicillin (100 µg/ml). When necessary, antibiotics were used at the following concentrations: ampicillin, 100 µg/ml; carbenicillin, 100 µg/ml; tetracycline, 10 µg/ml; kanamycin, 50 µg/ml. L-(+)-arabinose was added at varying concentrations (0-25 µM) to induce the expression of SlPatAGNAT cloned into the expression vector pBAD30 (16). Growth experiments were performed at 37°C using a microtiter plate and a microtiter plate reader (Bio-Tek Instruments).

Molecular techniques. DNA manipulations were performed using standard techniques (17). Restriction endonucleases were purchased from Fermentas. DNA was amplified using Pfu Ultra II Fusion DNA polymerase (Agilent) or Herculase II Fusion DNA polymerase (Agilent). Site-directed mutagenesis was performed using the Quikchange™ Site Directed Mutagenesis kit (Agilent). Plasmids were isolated using the Wizard Plus SV Miniprep kit (Promega) and PCR products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega). DNA sequencing was performed using BigDye® (ABI PRISM) protocols, and dye-terminator sequencing reactions were resolved and analyzed by capillary electrophoresis at the
University of Georgia Genomics Facility. Oligonucleotide primer sequences are listed in Table 3.

Construction of complementation plasmids. The allele encoding the GNAT domain of wild-type S/patA (residues M1-L194) was amplified from S. lividans TK24 genomic DNA using primers listed in Table 3. Primers used in the amplification changed the first codon of S/patA from TTG to the more common ATG start codon. The amplified fragment was cut with EcoRI and KpnI and ligated into pBAD30 cut with the same enzymes. The resulting plasmid, pSlPatA11, directed the synthesis of SlPatGNAT in response to L-(+)-arabinose. The plasmid directing the synthesis of variant SlPatAGNAT E123Q was generated by site-directed mutagenesis using primers listed in Table 3.

Plasmids for in vivo two-hybrid system assays. Alleles encoding the wild-type SlPatAGNAT domain, variant SlPatAGNAT E121R, and variant SlPatAGNAT D185R were amplified from plasmids pSlPatA27, pSlPatA33, and pSlPatA35, respectively, using primers listed in the Table 3. Alleles encoding wild-type SeAcsCTD (residues D518 to S652), variants SeAcsCTD R606E and SeAcsCTD R613D were amplified from plasmids pACS38, pACS42, and pACS44, respectively, using primers listed in the Table 3. DNA fragments were cut with NotI and BamHI, then ligated into plasmids pACTR-V-Zif-AP and pBRoGP, which had been cut with NotI and BamHI. The resulting plasmids are listed in Table 2.

SlPatAGNAT overproduction plasmids. The allele encoding the GNAT domain of SlPatA (residues M1-L194) was amplified from S. lividans TK24 genomic DNA with primers listed in Table 3. The primers used were designed to change the first codon of SlPatA from TTG to the more common ATG start codon. The amplified fragment was cut with KpnI and SalI and ligated into pKLD66 (18) cut with the same enzymes. The resulting plasmid pSlPatA14 directed the synthesis of wild-type SlPatAGNAT fused at its N-terminus to a H6-Maltose binding protein (MBP) tag cleavable by recombinant Tobacco Etch Virus (rTEV) protease using described protocols (19). Plasmids directing the synthesis of variants SlPatAGNAT S73C, SlPatAGNAT A110C, and SlPatAGNAT A164C were generated from plasmid pSlPatA14 using site-directed mutagenesis using primers listed in Table 3.

SeAcs C-terminal domain (SeAcsCTD) overexpression plasmids. The allele encoding SeAcsCTD (residues D518 to S652) was amplified from S. enterica LT2 DNA using primers listed in Table 3. The amplified fragment was cut with NheI and EcoRI and ligated into pTEV5 cut with the same enzymes. The resulting plasmid pACS38 directed the synthesis of wild-type SeAcsCTD fused at its N-terminus to a hexahistidine tag cleavable by rTEV protease. Plasmids directing synthesis of variants SeAcsCTD R606E, SeAcsCTD R613D, SeAcsCTD A238C, SeAcsCTD H567C, and SeAcsCTD D600C were generated from plasmid pACS38 by site-directed mutagenesis using primers listed in Table 3.

Full-length SeAcs overproduction plasmids. Plasmids directing the synthesis of variants SeAcsCTD R606E and SeAcsCTD R613D were generated from plasmid pACS33 (Table 2) using site-directed mutagenesis. The resulting plasmids pACS60 and pACS61 directed the synthesis of variants SeAcsCTD R606E and SeAcsCTD R613D, respectively, each fused to a N-terminal H6 tag cleavable by rTEV protease.

SlPatAGNAT overproduction and purification. Plasmids pSlPatA14, pSlPatA53, pSlPatA54, pSlPatA56 were transformed into the Δpka derivative of E. coli C41 (λDE3) (strain JE9314). The resulting strains were grown overnight and sub-cultured 1:100 (v/v) into 2 liters of LB containing ampicillin (100 µg/ml). The cultures were grown shaking at 25°C to A600 ~ 0.7 and H6-MBP-SlPatAGNAT synthesis was induced with IPTG (0.5 mM). Upon induction, the cultures were grown overnight at 25°C. Cells were harvested at 6000 x g for 10 min at 4°C in a Avanti J-2 XPI centrifuge fitted with rotor JLA-8.1000 (Beckman Coulter). Cell pellets were re-suspended in 30 ml cold His-Bind buffer A [tris(hydroxymethyl)aminomethane-HCl...
(Tris-HCl) buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (5 mM)) and Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (15 mM) and Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM). Cells were placed on ice and lysed by sonication for 1 min (2-s pulse followed by 4 s of cooling) at level 7 in a model 550 sonic dismembrator (Fisher). The extract was cleared by centrifugation at 4°C for 30 min at 43,367 × g. His6-tagged SlPatAGNAT were purified from the soluble fraction by Ni-affinity purification using a 1-ml bed volume of His-Pur Nickel-NTA Resin (Thermo). Unbound proteins were eluted off the column by extensive washing with buffer A. The column was then washed with buffer B1 [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (15 mM)] and SlPatAGNAT were purified from the soluble fraction by Ni-affinity purification using a 1-ml bed volume of His-Pur Nickel-NTA Resin (Thermo). Unbound proteins were eluted off the column by extensive washing with buffer A. The column was then washed with buffer B1 [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (250 mM)]. All fractions containing Hc-MBP-SlPatAGNAT were combined. The protease was added to Hc-MBP-SlPatAGNAT and the SlPatAGNAT/TEV mixture was incubated at 4°C overnight against buffer C (Tris-HCl (50 mM, pH 8), NaCl (500 mM), and imidazole (250 mM)). All fractions containing Hc-MBP-SlPatAGNAT were combined. The protease was added to Hc-MBP-SlPatAGNAT and the SlPatAGNAT/TEV mixture was incubated at room temperature for 3 h. The SlPatAGNAT/TEV mixture was dialyzed at 4°C against buffer C (Tris-HCl (50 mM, pH 8), NaCl (500 mM)) twice for 3 h and again against buffer C containing imidazole (5 mM) for 12 h. After cleavage and dialysis, protein mixtures were passed over the 1-ml HisTrap column using the buffers described above and the untagged proteins eluted in the flow-through and buffer A wash. Purified SlPatAGNAT was analyzed by SDS-PAGE. Fractions containing SlPatAGNAT were pooled together. SlPatAGNAT was stored in Tris-HCl buffer (50 mM, pH 8.0) containing NaCl (100 mM) and glycerol (10%, v/v). Tris(2-carboxyethyl)phosphine (TCEP, 0.3 mM) and ethylenediaminetetraacetate acid (EDTA, 0.5 mM) were included in the storage buffer for SlPatAGNAT cysteine variants. SlPatAGNAT concentration was determined by measuring absorbance at 280 nm. The molar extinction coefficient used to calculate SlPatAGNAT concentration was 17,420 M⁻¹ cm⁻¹. The purification protocol for SlPatAGNAT for crystallization was similar to the one described above except that the purified, untagged protein was dialyzed into Tris-HCl buffer (10 mM, pH 8.0) and concentrated to 11 mg/ml before flash freezing into liquid nitrogen.

Selenomethionine-labeled SlPatAGNAT was overproduced as follows. Plasmid pSlPatA14 was transformed into strain JE9314 and 1-L culture of the resulting strain was grown overnight in M9 glucose medium. The culture was used to re-inoculate 2 x 2L of fresh M9 glucose medium at a 1:100 inoculum and the culture was grown to an OD₆₀₀ of ∼1. The culture was cooled on ice to 16°C for 10 min, and a defined amino acid mixture containing lysine (100 mg/L), threonine (100 mg/l), phenylalanine (100 mg/L), leucine (50 mg/L), isoleucine (50 mg/L), valine (50 mg/L), and selenomethionine (50 mg/L) was added to suppress methionine biosynthesis. The culture was grown at 37°C for 30 min before the addition of IPTG (1 mM) to induce SlPatAGNAT expression. The culture was grown overnight at 37°C and the protein was purified and stored for crystallography as described above.

**Purification of SeAcsWT and SeAcsCTD.** Plasmids pACS38, pACS44, pACS56, pACS57, pACS58, pACS60, and pACS61 were transformed into a Δpka derivative of E. coli C41 (DE3) (JE9314) to prevent acetylation prior to overproduction. The resulting strains were grown overnight and sub-cultured 1:100 (v/v) into 2 liters of lysogeny broth (LB) containing ampicillin (100 µg/ml). The cultures were grown shaking at 25°C to A₆₀₀ ~ 0.7 and protein synthesis was induced with IPTG (0.25 mM). Upon induction, the cultures were grown overnight at 25°C. SeAcsWT and SeAcsCTD proteins were purified and stored as described above for SlPatAGNAT with modifications. During the first purification step, the His₆-SeAcs-bound resins were washed with buffer B [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (20 mM)] before His₆-SeAcs proteins were eluted with buffer C. His₆-SeAcsCTD-bound resins were washed with buffer B3 [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (40 mM)] before elution with buffer C. In the second purification step SeAcsWT proteins and SeAcsCTD proteins did not adsorb to the column and were present in the flow-through fractions. Proteins were stored in Tris-HCl buffer (50 mM, pH 8.0) containing NaCl (100 mM) and glycerol (10%, v/v). TCEP (0.3 mM) and EDTA (0.5 mM) were included in
the storage buffer for SlPatAGNAT cysteine variants. The molar extinction coefficients used to calculate protein concentrations were 72,152 M⁻¹ cm⁻¹ for SeAcsWT and 15,470 M⁻¹ cm⁻¹ for SeAcsCTD.

**Two-hybrid system assay and β-galactosidase activity measurement.** E. coli strain KDZif1ΔZ harboring compatible plasmids were grown overnight in 1 ml of nutrient broth (NB) supplemented with kanamycin and carbenicillin. The following day, strains were subcultured 1:100 into 200 ml of NB supplemented with kanamycin, carbenicillin, and isopropyl-β-D-1-thiogalactopyranoside (IPTG, 200 µM) to induce expression from two-hybrid screen plasmids. Strains were grown for 3.5 h at 37°C with shaking at medium intensity in a BioTek plate reader (BioTek Instruments, Inc.). Absorbance values of the cultures were measured in the microtiter plate at 650 nm wavelength using a Spectramax Plus 384 spectrophotometer (Molecular Devices).

The β-galactosidase activity of the cultures was measured as described (20, 21), with modifications. Cultures were diluted 20 ml into 80 ml of permeabilization solution (Na₂HPO₄ [100 mM], KCl [20 mM], MgCl₂ [2 mM], hexadecyltrimethylammonium bromide [CTAB, 0.06%], sodium deoxycholate [0.04%], 2-mercaptoethanol [38 mM]). Samples were shaken 5 s in the Spectramax Plus 384 spectrophotometer (Molecular Devices) to mix. Permeabilized cells were then diluted 25 ml into 150 ml of substrate solution (Na₂HPO₄ [60 mM], NaH₂PO₄ [40 mM], O-nitrophenyl-β-galactoside [ONPG, 1 mg/ml], 2-mercaptoethanol [38 mM]). The absorbance at 420 nm was monitored in a 96-well plate using the Spectramax Plus 384 spectrophotometer. A unit of activity was defined as the amount of enzyme required to catalyze the hydrolysis of 1 nmol of ONPG per min.

**AMP-forming CoA ligase assays.** SeAcsWT activity was measured using an NADH-consuming assay (22, 23). Reactions (100 µl total volume) contained 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (HEPES) buffer (50 mM, pH 7.5), tris(2-chloroethyl) phosphate (TCEP, 1 mM), ATP (2.5 mM), CoA (0.5 mM), MgCl₂ (5 mM), KCl (1 mM), phosphoenolpyruvate (3 mM), NADH (0.1 mM), pyruvate kinase (1 U), myokinase (5 U), lactate dehydrogenase (1.5 U) and acetate (0.2 mM). Reactions were started by the addition of SeAcsWT (15 nM). The absorbance at 340 nm was monitored in a 96-well plate using the Spectramax Plus UV-visible spectrophotometer (Molecular Devices). Enzyme specific activities were determined to be in the linear range for the assay and were calculated as described (22). Values are reported as the mean ± SD.

**In vitro acetylation.** Protein acetylation was monitored using radiolabeled [1-¹⁴C]-Ac-CoA (50 µM; 56.8 mCi mmol⁻¹) as described (23-25). Briefly, reactions contained SeAcsWT or SeAcsCTD (3 µM) and either SlPatWT or SlPatA (1 µM). Samples (10 µl) were removed and quenched with 2 µl loading buffer (glycerol (50%), EDTA (1 mM), bromophenol blue (0.25%), xylene cyanol (0.25%)) before being resolved and visualized with SDS-PAGE (26) and Coomassie Blue staining (27), respectively. Labeled proteins were visualized using a Typhoon FLA 9000 Variable Mode Imager (GE Healthcare) equipped with ImageQuant TL software (GE Healthcare). Acetylation was quantified as a measure of the radiolabel signal and reported as digital light units (DLU). Gels and phosphor images were cropped to the SeAcsWT, SeAcsCTD, or SeAcs variant bands.

The effect of acetylation on activities of SeAcsWT, SeAcsR⁶⁰⁶E, and SeAcsR⁶¹³D was determined as described (23) with modifications. SeAcsWT, SeAcsR⁶⁰⁶E, or SeAcsR⁶¹³D (3 µM) was incubated with SlPatA (1 µM) or SlPat (1 µM) and 50 µM Ac-CoA for 90 min at 30°C (SlPatA) or 37°C (SlPatA) using the buffer system described above. At 90 min, reactions were diluted 1:20 into HEPES buffer (50 mM, pH 7.5) at 4°C. SeAcs activity was measured as described above.

**SlPatAGNAT-SeAcsCTD cross-linking and acetylation assays.** Cysteine variants of wild-type SeAcsCTD were exchanged into cross-linking buffer (20 mM HEPES, pH 7.5, 100 mM...
NaCl, sparged with N₂) using Zeba Spin desalting columns (Pierce). SeAcs<sup>CTD</sup> proteins were immediately diluted into a 10-fold molar excess of aldrithiol-2 (AT-2, 2.05-Å cross-linker, Sigma), bis(maleimido)ethane (BMOE, 8.0 Å cross-linker, Thermo), 1,4-bis(maleimido)butane (BMB, 10.9 Å cross-linker, Thermo), bis(maleimido)hexane (BMH, 16.0 Å cross-linker, Thermo), or 1,4-di-(3’-[2’-pyridyldithio]-propionamido)butane (DPDPB, 19.9 Å cross-linker, bioWorld), and allowed to react for 1 h at room temperature. Activated SeAcs<sup>CTD</sup> variants and SlPatAGNAT cysteine variants were desalted into cross-linking buffer as described above. SlPatAGNAT variants were diluted into a 5-fold molar excess of SeAcs<sup>CTD</sup> and allowed to react for 2 h at 4°C. AT-2 and DPDPB crosslinking reactions were quenched with iodoacetamide (10 mM). BMOE, BMB, and BMH crosslinking reactions were quenched with dithiothreitol (DTT, 10 mM). Crosslinking reactions were analyzed by non-reducing SDS-PAGE to verify crosslinking and to quantify cross-linked species.

Cross-linked species were incubated with Bis-Tris (100 mM, pH 6.0) and [1-<sup>14</sup>C]-Ac-CoA (20 µM; 56.8 mCi mmol⁻¹) at 30°C for 1 h. Reactions were resolved by non-reducing SDS-PAGE and analyzed as described above. Gels and phosphor images were cropped to the wild-type SeAcs<sup>CTD</sup>-SlPatAGNAT complex bands.

**Crystallization of SlPatA<sup>GNAT</sup> and SlPatA-SmAcs<sup>CTD</sup> complex.** SlPatA<sup>GNAT</sup> crystals were grown by small-scale batch crystallization where 4 µL of concentrated protein at 10.7 mg/mL in Tris-HCl buffer (10 mM, pH 8.0) containing ethyl-CoA (2 mM) was mixed with 5 µL of PEG solution (22%, w/v) polyethylene glycol 5000 (22%, w/v), glycerol (20%, v/v), MOPS buffer (100 mM, pH 7.0) containing MgCl₂ (25 mM), 2,2,2-trifluoroethanol (2%, v/v), and ethyl-CoA (2 mM). Crystals were flash-frozen in liquid nitrogen. Selenomethionine protein crystals were grown under identical conditions except that all solutions lacked ethyl-CoA.

**SlPatA<sup>GNAT</sup>-SeAcs<sup>CTD</sup> complex crystals** were grown by hanging drop diffusion where 1 µL of concentrated protein at 10.0 mg/mL in Tris-HCl buffer (10 mM, pH 8.0) containing ethyl-CoA (2 mM) was mixed with 1 µL of 11.2% (w/v) polyethylene glycol 8000, 100 mM 1,4-piperazinediethanesulfonic acid pH 6.5, and 120 mM Li₂SO₄. Crystals formed spontaneously after 2 months and grew to a maximum dimension of 400 µm x 50 µm x 50 µm. For freezing, crystals were transferred to paratone-N and then flash frozen in liquid nitrogen.

**X-Ray data collection and structural refinement.** X-Ray diffraction data for the SlPatA<sup>GNAT</sup> crystals and SlPatA<sup>GNAT</sup>-SeAcs<sup>CTD</sup> complex crystals were collected at the SBC 19-ID and SBC 19-BM beamlines respectively (Advanced Photon Source, Argonne, IL). The data sets were integrated and scaled with the program HKL3000(29). X-ray data collection statistics are given in Table 4. The structure of the SlPatA<sup>GNAT</sup>-SeAcs<sup>CTD</sup> complex was solved by molecular replacement using SlPatA<sup>GNAT</sup> and SeAcs<sup>CTD</sup> (PDB # 1PG4) as search models where residues in loops were deleted. The structure of SlPatA<sup>GNAT</sup> was solved by single-wavelength anomalous diffraction (SAD) using crystals containing selenomethionine protein. The HKL3000 suite was used to build an initial model of seleno-methionine SlPatA<sup>GNAT</sup> utilizing the programs SHELX, mlphare, DM, and ARP/wARP (30-32). This initial structure was used without any further refinement as a molecular replacement model to determine the structure of the native protein using the program Phaser (32,33). The native structure of SlPatA<sup>GNAT</sup> and that of the SlPatA<sup>GNAT</sup>-SeAcs<sup>CTD</sup> complex were refined by iterative cycles of manual model building in Coot and restrained refinement in Refmac 5.6 (34,35). Data processing and refinement statistics are
Crystal structure of a protein acetylation complex

Results

SlPatA<sup>GNAT</sup> interacts with SeAcs<sup>CTD</sup>. To address the question of how GNAT specificity is achieved, the interaction between SlPatA<sup>GNAT</sup> with SeAcs<sup>CTD</sup> was investigated. Our attempts to crystallize SePat were not met with success, so the protein acetyltransferase from <i>S. lividans</i> (SlPatA<sup>GNAT</sup>) was chosen. SlPatA<sup>GNAT</sup> proved amenable to structural studies and previous studies showed that SlPatA<sup>WT</sup> acetylates SeAcs<sup>WT</sup> in vitro (10), even though the domain organization of SePat<sup>WT</sup> and SlPatA<sup>WT</sup> is reversed (Fig. 1B). SlPatA<sup>GNAT</sup> was sufficient for functionality in vivo through its ability to substitute for SePat<sup>WT</sup> during growth on 10 mM acetate. As expected, expression of SlPatA<sup>GNAT</sup> inhibited growth of a <i>S. enterica</i> Δpat ΔcobB strain, but allowed growth of a <i>S. enterica</i> Δpat cobB<sup>+</sup> strain, which retained the ability to deacetylate acetyllysine (Figs. 1A, 1E).

Interaction model for SlPatA<sup>GNAT</sup> and SeAcs<sup>CTD</sup>. Initial attempts to co-crystallize SlPatA<sup>GNAT</sup> with SeAcs<sup>CTD</sup> (contains the target K609) were unsuccessful. This problem was overcome by introducing a covalent linkage between the SlPatA<sup>GNAT</sup> and SeAcs<sup>CTD</sup> domains at a position that neither affected the enzymatic activity of SlPatA<sup>GNAT</sup> nor the formation of the SlPatA<sup>GNAT</sup>/SeAcs<sup>CTD</sup> binary complex (described below). The location for the linkage was identified by first creating a computational model for the complex from the high-resolution structures of the individual domains.

The three-dimensional structure of SlPatA<sup>GNAT</sup> was determined to 1.5 Å (Fig. 1C, 1D, Table 4). The structure of SlPatA<sup>GNAT</sup> revealed a characteristic mixed α/β GNAT fold that contained the conserved acetyl-CoA binding site including the catalytic residue E123. SlPatA<sup>GNAT</sup> residues F126 and M168 overlapped with the modeled Ac-CoA structure, thus are likely to undergo a shift upon Ac-CoA binding, as observed for the analogous residues of the Ac-CoA bound structure of MtPatA (F238 and M280). The most similar structure to SlPatA<sup>GNAT</sup> was that of the GNAT domain from Sulfolobus solfataricus Pat (PDB # 3F8K) with an RMSD of 1.32Å over 131 residues. The major secondary structure differences between the SsPat GNAT and SlPatA<sup>GNAT</sup> exist along the protein substrate-binding surface.

The structure of SlPatA<sup>GNAT</sup> determined here was combined with the previously reported structures of SeAcs<sup>WT</sup> (PDB # 1PG3, 1PG4) (13) to generate computer models of the interaction interface using the ClusPro 2.0 server (36-39). The models were evaluated by requiring that the distance between the SlPatA<sup>GNAT</sup> catalytic residue E123 and the ε-amino group of the target lysine in the PX<sub>K</sub>GK motif be similar to that of 8 Å observed in the crystal structure of the Tetrahymena Gen5 bound to the H3 peptide substrate (PDB # 1QSN). The best computational models placed residue E123 within 15Å of the α-carbon of the target lysine (K609 side chain was not resolved in the SeAcs<sup>WT</sup> structure) (40). The best model for SeAcs<sup>WT</sup>-SlPatA<sup>GNAT</sup> interactions is shown in figure 1F. Notably, SlPatA<sup>GNAT</sup> was predicted to interact predominantly with SeAcs<sup>CTD</sup>. Figure 2A shows the best model for the interaction between SlPatA<sup>GNAT</sup> domain and the SeAcs<sup>CTD</sup> (residues D518 to S652).

SlPatA<sup>GNAT</sup>-SeAcs<sup>CTD</sup> crosslinking reveals a selective orientation for enzyme-substrate interaction. The validity of the interaction model was tested by introducing a series of crosslinks between SlPatA<sup>GNAT</sup> and SeAcs<sup>CTD</sup> in nine orientations with linkers of distinct lengths, which included one orientation that approximated the ClusPro model shown in figure 2A. This strategy allowed us to examine the transfer of the acetyl moiety within the SlPatA<sup>GNAT</sup> and SeAcs<sup>CTD</sup> complexes as a function of cross-linker length and orientation. Residues were chosen at the periphery of the predicted interaction surfaces of SlPatA<sup>GNAT</sup> and SeAcs<sup>CTD</sup> to avoid substitutions that might otherwise disturb the putative interface. A single-cysteine variant was constructed for each chosen residue, which included residues S73,
A110, A164 in \( \text{SI} \text{Pat}^{\text{GNAT}} \) and residues A538, H567, D600 in \( \text{Se} \text{Ac}^{\text{CTD}} \) (Figs. 2A, 2B). Purified single-Cys variants were cross-linked using sulfhydryl-specific cross-linkers with reported lengths of 2.05 Å to 19.9 Å (Fig. 2C). Crosslinking in all nine configurations was successful with at least four different cross-linkers for each combination (Fig. 2E, Coomassie Blue-stained images).

The ability of the \( \text{SI} \text{Pat}^{\text{GNAT}} \) to acetylate \( \text{Se} \text{Ac}^{\text{CTD}} \) in each cross-linked complex was used as a measure of the biochemical relevance of the orientation allowed by the cross-linker length constraints. In the absence of a cross-link, \( \text{SI} \text{Pat}^{\text{GNAT}} \) acetylation of \( \text{Se} \text{Ac}^{\text{CTD}} \) was efficient only when \( \text{SI} \text{Pat}^{\text{GNAT}} \) was in 10-fold molar excess of the \( \text{Se} \text{Ac}^{\text{CTD}} \) (Fig. 2D). As seen in figure 2E, \( \text{SI} \text{Pat}^{\text{GNAT}} \) acetylated \( \text{Se} \text{Ac}^{\text{CTD}} \) in nearly all orientations tested when the spacer length was ≥8 Å. The single exception was the complex between variants \( \text{SI} \text{Pat}^{\text{GNAT}} \) A110C-\( \text{Se} \text{Ac}^{\text{CTD}} \) A538C, in which we did not observe acetylation for any of the crosslinked complexes tested. Notably, only complex \( \text{SI} \text{Pat}^{\text{GNAT}} \) S73C-\( \text{Se} \text{Ac}^{\text{CTD}} \) H567C resulted in detectable acetylation when a direct disulfide bond between the cysteine residues of each protein held the complex together. This direct disulfide bond severely restricted the interactions between \( \text{SI} \text{Pat}^{\text{GNAT}} \) and \( \text{Se} \text{Ac}^{\text{CTD}} \), and yet the acetylation signal was strong (Fig. 2E). Thus, we hypothesized that the protein orientations in the \( \text{SI} \text{Pat}^{\text{GNAT}} \) S73C-\( \text{Se} \text{Ac}^{\text{CTD}} \) H567C complex reflected the interactions of these proteins in vivo. Significantly, residues S73 of \( \text{SI} \text{Pat}^{\text{GNAT}} \) and H567 of \( \text{Se} \text{Ac}^{\text{CTD}} \) were positioned near one another in the ClusPro interaction model (Fig. 2A).

**Crystal structure of the \( \text{SI} \text{Pat}^{\text{GNAT}}-\text{Se} \text{Ac}^{\text{CTD}} \) complex at 1.9 Å resolution.** To visualize how variants \( \text{SI} \text{Pat}^{\text{GNAT}} \) S73C and \( \text{Se} \text{Ac}^{\text{CTD}} \) H567C interacted, we crystallized the \( \text{SI} \text{Pat}^{\text{GNAT}} \) S73C-\( \text{Se} \text{Ac}^{\text{CTD}} \) H567C complex formed with an 8-Å linker, and the structure was determined to 1.9 Å (Fig. 3A-D, Table 4). In the of \( \text{SI} \text{Pat}^{\text{GNAT}} \) S73C-\( \text{Se} \text{Ac}^{\text{CTD}} \) H567C structure, the catalytic E123 residue of \( \text{SI} \text{Pat}^{\text{GNAT}} \) was 4.7 Å away from the target K609 of \( \text{Se} \text{Ac}^{\text{CTD}} \) (Fig. 3D,E), a distance within the range observed for Tetrahymena Gcn5 bound to a lysine-containing peptide substrate (40,41).

The interface between \( \text{SI} \text{Pat}^{\text{GNAT}} \) and \( \text{Se} \text{Ac}^{\text{CTD}} \) included more than just the interactions between the K609-containing \( \text{Se} \text{Ac}^{\text{CTD}} \) loop and the primary active site of \( \text{SI} \text{Pat}^{\text{GNAT}} \), for example, it included interactions that are well separated in the primary sequence. The interface shows good shape complementarity with a shape correlation statistic (Sc) of 0.54 that is similar to 0.60 observed in the Gcn5 H3 peptide complex (1QSN) (42). These values fall within the expected range for this type of complex where an antibody/antigen complex results in an Sc value from 0.64 to 0.68 and while an aberrant interface will result in an Sc of around 0.35. The \( \text{SI} \text{Pat}^{\text{GNAT}}-\text{Se} \text{Ac}^{\text{CTD}} \) interaction surface was distinct and larger than that of the Gcn5-H3 peptide complex (Fig. 3B,C,F). The interface between \( \text{SI} \text{Pat}^{\text{GNAT}} \) and \( \text{Se} \text{Ac}^{\text{CTD}} \) buried a total surface area of 2150 Å² where this was 48% polar and 52% non-polar, which is typical for recognition surfaces. The distribution of hydrophobicity in the interface is roughly the same as across the total surface area of either protein. This is consistent with a transient interface in which specificity is driven by charge-charge interactions with minimal hydrophobic contributions. The size and disposition of the residues in the binding interface was consistent with the hypothesis that Pat substrate specificity involves elements outside the simple PX4GK loop motif (12).

However, the PX4GK loop does play a structural role in positioning K609 into the active site cleft of \( \text{SI} \text{Pat}^{\text{GNAT}} \). The carbonyl oxygen of the preceding glycine residue hydrogen bonded to residue R64 of \( \text{SI} \text{Pat}^{\text{GNAT}} \) and facilitated a bend in the backbone loop conformation (Φ 97°, ψ 9°). Also located within the PX4GK loop, the carbonyl oxygens of \( \text{Se} \text{Ac}^{\text{CTD}} \) R606 and S607 hydrogen bond with the positively charged side chains of \( \text{SI} \text{Pat}^{\text{GNAT}} \) residues R79 and R64, respectively (Fig. 3E). Similar to the Tetrahymena Gcn5 H3 complex, hydrophobic interactions were also involved in positioning the target lysine (40). \( \text{SI} \text{Pat}^{\text{GNAT}} \) residue F66 packed against the methylene groups of \( \text{Se} \text{Ac}^{\text{CTD}} \) residue K609 (Fig. 3E).

In addition to the interactions with the PX4GK motif there were complementary ionic...
interactions between the protein domains, where a large negatively charged surface patch on \( \text{SI} \text{PatA}^{\text{GNAT}} \) interacted with a complementary positive patch on \( \text{SeAcs}^{\text{CTD}} \) (Figs. 3B, 3C). A prominent group of arginine residues in \( \text{SeAcs}^{\text{CTD}} \) (R612, R613, and R616) lay on the surface of a short \( \alpha \)-helix that followed the PXxGK motif (Figs. 3B), where these are conserved in Acs homologues from bacteria, archaea, and eukaryotes (Fig. 4A) (43). These arginines interacted with a negative patch on \( \text{SI} \text{PatA}^{\text{GNAT}} \) that included residues F66, E160, and E184 (Figs. 3C). These interactions most likely contribute to the specificity of the \( \text{SI} \text{PatA}^{\text{GNAT}} \) domain for its substrate.

\( \text{SI} \text{PatA}^{\text{GNAT}} \) homologs shown to acetylate the cognate Acs from the same organism exhibit amino acid sequence conservation at several of the residues noted above (Fig. 4B).

In vitro and in vivo evidence that amino acid charge reversals at the \( \text{SI} \text{PatA}^{\text{GNAT}}-\text{SeAcs}^{\text{CTD}} \) interface disrupt interactions. The protein:protein interactions observed in the crystal structure were tested with a bacterial two-hybrid assay in vivo by mutating charged surface residues (Fig. 5). Introduction of an opposing charge into the interacting surface of the \( \text{SI} \text{PatA}^{\text{D185R}} \) or \( \text{SeAcs}^{\text{CTD}} \) (e.g., \( \text{SeAcs}^{\text{R606E}}, \text{SeAcs}^{\text{R613D}} \)) significantly reduced interactions of those proteins with \( \text{SI} \text{PatA}^{\text{GNAT}} \) or \( \text{SI} \text{PatA}^{\text{GNAT}} \), respectively. Conversely, as a control, substitution of a residue near, but outside of the interaction interface (\( \text{SI} \text{PatA}^{\text{GNAT}} \text{E121R} \)) did not significantly affect the \( \text{SI} \text{PatA}^{\text{GNAT}}-\text{SeAcs}^{\text{CTD}} \) interaction, supporting the orientation of the domains in the X-ray structural model.

Importantly, the bacterial two-hybrid system results were reproduced both in vitro and in vivo when non-truncated forms of \( \text{SeAcs} \) and \( \text{SI} \text{PatA} \) were used, and when the \( \text{SI} \text{PatA} \) homologue from \( S. \text{enterica} \) (\( \text{SePat} \)) was used (Fig. 6, Fig. 7). Full-length \( \text{SeAcs}^{\text{WT}} \) variants \( \text{SeAcs}^{\text{R606E}} \) and \( \text{SeAcs}^{\text{R613D}} \) retained activity in vitro despite amino acid substitutions near the catalytic K609 [(\( \text{SeAcs}^{\text{WT}}, 8.3 \pm 0.5 \) \( \mu \)mol AMP min\(^{-1}\) mg\(^{-1}\); \( \text{SeAcs}^{\text{R606E}}, 5.1 \pm 0.2 \) \( \mu \)mol AMP min\(^{-1}\) mg\(^{-1}\); and \( \text{SeAcs}^{\text{R613D}}, 2.8 \pm 0.2 \) \( \mu \)mol AMP min\(^{-1}\) mg\(^{-1}\); mean ± S.D., n=9)], however \( \text{SI} \text{PatA}^{\text{WT}} \) and its \( S. \text{enterica} \) homologue \( \text{SePat}^{\text{WT}} \) acetylated these proteins less efficiently (Fig. 6). Amino acid substitutions near the active site lysine of AMP-forming CoA ligases have been shown to affect activity (11).

We also demonstrated that variant \( \text{SI} \text{PatA}^{\text{D185R}} \) E121R interacted with \( \text{SeAcs}^{\text{WT}} \) in vivo and inhibited growth of a \( S. \text{enterica} \Delta \text{pat} \Delta \text{cobB} \) strain, but not the growth of a \( S. \text{enterica} \Delta \text{pat} \text{cobB} \) strain. In contrast, variant \( \text{SI} \text{PatA}^{\text{D185R}} \) only slightly inhibited growth of the \( S. \text{enterica} \Delta \text{pat} \Delta \text{cobB} \) strain (Fig. 7). These data were consistent with the observation that the \( \text{SI} \text{PatA}^{\text{D185R}} \) variant exhibited significantly weaker interactions with \( \text{SeAcs}^{\text{CTD}} \) in the bacterial-two-hybrid assay (Fig. 5). When higher levels of variant \( \text{SI} \text{PatA}^{\text{GNAT}} \text{D185R} \) were present, growth of \( S. \text{enterica} \) on 10 mM acetate was inhibited (Fig. 7). These results were consistent with the idea that variant \( \text{SI} \text{PatA}^{\text{GNAT}} \text{D185R} \) and \( \text{SeAcs}^{\text{CTD}} \) interactions were weakened but not abolished.

**Discussion**

Mapping the interaction surface between the \( \text{SI} \text{PatA}^{\text{GNAT}} \) and the globular protein substrate \( \text{SeAcs}^{\text{CTD}} \) is a significant advance in our understanding of how protein lysine acetyltransferases recognize—with specificity—large globular protein targets. To generate an interaction model of a GNAT with a protein substrate, the crystal structure of \( \text{SI} \text{PatA}^{\text{GNAT}} \) was first solved. The structure of \( \text{SI} \text{PatA}^{\text{GNAT}} \) and \( \text{SePat}^{\text{WT}} \), the most closely related structure, differed significantly at the predicted substrate interface. \( \text{SePat}^{\text{WT}} \) was crystallized after limited proteolysis, and the absence of amino acids (specifically residues 42-52 that link helix \( \alpha 2 \) to strand \( \beta 2 \)) may distort the active site cleft. \( \text{SePat}^{\text{WT}} \) acetylates the DNA binding protein Alba, but it is unknown whether it can acetylate AMP-forming CoA ligases. Thus, we cannot rule out the possibility that the structural differences surrounding the \( \text{SePat}^{\text{WT}} \) substrate-binding cleft may represent distinct substrate specificity.

Interestingly, the factors that contribute to specificity observed for \( \text{SI} \text{PatA} \) are different from those reported for methylmalonyl-CoA synthetase of *Rhodopseudomonas palustris*. In that case, a loop spanning residues 447-450 in
the *R. palustris* methylmalonyl-CoA synthetase (*RpMatB*) was identified as containing elements critical for binding to, and subsequent acetylation by the *R. palustris* protein acetyltransferase (*RpPat*) enzyme (12). The equivalent loop in *SeAcs*<sub>CTD</sub> (residues 565-568) is not involved in the *SlPatA*<sub>GNAT</sub>*-SeAcs*<sub>CTD</sub> interface, which highlights the complexity and difficulty in predicting GNAT substrate specificity from first principles.

The structure of the acetylation complex revealed complementary electrostatic interactions between *SlPatA*<sub>GNAT</sub> and *SeAcs*<sub>CTD</sub>. The charged residues involved in the *SlPatA*<sub>GNAT</sub>*-SeAcs*<sub>CTD</sub> interface are conserved in only some species (Fig. 4A,B). The positively charged *SeAcs* residues R612, R613, and R616 are seen in homologues in all domains of life, yet further examination of the exceptions may reveal differences in PatA-Acs interactions among species. Likewise, few of the PatA residues involved in the PatA-Acs interaction are conserved. Considering that the GNAT family of acetyltransferases is noted for its lack of primary sequence conservation (7), predicting interacting residues of GNATs in the absence of structural data remains challenging.

The extensive interaction surface observed at the *SlPatA*<sub>GNAT</sub>*-SeAcs*<sub>CTD</sub> may be a key feature of PatA-Acs interactions. A large interaction surface would facilitate evolution of distinct constellations of interactions between each GNAT and its protein substrate(s). Continued structural analysis of GNAT-substrate complexes will reveal the range of interactions that occur between GNATs and their protein substrates.

The GNAT-substrate interface identified in the *SlPatA*<sub>GNAT</sub>*-SeAcs*<sub>CTD</sub> crystal structure is remarkably distinct from the Gcn5-H3 peptide complexes reported to date (Figs. 3C, 3E), and reveals structural roles of residues within and distant from the PX<sub>4</sub>GK motif found in substrates of Pat-type GNATs. The structure of the *SlPatA*<sub>GNAT</sub>*-SeAcs*<sub>CTD</sub> interaction will serve as a model to further identify, validate, and engineer (11) specific globular protein targets of GNAT protein acetyltransferases.

**Acknowledgments**

The authors thank Simon Dove for plasmids and strains. This work was supported by USPHS grant R01 GM062203 to J.C.E.-S. and by grant R01 GM083987 to I.R. Part of the results shown in this report were obtained at the Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. The University of Chicago Argonne, LLC operates ANL for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. K.C.R. was supported by a NIH training grant (T32-GM07215).

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the RCSB Protein Data Bank under the accession codes 4NXY (*SlPatA*<sub>GNAT</sub>) and 4U5Y (*SlPatA*<sub>GNAT</sub>*-SeAcs*<sub>CTD</sub> complex.)

**References**

1. Kouzarides, T. (2007) Chromatin modifications and their function. *Cell* 128, 693-705
2. Glozak, M. A., Sengupta, N., Zhang, X., and Seto, E. (2005) Acetylation and deacetylation of non-histone proteins. *Gene* 363, 15-23
3. Liang, W., Malhotra, A., and Deutscher, M. P. (2011) Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. *Mol. Cell* 44, 160-166
4. Crosby, H. A., Pelletier, D. A., Hurst, G. B., and Escalante-Semerena, J. C. (2012) System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J. Biol. Chem.* 287, 15590-15601
5. Hubbert, C., Guardiola, A., Shao, R., Kawaguchi, Y., Ito, A., Nixon, A., Yoshida, M., Wang, X. F., and Yao, T. P. (2002) HDAC6 is a microtubule-associated deacetylase. *Nature* 417, 455-458
6. Friedmann, D. R., and Marmorstein, R. (2013) Structure and mechanism of non-histone protein acetyltransferase enzymes. *FEBS J.* 280, 5570-5581
Crystal structure of a protein acetylation complex

7. Vetting, M. W., Carvalho, L. P. S. d., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L., and Blanchard, J. S. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* **433**, 212-226

8. Starai, V. J., Celic, I., Cole, R. N., Boeke, J. D., and Escalante-Semerena, J. C. (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science* **298**, 2390-2392

9. Thao, S., and Escalante-Semerena, J. C. (2011) Control of protein function by reversible N(epsilon)-lysine acetylation in bacteria. *Curr. Opin. Microbiol.* **14**, 200-204

10. Tucker, A. C., and Escalante-Semerena, J. C. (2013) Acetoacetyl-CoA synthetase activity is controlled by a protein acetyltransferase with unique domain organization in *Streptomyces lividans*. *Mol. Microbiol.* **87**, 152-167

11. Crosby, H. A., and Escalante-Semerena, J. C. (2014) The acetylation motif in AMP-forming acyl coenzyme A synthetases contains residues critical for acetylation and recognition by the protein acetyltransferase Pat of *Rhodopseudomonas palustris*. *J. Bacteriol.* **196**, 1496-1504

12. Crosby, H. A., Rank, K. C., Raymond, I., and Escalante-Semerena, J. C. (2012) Structural insights into the substrate specificity of the protein acetyltransferase *Rp*Pat: identification of a loop critical for recognition by *RpPat*. *J. Biol. Chem.* **287**, 41392-41404

13. Gulick, A. M., Starai, V. J., Horswill, A. R., Homick, K. M., and Escalante-Semerena, J. C. (2003) The 1.75Å crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A. *Biochemistry* **42**, 2866-2873

14. Bertani, G. (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293-300

15. Berkowitz, D., Hushon, J. M., Whitfield, H. J., Jr., Roth, J., and Ames, B. N. (1968) Procedure for identifying nonsense mutations. *J. Bacteriol.* **96**, 215-220

16. Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121-4130

17. Elion, E. A., Marina, P., and Yu, L. (2007) Constructing recombinant DNA molecules by PCR. in *Current protocols in molecular biology* (Ausubel, F. M., R. Brent, R. E., Kingston, D. D., Moore, J. G., Seidman, J. A., Smith, a., and Struhl, K. eds.), Greene Publishing Associates & Wiley Interscience, New York, N.Y. pp Unit 3.17.11-13.17.12

18. Rocco, C. J., Dennison, K. L., Klenchin, V. A., Raymond, I., and Escalante-Semerena, J. C. (2008) Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*. *Plasmid* **59**, 231-237

19. Blommel, P. G., and Fox, B. G. (2007) A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr. Purif.* **55**, 53-68

20. Griffith, K. L., and Wolf, R. E., Jr. (2002) Measuring beta-galactosidase activity in bacteria: cell growth, permeabilization, and enzyme assays in 96-well arrays. *Biochem. Biophys. Res. Commun.* **290**, 397-402

21. Zhang, X., and Bremer, H. (1995) Control of the *Escherichia coli* rnmB P1 promoter strength by ppGpp. *J. Biol. Chem.* **270**, 11181-11189

22. Garrity, J., Gardner, J. G., Hawse, W., Wolberger, C., and Escalante-Semerena, J. C. (2007) N-lysine propionylation controls the activity of propionyl-CoA synthetase. *J. Biol. Chem.* **282**, 30239-30245

23. Crosby, H. A., Heiniger, E. K., Harwood, C. S., and Escalante-Semerena, J. C. (2010) Reversible N(epsilon)-lysine acetylation regulates the activity of acyl-CoA synthetases involved in anaerobic benzoate catabolism in *Rhodopseudomonas palustris*. *Mol. Microbiol.* **76**, 874-888

24. Starai, V. J., and Escalante-Semerena, J. C. (2004) Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in *Salmonella enterica*. *J. Mol. Biol.* **340**, 1005-1012

25. Tucker, A. C., and Escalante-Semerena, J. C. (2010) Biologically active isoforms of CobB sirtuin deacetylase in *Salmonella enterica* and *Erwinia amylovora*. *J. Bacteriol.* **192**, 6200-6208

26. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685

27. Sasse, J. (1991) Detection of proteins, in *Current Protocols in Molecular Biology* (Ausubel, F. A., Brent, R., Kingston, R. E.,
Crystal structure of a protein acetylation complex

Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. eds.), Wiley Interscience, New York. pp 10.16.11-10.16.18

28. Rayment, I. (2002) Small-scale batch crystallization of proteins revisited: an underutilized way to grow large protein crystals. *Structure (Camb)* **10**, 147-151

29. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326

30. Sheldrick, G. M. (2008) A short history of SHELX. *Acta Crystallogr. A.* **64**, 112-122

31. Cowtan, K., and Main, P. (1998) Miscellaneous algorithms for density modification. *Acta Crystallographica D* **54**, 487-493

32. CCP4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760-763

33. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674

34. Skubák, P., Murshudov, G. N., and Pannu, N. S. (2004) Direct incorporation of experimental phase information in model refinement. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2196-2201

35. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132

36. Kozakov, D., Brenke, R., Comeau, S. R., and Vajda, S. (2006) PIPER: an FFT-based protein docking program with pairwise potentials. *Proteins* **65**, 392-406

37. Kozakov, D., Beglov, D., Bohnood, T., Mottarella, S. E., Xia, B., Hall, D. R., and Vajda, S. (2013) How good is automated protein docking? *Proteins*

38. Comeau, S. R., Gatechell, D. W., Vajda, S., and Camacho, C. J. (2004) ClusPro: a fully automated algorithm for protein-protein docking. *Nucleic Acids Res.* **32**, W96-99

39. Comeau, S. R., Gatechell, D. W., Vajda, S., and Camacho, C. J. (2004) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* **20**, 45-50

40. Rojas, J. R., Trivel, R. C., Zhou, J., Mo, Y., Li, X., Berger, S. L., Allis, C. D., and Marmorstein, R. (1999) Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. *Nature* **401**, 93-98

41. Clements, A., Poux, A. N., Lo, W. S., Pilulus, L., Berger, S. L., and Marmorstein, R. (2003) Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase. *Mol. Cell.* **12**, 461-473

42. Lawrence, M. C., and Colman, P. M. (1993) Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946-950

43. Starai, V. J., and Escalante-Semerena, J. C. (2004) Acetyl-coenzyme A synthetase (AMP forming). *Cell. Mol. Life Sci.* **61**, 2020-2030

44. Xu, H., Hegde, S. S., and Blanchard, J. S. (2011) Reversible acetylation and inactivation of *Mycobacterium tuberculosis* acetyl-CoA synthetase is dependent on cAMP. *Biochemistry* **50**, 5883-5892

45. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948

46. Castang, S., and Dove, S. L. (2010) High-order oligomerization is required for the function of the H-NS family member MvaT in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **78**, 916-931

Footnote

1Abbreviations: GNAT, Gcn5-related acetyltransferase; Pat, protein acetyltransferase; *SePat*, *Salmonella enterica* Pat; *SIPatA*, *Streptomyces lividans* PatA; *SIPatA*GNAT, GNAT domain of the *SIPatA* enzyme; *SeAcS*, *Salmonella enterica* (AMP-forming) acetyl-coenzyme A synthetase; *SeAcS*CTD, C-terminal domain of *SeAcS*; *RpPat*, *Rhodopseudomonas palustris* Pat enzyme; *SvPat*, *Sulfolobus sulfataricus* Pat enzyme; PEG, polyethylene glycol; NB, nutrient broth; LB, lysogeny broth; ONPG, *O*-nitrophenyl-β-galactoside; BMOE, *bis*(maleimido)ethane; BMH, *bis*(maleimido)hexane; BMB, 1,4-*bis*(maleimido)butane; DPDPB, 1,4-di-(3'-pyridylidithio)-propionamido); HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; MOPS, 3-(N-morpholino)propanesulfonic acid,
H$_e^+$, hexahistidine tag; rTEV, recombinant tobacco etch virus protease.

**FIGURE LEGENDS**

**Fig. 1. SLPatAGNAT domain and in vivo function.** (A) Regulation of SeAc$_{\text{WT}}$ activity by SePat$_{\text{WT}}$ SLPatAGNAT in *S. enterica*. (B) Domain arrangements of characterized wild-type protein acetyltransferases are listed for *Streptomyces lividans* (SLPatA), *Salmonella enterica* (SePat), and *Mycobacterium tuberculosis* (MtPatA). GNAT, Gen5-related N-acetyltransferase; cNMP, cyclic mononucleotide monophosphate-binding domain (C). Ribbon representation of SLPatAGNAT with its catalytic residue E123 shown in purple (PDB 4NXY). Ac-CoA was modeled into the SLPatAGNAT structure using the closely related MtPatA$_{\text{WT}}$ structure as a guide (44) (PDB # 4AVB). (D) Alignment of SLPatAGNAT (beige) with *Mycobacterium tuberculosis* PatA (teal) shows the location of SLPatAGNAT residues F126 and M168 that overlap with the modeled Ac-CoA. Homologous residues of MtPatA (F238 and M280) are shown and accommodate the bound Ac-CoA. SLPatAGNAT active site residue E123 is colored purple and putative active site MtPatA E235 is shown in dark teal. (E) Growth of *S. enterica* Δpat cobB$^+$ (open symbols) and Δpat ΔcobB strains (filled symbols) producing SLPatAGNAT (squares), SLPatAGNAT active site variant E123Q (triangles), or an empty vector (circles) on acetate. (F) ClusPro interaction model of SLPatAGNAT (beige) with SeAc$_{\text{N-terminal domain in green, C-terminal domain in blue}}$ with active site residue E123 (red) and SeAc$_{\text{WT target residue K609 (blue)}}$ shown as spheres.

**Fig. 2. Cross-linking of variants SLPatAGNAT S73C and SeAc$_{\text{CTD}}$ H567C with a disulfide bond results in a productive complex.** (A) ClusPro model of the SLPatAGNAT-SeAc$_{\text{interface}}$. SLPatAGNAT active site residue E123 (red) and SeAc$_{\text{target residue K609 (blue)}}$ are shown as spheres. Cross-linked residues are highlighted. (B) Schematic of SeAc$_{\text{CTD}}$ (light blue) and SLPatAGNAT$_{\text{beige}}$ with relative location of the cross-linked residues (spheres). Residues E123 of SLPatAGNAT and K609 of SeAc$_{\text{CTD}}$ are red and blue spheres, respectively. (C) Chemical structures of 2.0 - 19.0Å cross link where atoms and bonds in red are directly inserted between the cysteine residues (AT-2, aldrithiol-2; BMOE, bis[maleimido]ethane; BMB, 1,4-[bis[maleimido]butane; BMH, bis[maleimido]hexane; DPDPB, 1,4-di-[3'-[2'-pyridyldithio]-propionamido]. (D) SLPatAGNAT was incubated with SeAc$_{\text{CTD}}$ at molar ratios of 1:3, 1:1, 10:1 (SLPatAGNAT$^{-}$SeAc$_{\text{CTD}}$) in the presence of [1-14C]-acetyl-CoA to visualize acetyltransfer to SeAc$_{\text{CTD}}$ (phosphor images labeled “[14C] Acetylation”). Samples were quenched after 60 min, separated by SDS-PAGE and stained with Coomassie Blue to visualize proteins (labeled “SDS-PAGE”). Full length SLPatA was incubated with SeAc$_{\text{CTD}}$ at a molar ratio of 1:3 (SLPatA$^{-}$:SeAc$_{\text{CTD}}$) for reference. Images of Coomassie Blue-stained gels and phosphor images were cropped to bands corresponding to the SeAc$_{\text{CTD}}$. Acetylation was quantified relative to the signal obtained with SLPatA plus SeAc$_{\text{CTD}}$, and is reported as the mean (n=3). S.D. was ≤18% of the mean value. (E) Transfer of the acetyl moiety from [1-14C]-acetyl-CoA to the SeAc$_{\text{CTD}}$ was tested (phosphor images) for each of the SLPatAGNAT$^{-}$SeAc$_{\text{CTD}}$ complexes. Images of Coomassie Blue-stained gels (labeled “SDS-PAGE”) and phosphor images (labeled “[14C] Acetylation”) were cropped to bands corresponding to the SLPatAGNAT$^{-}$SeAc$_{\text{CTD}}$ heterodimers.

**Fig. 3. Charged surface residues mediate interactions between SLPatAGNAT and SeAc$_{\text{CTD}}$.** (A) Ribbon representation of the SLPatAGNAT S73C- SeAc$_{\text{CTD}}$ H567C crystal structure with residues E123 of SLPatAGNAT and K609 of SeAc$_{\text{CTD}}$ shown as sticks in red and blue, respectively. (B) and (C) the electrostatic potential distribution at the interface for SeAc$_{\text{CTD}}$ and SLPatAGNAT colored red for negatively charged, white for neutral, and blue for positively charged areas. (D) Interaction of the SLPatAGNAT active site cleft with SeAc$_{\text{CTD}}$ (R606-R616 shown in blue). The region shown in the black box is enlarged (E) to show specific interaction of SeAc$_{\text{CTD}}$ R606-R616 with SLPatAGNAT. (F) The electrostatic potential distribution at the interface for Gcn5 interaction with substrate H3 peptide (PDB # 1QSN) colored as described above. The H3 peptide is colored green with the target Lys14 shown in blue.
Fig. 4. SlPatAGNAT and SeAcsCTD residues at the interaction interface are conserved. (A) Alignment of sequences in and around the AcsCTD PX4GK motif (black box) from S. enterica (SeAcs, accession # NP_463140), Saccharomyces cerevisiae (Acs2p, accession # NP_013254), Halobacterium salinarum (HsAcs, accession # WP_0109027), and S. lividans (SlAcs, accession # EF668454). Blue shaded boxes indicate conserved positively-charged residues. “*“ indicates a fully conserved residue; “:“ indicates residues with high similarity; “.” indicates residues with low similarity. (B) Alignment of GNAT domain from homologs of SlPatA (accession # EF666247) from S. enterica (SePat, accession # XNP_461586), R. palustris (RpPat, accession # WP_0494576), and Mycobacterium tuberculosis (MtPat, accession # WP_003906490). Notation is described as above. Red and green shaded boxes indicate negatively charged and hydrophobic residues, respectively, observed at the SlPatAGNAT- SeAcsCTD interaction interface. Sequence alignment generated in ClustalW2 (45).

Fig. 5. Reversing charges at the interaction surface disrupts interactions between SlPatAGNAT and SeAcsCTD. (A) Effect of different Zif-SlPatAGNAT variants and empty vector on transcription in vivo from promoter PlacZif 1-61 with ω-SeAcsCTD variants or empty vector. (B) Data from the reciprocal bait-prey experiment (Zif-SeAcsCTD and ω-SlPatAGNAT). “ω only” refers to empty prey plasmid expressing only the ω subunit of the RNA Polymerase alpha subunit. “Zif only” refers to empty bait plasmid expressing only the Zinc finger protein. * indicates p<0.0001; † interactions were detected in bait-prey reciprocal experiments.

Fig. 6. Full length variants SeAcsR606E and SeAcsR613D are poorly acetylated by protein acetyltransferases. (A) SeAcsWT, SeAcsR606E, and SeAcsR613D were incubated with SlPatA or SePat in a 3:1 molar ratio (SeAcs:SlPatA or SePat) the presence of [1-14C]-Ac-CoA. Samples were quenched at 0, 15, 30, 60, and 90 min and separated by SDS-PAGE and stained with Coomassie Blue to visualize proteins (labeled “SDS-PAGE”). Acetylation was visualized by phosphor imaging (labeled “[14C] Acetylation”). (B) Reaction controls lacking SePat were incubated for 90 minutes and imaged as described above. Gels and phosphor images were cropped to the SeAcs bands and labeled as described in panel A. (C) Phosphor signal associated with each band in panels A and B was quantified as described in the experimental procedures. (D) SeAcs, SeAcsR606E, and SeAcsR613D were incubated with SlPatA (white bars) or SePat (gray bars) at the ratio described above in the presence or absence of acetyl-CoA. After 90 min, SeAcs activity was measured in an NADH-consumption assay. All data points are mean ± standard deviation (n=6).

Fig. 7. SlPatAGNAT D185R interacts poorly with SeAcs in vivo. S. enterica Δpat cobB+ (open shapes) and S. enterica Δpat ΔcobB (filled shapes) strains producing wild-type SlPatAGNAT (pGNATWT, circles), and variants SlPatAGNAT E121R (pGNATE121R, squares) and SlPatAGNAT D185R (pGNATD185R, triangles) in the absence of inducer during growth on NCE minimal medium supplemented with acetate (10 mM). (B) Growth of S. enterica Δpat cobB+ (squares) and S. enterica Δpat ΔcobB (circles) strains producing variant SlPatAGNAT D185R (pGNATD185R) in the presence of inducer (no inducer [white], 5 µM inducer [light gray], 10 µM inducer [dark gray], 25 µM inducer [black]) during growth on NCE minimal medium supplemented with acetate (10 mM). Growth experiments were performed at 37°C using a microtiter plate and a microtiter plate reader (Bio-Tek Instruments). All data points represent mean value. All standard deviations <0.015 absorbance units (n=4).
### Table 1. Strains used in this study

| Strain       | Genotype                               | Source                  |
|--------------|----------------------------------------|-------------------------|
| **E. coli strains** |                                        |                         |
| JE9314       | C41 (λDE3) pka12::kan<sup>+</sup>      | Laboratory collection   |
| JE14224      | KDZif1ΔZ                               |                         |
| JE18685      | KDZif1ΔZ / pAC-WSVG-Zif-AP pBR<sup>+</sup> kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18687      | KDZif1ΔZ / pAC-WSVG-Zif-AP pSlPatA46 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18688      | KDZif1ΔZ / pAC-WSVG-Zif-AP pSlPatA47 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18689      | KDZif1ΔZ / pAC-WSVG-Zif-AP pSlPatA48 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18690      | KDZif1ΔZ / pAC-WSVG-Zif-AP pACS53 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18691      | KDZif1ΔZ / pAC-WSVG-Zif-AP pACS54 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18692      | KDZif1ΔZ / pAC-WSVG-Zif-AP pACS55 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18693      | KDZif1ΔZ / pAC-WSVG-Zif-MvaT pBR<sup>+</sup>-MvaT kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18695      | KDZif1ΔZ / pSlPatA43 pBR<sup>+</sup>-GP kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18696      | KDZif1ΔZ / pSlPatA44 pACS53 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18697      | KDZif1ΔZ / pSlPatA44 pACS54 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18698      | KDZif1ΔZ / pSlPatA44 pACS55 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18699      | KDZif1ΔZ / pSlPatA44 pBR<sup>+</sup>-GP kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18700      | KDZif1ΔZ / pSlPatA44 pACS53 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18701      | KDZif1ΔZ / pSlPatA44 pACS54 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18702      | KDZif1ΔZ / pSlPatA44 pACS55 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18703      | KDZif1ΔZ / pSlPatA45 pBR<sup>+</sup>-GP kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18704      | KDZif1ΔZ / pSlPatA45 pACS53 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18705      | KDZif1ΔZ / pSlPatA45 pACS54 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18706      | KDZif1ΔZ / pSlPatA45 pACS55 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18707      | KDZif1ΔZ / pACS50 pBR<sup>+</sup>-GP kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18708      | KDZif1ΔZ / pACS50 pSlPatA46 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18709      | KDZif1ΔZ / pACS50 pSlPatA47 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18710      | KDZif1ΔZ / pACS50 pSlPatA48 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18711      | KDZif1ΔZ / pBR<sup>+</sup>-GP pACS51 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18712      | KDZif1ΔZ / pBR<sup>+</sup>-GP pACS52 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18713      | KDZif1ΔZ / pSlPatA46 pACS51 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18714      | KDZif1ΔZ / pSlPatA46 pACS52 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18715      | KDZif1ΔZ / pSlPatA47 pACS51 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18716      | KDZif1ΔZ / pSlPatA47 pACS52 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18717      | KDZif1ΔZ / pSlPatA48 pACS51 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| JE18718      | KDZif1ΔZ / pSlPatA48 pACS52 kan<sup>+</sup> bla<sup>+</sup> tet<sup>+</sup> |
| **S. enterica strains** |                                        |                         |
| JE16236      | metE205 ara-9 Δpat3                       |                         |
| JE16237      | metE205 ara-9 Δpat2 ΔcobB1375              |                         |
| **Derivatives of JE16236** |                                        |                         |
| JE16146      | metE205 ara-9 Δpat3 / pBAD30 bl<sup>+</sup>  |                         |
| JE17191      | metE205 ara-9 Δpat3 / pSlPatA11 bl<sup>+</sup> |                         |
| JE17922      | metE205 ara-9 Δpat3 / pSlPatA21 bl<sup>+</sup> |                         |
| JE18755      | metE205 ara-9 Δpat3 / pSlPatA51 bl<sup>+</sup> |                         |
| JE18756      | metE205 ara-9 Δpat3 / pSlPatA52 bl<sup>+</sup> |                         |
| **Derivatives of JE16237** |                                        |                         |
| JE16419      | metE205 ara-9 Δpat2 ΔcobB1375 / pBAD30 bl<sup>+</sup> |                         |
| Strain   | Genotype                        | Plasmid          |
|----------|---------------------------------|------------------|
| JE17193  | metE205 ara-9 Δpat2 ΔcobB1375 / pSI/PatA11 bla<sup>+</sup> |                  |
| JE17894  | metE205 ara-9 Δpat2 ΔcobB1375 / pSI/PatA21 bla<sup>+</sup> |                  |
| JE18757  | metE205 ara-9 Δpat2 ΔcobB1375 / pSI/PatA51 bla<sup>+</sup> |                  |
| JE18758  | metE205 ara-9 Δpat2 ΔcobB1375 / pSI/PatA52 bla<sup>+</sup> |                  |

*Unless otherwise indicated, all strains and plasmids were constructed during the course of this work*
Table 2. Plasmids used in this study

| Plasmid       | Genotype                                                                 | Referencea |
|---------------|---------------------------------------------------------------------------|------------|
| pTEV5         | N-terminal, rTEV-cleavable His6-tag overexpression vector, bla<sup>+</sup>  | (18)       |
| pBAD30        | P<sub>araBAD</sub> expression vector, bla<sup>+</sup>                      |            |
| pKLD66        | N-terminal, rTEV-cleavable His<sub>6</sub>-MBP-tag overexpression vector,  |            |
|               | bla<sup>+</sup>                                                          |            |
| pBR<sub>ω</sub>GP | N-terminal RNAP ω fusion vector for bacterial 2-hybrid assay, bla<sup>+</sup> | (46)       |
| pACTR-V-Zif-AP | N-terminal VSV-G-Zif268 fusion vector for bacterial 2-hybrid assay, tet<sup>+</sup> | (46)       |

**pTEV5 derivatives**

| Plasmid | Genotype                                                                 |
|---------|---------------------------------------------------------------------------|
| pACS38  | Encodes the wild-type SeAcs CTD (SeAcs<sup>CTD</sup>)                     |
| pACS42  | Encodes variant SeAcs<sup>CTD</sup> R606E                                |
| pACS44  | Encodes variant SeAcs<sup>CTD</sup> R613D                                |
| pACS56  | Encodes variant SeAcs<sup>CTD</sup> A538C                                |
| pACS57  | Encodes variant SeAcs<sup>CTD</sup> H567C                                |
| pACS58  | Encodes variant SeAcs<sup>CTD</sup> D600C                                |
| pACS60  | Encodes variant SeAcs<sup>CTD</sup> R606E                                |
| pACS61  | Encodes variant SeAcs<sup>CTD</sup> R613D                                |
| pSlPatA27| Encodes the wild-type GNAT domain of SlPatA                              |
| pSlPatA33| Encodes variant SlPatA<sup>GNAT</sup> E121R                             |
| pSlPatA35| Encodes variant SlPatA<sup>GNAT</sup> D185R                             |

**pKLD66 Derivatives**

| Plasmid | Genotype                                                                 |
|---------|---------------------------------------------------------------------------|
| pSlPatA14 | Encodes the wild-type GNAT domain of SlPatA (SlPatA<sup>GNAT</sup>)   |
| pSlPatA53 | Encodes variant SlPatA<sup>GNAT</sup> S73C                               |
| pSlPatA54 | Encodes variant SlPatA<sup>GNAT</sup> A110C                              |
| pSlPatA56 | Encodes variant SlPatA<sup>GNAT</sup> A164C                              |

**pBAD30 Derivatives**

| Plasmid | Genotype                                                                 |
|---------|---------------------------------------------------------------------------|
| pSlPatA11 | Encodes the wild-type GNAT of SlPatA (SlPatA<sup>GNAT</sup>)             |
| pSlPatA21 | Encodes variant SlPatA<sup>GNAT</sup> E123Q                              |
| pSlPatA51 | Encodes variant SlPatA<sup>GNAT</sup> E121R                              |
| pSlPatA52 | Encodes variant SlPatA<sup>GNAT</sup> D185R                              |

**pBR<sub>ω</sub>GP derivatives**

| Plasmid | Genotype                                                                 |
|---------|---------------------------------------------------------------------------|
| pSlPatA46 | Encodes the wild-type SlPatA GNAT domain (SlPatA<sup>GNAT</sup>)        |
| pSlPatA47 | Encodes variant SlPatA<sup>GNAT</sup> E121R                              |
| pSlPatA48 | Encodes variant SlPatA<sup>GNAT</sup> D185R                              |
| pACS53  | Encodes the wild-type CTD of SeAcs                                       |
| pACS54  | Encodes the SeAcs<sup>CTD</sup> CTD variant                              |
| pACS55  | Encodes the SeAcs<sup>CTD</sup> CTD variant                              |

**pACTR-V-Zif-AP derivatives**

| Plasmid | Genotype                                                                 |
|---------|---------------------------------------------------------------------------|
| pSlPatA43 | Encodes the wild-type GNAT domain of SlPatA (SlPatA<sup>GNAT</sup>)   |
| pSlPatA44 | Encodes variant SlPatA<sup>GNAT</sup> E121R                              |
| pSlPatA45 | Encodes variant SlPatA<sup>GNAT</sup> D185R                              |
| pACS50  | Encodes the wild-type CTD of SeAcs (SeAcs<sup>CTD</sup>)                 |
| pACS51  | Encodes variant SeAcs<sup>CTD</sup> R606E                                |
| pACS52  | Encodes variant SeAcs<sup>CTD</sup> R613D                                |

*Unless noted, all plasmids were constructed during the course of this study.*
### Table 3. Oligonucleotides used in this study

| Primer Name       | Primer Sequence<sup>a,b,c</sup> |
|-------------------|----------------------------------|
| SeAcs CTD pTEV5 F | GTAGCTAGCGACGTGTAAACGTCGCCATTCGC|
| SeAcs CTD pTEV5 R | ATGGAATTTTCTATGAGCCGATCGCATGAGC|
| SeAcs<sup>R606E</sup> F | CTCACCTGCCAAAAACGGAATCCGGCAGAATTATGCGC |
| SeAcs<sup>R606E</sup> R | GCCGATAAATTTCGCAGCAGATTTTCGTCAGATGAGC |
| SeAcs<sup>R613D</sup> F | CCGGCAAATTATCGCAGCAGATTTTTCGCAAATC |
| SeAcs<sup>R613D</sup> R | CGGCAATTTTCCGCAAATATCGCATATTTTGCAGC |
| SeAcs<sup>A538C</sup> F | GCTTCCGGCAGATTTCCGAGTCGACTCGATC |
| SeAcs<sup>A538C</sup> R | GTTTGGCAGATGAGAAGCGCAGATTCGCAAATC |
| SeAcs<sup>H567C</sup> F | GTGACGCTCAACGGCGAGGTTCGCGCGATC |
| SeAcs<sup>H567C</sup> R | GGCTCCTCGCCGAGTTGAGCGC |
| SeAcs<sup>R606E</sup> F | CTCACCTGCCAAAAACGGAATCCGGCAGAATTATGCGC |
| SeAcs<sup>R606E</sup> R | GCCGATAAATTTCGCAGCAGATTTTCGTCAGATGAGC |
| SeAcs<sup>R613D</sup> F | CCGGCAAATTATCGCAGCAGATTTTTCGCAAATC |
| SeAcs<sup>R613D</sup> R | CGGCAATTTTCCGCAAATATCGCATATTTTGCAGC |
| GNAT pTEV5 F | GTAGCTAGCAATGTCGTACGCGAGCCGTAC |
| GNAT pTEV5 R | ATGGAATTTTCTATGAGCCGATCGCATGAGC |
| GNAT pKLD66 F | GTAGGTACCATGTCGTACGCGAGCCGTACTCTTG |
| GNAT pKLD66 R | ACTGTCGACTCAGATCGAATCCGAGCGC |
| GNAT<sup>S73C</sup> F | CGTACCTCGCTGTCGCGGAGGACGTGAC |
| GNAT<sup>S73C</sup> R | GTTGATGCCGAATGAGACGTGAC |
| GNAT<sup>A110C</sup> F | GACCGATCGCAGGCGCGCAACAC |
| GNAT<sup>A110C</sup> R | GTGTTCCGCCGACCGC |
| GNAT<sup>A116C</sup> F | CTACGCAGCGGATCGGTTGCGCGCGAGC |
| GNAT<sup>A116C</sup> R | GCGGGTTTGGCCTCGCAGCAG |
| GNAT pBAD30 F | ACCGAAATTCAGGGAACGCGGATGTCGTACGCGAGGGTCGCTAC |
| GNAT pBAD30 R | GTAGGTACCTCGAGATCGAATCCGAGCGC |
| GNAT<sup>E112Q</sup> F | GCGGCGAGGCGGCGGAGGCGCGATCGCCTG |
| GNAT<sup>E112Q</sup> R | GAGGAGGCGCAGGTGCGC |
| GNAT<sup>E112R</sup> F | CACCGGACCGCGGCAAGTGCGCAGGTCGCTCT |
| GNAT<sup>E112R</sup> R | GAAAGCGACCTCGGACACG |
| GNAT<sup>D135K</sup> F | GAGAAGCGGAAGCTCGAGGTTGCGGCTGCGG |
| GNAT<sup>D135K</sup> R | CAGGCGGACACCCCGGACCTGCGGCTGCGG |
| GNAT TwoHyb F | GCTGCGGCCGCAATGTCGTACGCGAGCCGTAC |
| GNAT TwoHyb R | ATGGAATTTTCTATGAGCGAATCCGAGCG |
| SeAcs CTD TwoHyb F | CAGGCGGCAGCGCCGAGCTTAAACGTCGTC |
| SeAcs CTD TwoHyb R | ATGGAATTTTCTATGAGCGAATCCGAGCG |

<sup>a</sup> Restriction nuclease sites are underlined
<sup>b</sup> Bold typeface indicates the codon introduced during site-directed-mutagenesis.
<sup>c</sup> Italic typeface indicates a 13 nt region found 5’ of *S. enterica* *pat* that includes the *pat* RBS
Table 4. Crystallographic Statistics.

| Data Collection | SlPatAGNAT-SeAcs<sup>CTD</sup> | SlPatAGNAT<sup>GNAT</sup> |
|----------------|------------------|------------------|
| Space group    | C222₁            | P2₁2₁2₁          |
| Cell dimensions|                  |                  |
| a, b, c (Å)    | 62.5, 66.5, 138.5| 40.1, 56.7, 74.0 |
| α, β, γ (°)    | 90, 90, 90       | 90, 90, 90       |
| Wavelength (Å) | 0.9789           | 0.9793           |
| Resolution (Å)<sup>a</sup> | 50-1.9 (1.95-1.92) | 1.45             |
| R<sub>merge</sub><sup>a</sup> | 6.7 (33.1)       | 3.8 (21.2)       |
| <I>/<σ><sup>a</sup> | 30.5 (4.9)       | 35.4 (10.1)      |
| Completeness (%)<sup>a</sup> | 99.7 (98.4)     | 99.8 (98.7)      |
| Redundance<sup>a</sup> | 10.6 (7.3)      | 8.7 (6.6)        |
| Beamline       | 19-BM            | 19-ID            |

**Refinement**

| Resolution (Å)<sup>a</sup> | 22-1.9 (2.03-1.92) | 25-1.45 (1.48-1.45) |
| No. reflections<sup>b</sup> | 22055 (1035)       | 29002 (1540)        |
| R<sub>work</sub>/R<sub>free</sub><sup>c</sup> | 0.19/0.24          | 0.18/0.22           |
| No. atoms       |                  |                  |
| Protein         | 2201             | 1412             |
| Water           | 184              | 190              |
| Ligand          | 10               | 12               |
| Average B-factors (Å<sup>2</sup>) | 18.8            | 15.4             |
| R.m.s. deviations|                  |                  |
| Bond lengths (Å)| 0.008            | 0.025            |
| Bond angles (°) | 1.04             | 2.4              |
| Ramachandran (%)|                  |                  |
| Most favored    | 98.9             | 98.25            |
| Allowed         | 1.1              | 1.75             |
| Disallowed      | 0                | 0                |
| PDB Code        | 4U5Y             | 4NXYY            |

<sup>a</sup>Data in parentheses represent the highest resolution shell.

<sup>b</sup>Data in parentheses represent the number of reflections used for the calculation of R<sub>free</sub>.

<sup>c</sup>R<sub>factor</sub> = \( \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|} \)

Where R<sub>work</sub> refers to the R<sub>factor</sub> for the data utilized in the refinement and R<sub>free</sub> refers to the R<sub>factor</sub> for 5% of the data that were excluded from the refinement.
Crystal structure of a protein acetylation complex

Figure 1

A

B

C

D

E

F
Crystal structure of a protein acetylation complex

Figure 2

A ClusPro 2.0 Interaction Model

B

Acetyltransferase:

| Molar Ratio: (SeAc-CoA/SpaA^NAD^+) | SDS-PAGE | [%] Acetylation |
|-------------------------------------|----------|-----------------|
| - 3.1:1 3.1:1 1.1:1 1:10            |          |                 |

Relative Acetylation: - 100 5 15 95

C

Crosslinkers:

- 2.05 Å Crosslinker (AT-2)
- 8.09 Å Crosslinker (BMOE)
- 10.9 Å Crosslinker (BMBI)
- 13.0 Å Crosslinker (BMH)
- 19.0 Å Crosslinker (DPDCB)

D

E

Downloaded from http://www.jbc.org/ by guest on July 8, 2020
Crystal structure of a protein acetylation complex

Figure 3
Crystal structure of a protein acetylation complex

Figure 5

A

| $\omega$-SeAc$\text{CTD}$ | Zif-S/PatA$\text{GNAT}$ | Bait: Zif-S/PatA$\text{GNAT}$ | Prey: $\omega$-SeAc$\text{CTD}$ |
|---------------------------|-------------------------|-------------------------------|-------------------------------|
| wildtype                  | Zif only                | ~                              | ~                            |
|                           | wildtype                | ~                              | ~                            |
|                           | E121R                   | *                             | *                            |
|                           | D185R                   | *                             | *                            |
| R606E                     | Zif only                |                               |                              |
|                           | wildtype                |                               |                              |
| R613D                     | Zif only                |                               |                              |
|                           | wildtype                |                               |                              |
| $\omega$ only             | Zif only                |                               |                              |
|                           | wildtype                |                               |                              |

Units $\beta$-glucosidase activity

B

| Zif-SeAc$\text{CTD}$ | $\omega$-S/PatA$\text{GNAT}$ | Bait: Zif-SeAc$\text{CTD}$ | Prey: $\omega$-S/PatA$\text{GNAT}$ |
|-----------------------|-------------------------------|-------------------------------|-------------------------------|
| wildtype              | $\omega$ only                | ~                             | ~                            |
|                       | wildtype                     | ~                             | ~                            |
|                       | E121R                        | ~                             | ~                            |
|                       | D185R                        | ~                             | *                            |
| R606E                 | $\omega$ only                | ~                             | ~                            |
|                       | wildtype                     | ~                             |                              |
| R613D                 | $\omega$ only                | ~                             | ~                            |
|                       | wildtype                     | ~                             |                              |
| Zif only              | $\omega$ only                | ~                             | ~                            |
|                       | wildtype                     | ~                             |                              |

Units $\beta$-glucosidase activity
Crystal structure of a protein acetylation complex

Figure 6

A

B

C

D

[Image of figures and graphs related to protein acetylation complex]
Crystal structure of a protein acetylation complex

Figure 7

[Graph showing absorption at 630 nm over time for different conditions involving cobB and pGNAT variants, with annotations indicating strains and conditions such as "cobB* / pGNAT^{D185R} (25 \mu M inducer)".]
Insights Into the Specificity of Lysine Acetyltransferases
Alex C. Tucker, Keenan C. Taylor, Katherine C. Rank, Ivan Rayment and Jorge C. Escalante-Semerena

J. Biol. Chem. published online November 7, 2014

Access the most updated version of this article at doi: 10.1074/jbc.M114.613901

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts