REVERSIBLE LYSINE ACETYLATION IN

STREPTOMYCES LIVIDANS AND *SALMONELLA ENTERICA*

by

ALEX CLAYTON TUCKER

(Under the Direction of JORGE C. ESCALANTE-SEMERENA)

ABSTRACT

Lysine acetylation is a post-translational modification conserved in all domains of life. Lysine acetylation occurs lysine side chains and affects functions of proteins involved in diverse cellular pathways including transcription, chromatin maintenance, cell structure, and central metabolism. Acetyltransferase enzymes carry out acetylation of specific lysine residues in target proteins. The protein acetyltransferase Pat enzymes from the γ-proteobacterium *Salmonella enterica* and the α-proteobacterium *Rhodopseudomonas palustris* contain 2 domains: an *N*-terminal domain of unknown function and a *C*-terminal GNAT domain. *S. enterica* Pat and *R. palustris* Pat are the best-characterized bacterial protein acetyltransferases. Pat homologs, along with protein deacetylases, regulate the activity of AMP-forming CoA ligases (ACLs) and structurally unrelated targets *in vivo*. The first part of this work describes the identification and characterization of two active isoforms of CobB sirtuin-type (NAD^+-dependent) deacetylase expressed as part of the *S. enterica* reversible lysine acetylation system.

Reversible lysine acetylation is poorly understood in Gram-positive bacteria. The actinomycete *S. lividans* encodes a Pat homolog, *Sl*PatA, with reversed domain order (i.e. *N*-terminal GNAT domain and *C*-terminal domain of unknown function). Experiments here demonstrate that *Sl*PatA regulates acetoacetyl-CoA synthetase (AacS), a member of the ACL family, in *S. lividans*. Acetylation of the active site lysine inactivates AacS. A proteomic-based study of acetylated proteins in *S. lividans* reveals that

*Sl*PatA regulates the acetylation of four additional ACL enzymes in *S. lividans*. These data suggest that a major role of *Sl*PatA is to regulate ACL enzymes.

These studies also describe how *Sl*PatA recognizes a protein substrate. *Sl*PatA does not efficiently acetylate the *S. lividans* ACL enzyme acetyl-CoA synthetase (Acs) *in vitro*. By swapping regions of *S. lividans* Acs protein with a known protein substrate of *Sl*PatA, we generated a *S. lividans* Acs variant that is efficiently acetylated *in vitro*. This work defined elements in protein substrates recognized by *Sl*PatA. Structure guided studies of the *Sl*PatA GNAT domain, reveal important electro static interactions between the *Sl*PatA GNAT domain and a protein substrate. A crosslinked complex composed of the GNAT and protein substrate was generated that likely resembles a biologically acetylation complex.

INDEX WORDS: post-translational modification, acetylation, bacteria, *Salmonella*, *Streptomyces*

REVERSIBLE LYSINE ACETYLATION IN *STREPTOMYCES LIVIDANS* AND *SALMONELLA ENTERICA*

by

ALEX CLAYTON TUCKER

B.S., Auburn University, 2007

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2013

© 2013

Alex Clayton Tucker

All Rights Reserved

REVERSIBLE LYSINE ACETYLATION IN

STREPTOMYCES LIVIDANS AND *SALMONELLA ENTERICA*

by

ALEX CLAYTON TUCKER

Major Professor:	Jorge C. Escalante-Semerena
Committee:	Diana Downs Lawrence Shimkets Janet Westpheling

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2013

ACKNOWLEDGEMENTS

The faculty members at the University of Wisconsin and the University of Georgia have been helpful with feedback on my research projects. I especially like to thank Dr. Jorge C. Escalante-Semerena for creating a stimulating scientific environment, encouraging my scientific development, and supporting my work in his lab. I would like to thank Dr. Diana Downs, Dr. Karen Wassarman, Dr. Margaret McFall-Ngai, and Dr. Ned Ruby for serving on my thesis committee at UW. I would also like to thank Dr. Lawrence Shimkets and Dr. Janet Westpheling for serving on my thesis committee at UGA. Dr. Michael Thomas (UW) and Dr. Westpheling (UGA) passed on much knowledge about handling *Streptomyces*. I thank Dr. Nicole Perna at UW for her gift of *Erwinia amylovora* DNA. I would also like to thank my former scientific advisor Dr. Sang-Jin Suh for his patience and encouragement in my earliest days of scientific research.

My colleagues, friends, and family have all played a role in this work. Members of the Escalante and Downs labs provided much scientific feedback during the course of this work. I'd especially like to thank Dr. Heidi Crosby, Dr. Chi Ho Chan, Dr. Paola Mera, Norbert Tavares, and Lauren Palmer for help with experimental design, data analysis, and manuscripts. I would also like to thank John Barkei and Dr. Matt McMahon for valuable advice about working with *Streptomyces.* I'd also like to thank all of my friends that have supported me in many ways related to this work. Lastly, I'd like to thank my family for their love and support.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iv

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

2 BIOLOGICALLY ACTIVE ISOFORMS OF COBB SIRTUIN DEACETYLASE IN *SALMONELLA ENTERICA* AND *ERWINIA AMYLOVORA* 23

3 ACETOACETYL-COA SYNTHETASE ACTIVITY IS CONTROLLED BY A PROTEIN ACETYLTRANSFERASE WITH UNIQUE DOMAIN ORGANIZATION IN *STREPTOMYCES LIVIDANS* 52

4 IDENTIFICATION OF ACETYLATION TARGETS OF THE PROTEIN ACETYLTRANSFERASE PATA IN *STREPTOMYCES LIVIDANS* 98

5 IDENTIFICATION OF DETERMINANTS WITHIN THE WILD-TYPE *C*-TERMINAL DOMAIN OF THE *STREPTOMYCES LIVIDANS* ACETYL-COA SYNTHETASE ENZYME THAT BLOCK ACETYLATION OF ITS ACTIVE SITE LYSINE 125

6 INTERACTIONS OF A PROTEIN ACETYLTRANSFERASE DOMAIN WITH AN ACID:COA LIGASE (AMP-FORMING) 160

7 CONCLUSIONS AND FUTURE DIRECTIONS 203

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

OVERVIEW

Lysine acetylation is a conserved post-translational modification. Lysine acetylation occurs on the side chain of lysine residues, and is known to affect function of proteins involved in diverse cellular pathways. Acetylomics, the proteomic-based study of acetylated proteins, has revealed a large number of acetylation targets in diverse organisms from bacteria to humans. Acetyltransferases enzymes catalyze lysine acetylation. Gcn5-related *N*-acetyltransferases (GNATs) are the only protein acetyltransferases indentified in all domains of life. While some bacterial GNATs exist as single-domain acetyltransferases, some bacteria encode multi-domain acetyltransferases and activity of these enzymes is allosterically regulated in some cases. Bacterial protein acetyltransferases acetylate members of the AMP-forming CoA ligase protein family, and some acetyltransferases modify structurally unrelated targets. GNATs exhibit substrate specificity by interacting with regions of the protein that are close to and distant from the target lysine. Further biochemical and structural studies will indentify conserved and distinct features in GNATs and their substrates that define reversible lysine acetylation in bacterial species.

LYSINE ACETYLATION IS A CONSERVED POST-TRANSLATIONAL MODIFICATION

Discovery, distribution, and effect of lysine acetylation. Lysine acetylation occurs on the ε-amino group of lysine residues, increasing the bulk of the lysine side chain and decreasing the relative reactivity of the ε-amine by converting it to an amide (3). Lysine acetylation was first documented in 1963 as a modification on the *N*-terminal tails of eukaryotic histones (49). Histones are major protein component of the compact DNA structure in eukaryotes known as chromatin, and acetylation of these proteins is associated with decreased DNA binding and increase in transcription (27, 58). Lysine

acetylation has since been observed in bacteria, archaea, and eukaryotes (53), and affects cellular pathways including chromatin maintenance (32), transcriptional regulation (16), primary metabolism (9), and cell structure (23). Consequently, lysine acetylation has been suggested to rival phosphorylation in the diversity of protein targets (31). Acetylation was thought to occur only through the action of enzymes known as histone acetyltransferase (HATs) until recently. Recent work in *Escherichia coli* suggests that global, low-level lysine acetylation is also mediated by the reactive, high-energy metabolite acetyl-phosphate, and this acetylation is globally regulated by growth phase and metabolism (66). Direct acetylation by acetyl-Coenzyme A has also been suggested as a mechanism of non-enzymatic acetylation in eukaryotes (18). These findings suggest complex control of protein lysine acetylation that occurs in a diversity of biological pathways and processes.

Identification of acetylated proteins in biological systems. Characterization of total protein lysine acetylation (termed the "acetylome") in bacterial and eukaryotic systems has been accelerated by the development of proteomics-based approaches to enrich for and detect lysine acetylation (29). Monoclonal anti-acetyllysine antibodies and mass spectrometry are the basis for most current proteomic approaches. The first acetylomic study was based on the 2-dimensional separation of proteins from mouse tissue and detection of acetyllysine residues by immunoblotting (26). Subsequent eukaryotic acetylomic studies employed immunoprecipitation to enrich for acetyllysine peptides in tryptic digests of total cellular proteins from human cells and mouse mitochondria (30, 51, 76). In all cases, mass spectrometry has been used to identify the acetylated peptides and proteins. These total protein studies identified acetylation targets involved in diverse cellular pathways including metabolism, cell structure, cellular trafficking, and chromatin maintenance, confirming a broad role of acetylation in eukaryotic biology.

Bacteria acetylomes are characterized from the model organisms *Escherichia coli* (73, 75) and *Bacillus subtilis* (28) to the economically important plant pathogen *Erwinia amylovora* (69) and the thermophile *Geobacillus kaustophilus* (34). Studies in these and other bacteria have revealed many new targets for acetylation, and have raised important questions about the physiological relevance of the acetylation detected by mass spectrometry. Notably, two studies of acetylated proteins in *E. coli* identified

85 and 91 acetylation targets (73, 75), and only 6 acetylation sites overlapped between the two studies (61). An acetylome of the closely related Gram-negative enterobacterium *Salmonella enterica* included 191 acetylated proteins. Though several of these targets were validated *in vitro* (64), subsequent reports failed to reproduce those observations (9). Comparisons among the wild-type *Rhodopseudomonas palustris*.acetylome compared to the acetylomes of *R. palustris* strains lacking one or more acetyltransferases revealed acetylation targets that were validated *in vitro* and *in vivo,* and activity of each of the targets was affected by acetylation (10). Thus, a combination of genetics and *in vitro* validation enhance understanding of bacterial acetylomes.

Classes of histone acetyltransferase enzymes. HATs mediate acetylation of specific lysine residues within the target protein (2, 9, 33). Despite the nomenclature, many non-histone substrates have been identified for these enzymes (described above). HATs catalyze the transfer of the acetyl moiety from acetyl-CoA (Ac-CoA) to the ε-amine of lysine side chains. At least five major families of HATs have been identified based on sequence divergence. HAT families are named after their founding member(s) including Gcn5-related *N*-acetyltransferases (GNAT, named after yeast yGcn5p), histone acetyltransferase 1 (HAT1, named after yeast yHAT1p), MYST (named after human MOZ, yeast Ybf2/Sas3, yeast Sas2, and human Tip60), p300/CBP (named for human hp300 and hCBP), and Rtt109 (yeast Rtt109) (74). The acetyltransfer mechanism used by GNATs and MYST HATs involves a catalytic glutamate that acts a general base (50, 71). GNATs transfer the acetyl group from Ac-CoA directly to the substrate lysine. Members of the MYST family employ either a catalytic mechanism similar to GNATs or a ping-pong mechanism involving and acetylated enzyme intermediate (72). The p300/CBP HATs do not use a catalytic base for acetyltransfer. Instead p300 /CBP HATs are though to employ a mechanistically distinct “hit-and-run” type acetyltransfer reaction (38). The acetyltransfer mechanisms of the HAT1 and Rtt109 HATs are not well characterized. However, HAT1 is thought to use a general base for catalysis (68).

GNAT protein acetyltransferases in bacteria. GNATs are the only lysine acetyltransferase family identified in bacteria, eukaryotes, and archaea (61), and members of GNAT family can acetylate

non-protein substrates (63). The model bacterium *E. coli* encodes up to 22 GNATs, many of which target primary amines other than lysine side chains (21, 66), including the *N*-termini of proteins (11, 25), aminoglycoside antibiotics (13), glutamate (41), and transfer RNAs (24). As reference points for the diversity of GNATs, the Gram-positive actinomycete *Streptomyces lividans* encodes 72 putative GNAT enzymes and the purple non-sulfur α-proteobacterium *Rhodopseudomonas palustris* encodes 26 putative GNATs (Pfam00583) (9, 62), most of which have no known function. Here, we focus on the bacterial and archaeal GNATs that catalyze protein lysine acetylation.

Protein lysine acetyltransferases have been identified in bacteria and archaea, and this family of related enzymes exhibits diverse domain architecture (Fig. 1.1). The *Salmonella enterica* protein acetyltransferase *Se*Pat was the first protein acetyltransferase identified in bacteria (56). *Se*Pat is a two-domain protein acetyltransferase. The *N*-terminal domain of *Se*Pat (~700 a.a.) is homologous to the NDP-forming CoA ligase family of enzymes. The catalytic GNAT domain (~200 a.a.) is found at the *C*-terminus of *Se*Pat. Despite its homology to a metabolic enzyme, the large *N*-terminal domain of *Se*Pat has no known function. *Se*Pat homologs have been identified in *E. coli* (*Ec*Pka) (37), and *Rhodopseudomonas palustris* (*Rp*Pat) (8). *S. lividans* encodes an *Se*Pat homolog (*Sl*PatA) in which the domain order is reversed (*N*-terminal GNAT domain, *C*-terminal large domain of unknown function) (62). *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* each encode a two-domain protein acetyltransferase (*Mt*PatA and *Ms*PatA, respectively) in which the GNAT domain is fused to the *C*-terminus of a cyclic nucleotide-binding domain (discussed below). Single-domain GNAT protein lysine acetyltransferases have been identified in *Bacillus subtilis* (*Bs*AcuA), *Rhodopseudomonas palustris* (*Rp*KatA), and the archaeon *Sulfolobus solfataricus* (*Ss*Pat). The range of domain architectures and organization in the bacterial and archaeal protein lysine acetyltransferase reveals that the lysine acetylation is regulated in diverse ways within these organisms. Understanding the regulatory inputs of these enzymes will allow us to draw connections from cellular signals to lysine acetylation and the effects acetylation has on target proteins.

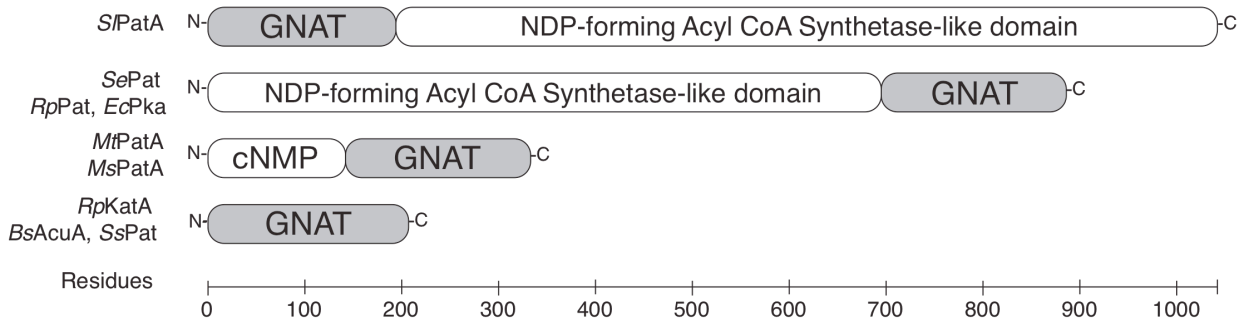


Figure 1.1. Protein acetyltransferases in bacteria and archaea. Domain arrangements of characterized protein acetyltransferases are listed for *Streptomyces lividans* (*Sl*PatA), *Rhodopseudomonas palustris* (*Rp*Pat, *Rp*KatA), *Salmonella enterica* (*Se*PatA), *Escherichia coli* (*Ec*Pka), *Mycobacterium tuberculosis* (*Mt*PatA), *Mycobacterium smegmatis* (*Ms*PatA), *Bacillus subtilis* (*Bs*AcuA), and *Sulfolobus solfataricus* (*Ss*Pat). GNAT, Gcn5-related *N*-acetyltransferase; cNMP, cyclic mononucleotide monophosphate-binding domain.

REVERSIBLE LYSINE ACETYLATION IN BACTERIA AND ARCHAEA

***Regulation of acetyl-CoA synthetase in* S. enterica.** *Se*Pat was first identified for its ability to acetylate and inactivate the enzyme acetyl-CoA synthetase (Acs) involved in acetate utilization in *S. enterica* (56) (Fig. 1.2). Acs is a member of the AMP-forming CoA ligase (ACL) family of enzymes (55). ACL enzymes convert organic acids to their CoA thioesters via an adenylated intermediate (19). A conserved lysine in ACL enzymes is required for adenylation of organic acids. Acetylation of this lysine prevents adenylylation of the acid substrate and blocks enzyme activity. Thus, *Se*Pat acetylation of the Acs active site lysine blocks Acs conversion of acetate to acetyl-AMP (54), the first half-reaction of Acs. Acetylation of Acs homologs has subsequently been reported in eukaryotes (20, 57), demonstrating that this is a conserved regulatory mechanism.

Acetylation of Acs can be reversed by sirtuin CobB, a NAD^+-dependent protein deacetylase in *S. enterica*. CobB consumes NAD^+ as a co-substrate in the deacetylation reaction, releasing deacetylated protein substrate, nicotinamide, and *O*-acetyl-ADP-ribose (15). *S. enterica* produces two isoforms of CobB from the same open reading frame using two independent start codons (see Chapter 2). The isoforms differ in length by 37 amino acids, resulting in a long isoform of CobB ($CobB_L$, 273 aa) and a short isoform of CobB ($CobB_S$, 236 aa). Both CobB isoforms are active, and $CobB_S$ is produced at ~10-fold greater levels than $CobB_L$ *in vivo*. The dual, in-frame start codons are conserved in the enterobacteria, and the role physiological relevance of the two isoforms has not been determined.

***Regulation of Acetyl-CoA synthetase in* B. subtilis.** A single-domain GNAT protein acetyltransferase AcuA regulates the activity of the acetyl-CoA synthetase AcsA in the soil-dwelling, Gram-positive bacterium *B. subtilis*. AcsA is deacetylated by both sirtuin SrtN or Zn^{2+}-dependent deacetylase AcuC.

***Regulation of ACL enzymes in* R. palustris.** Lysine acetylation regulates anaerobic catabolism of aromatic organic acids and other diverse fatty acids in the purple non-sulfur photosynthetic α-proteobacterium *R. palustris*. *R. palustris* encodes an *Se*Pat homolog, *Rp*Pat, and a single-domain GNAT *Rp*KatA. *R. palustris* encodes a sirtuin deacetylase, SrtN, and a Zn^{2+}-dependent lysine deacetylase

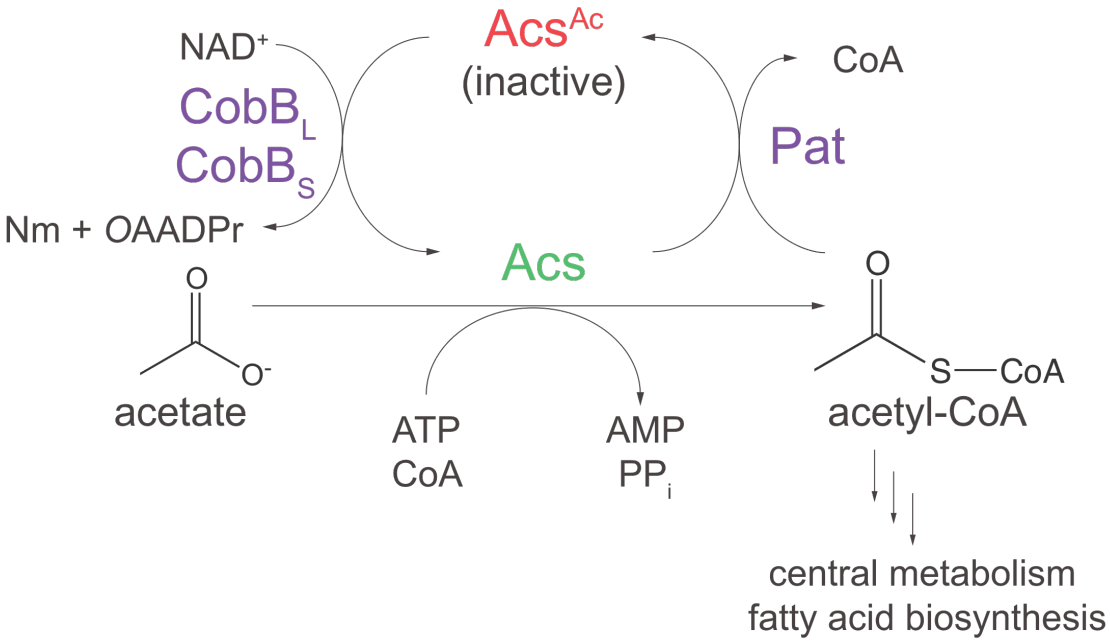


Figure 1.2. Lysine acetylation regulates Acs activity in *S. enterica*. Acetyl-CoA synthetase converts acetate to acetyl-CoA. Pat acetylates and inactivates Acs. Acs is deacetylated and reactivated by sirtuin isoforms $CobB_L$ or $CobB_S$ in a NAD^+-dependent manner, releasing nicotinamide (Nm) and *O*-acetyl-ADP-ribose (*O*AADPr).

(LdaA). *Rp*Pat and *Rp*KatA together with SrtN and LdaA regulate members of the ACL family of enzymes in *R. palustris*. *Rp*Pat and *Rp*KatA acetylate the conserved catalytic lysine of ACL enzymes that activate short-, medium-, long-, and branched-chain fatty acids in addition to aromatic organic acids (8, 9). *Rp*Pat and *Rp*KatA display different substrate specificities for the ACL enzymes of *R. palustris*, and not all ACL enzymes in *R. palustris* are regulated by reversible lysine acetylation. Together, these finding suggest that GNATs exhibit substrate specificity among members of the ACL family (9).

Regulation of ACL enzymes in* Streptomyces *spp. Metabolic regulation of actinomyces is of interest because of the diverse natural products they produce (6, 7, 44, 52). The Gram-positive actinomycete *S. lividans* encodes a *Se*Pat homolog, *Sl*PatA, in which the GNAT domain is found at the *C*-terminal domain of the large domain of unknown function. Despite domain organization, *Sl*PatA also acetylates ACL enzymes, namely *S. lividans* acetoacetyl-CoA synthetase and *S. enterica* Acs (62). *S. lividans* encodes two sirtuins CobB1 and CobB2, and a Zn^{+}-dependent deacetylase. Work in the closely related actinomycete *Streptomyces coelicolor* demonstrated that Acs is acetylated in this organism and that CobB1 deacetylates Acs *in vitro* (45). However, the acetyltransferase responsible for acetylation of *S. coelicolor* Acs is unknown.

Reversible lysine acetylation in mycobacterial species. *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* encode 2-domain protein lysine acetyltransferases (*Mt*PatA and *Ms*PatA, respectively) in which the GNAT domain is fused to the *C*-terminus of a cyclic AMP (cAMP) binding domain (47, 70). Binding of cAMP to the *Ms*PatA enhances acetyltransferase activity >2-fold (47). Universal stress protein (USP) was identified as a substrate for *Ms*PatA. USP is not conserved in Mycobacterial species that encode *Ms*PatA homologues, leading Xu *et al.* to test the ACL enzyme Acs as a substrate for *Ms*PatA (70). *M. smegmatis* Acs could not be purified, but the authors were able to show that *Ms*PatA acetylates the *M. tuberculosis* Acs in a cAMP-dependent manner. *Mt*PatA was inactive when purified in *vitro*. Eight additional *M. tuberculosis* ACL enzymes were identified as substrates of *Ms*PatA. Acetylation of all 8 ACL enzymes and the Acs enzyme are reversible by *M. tuberculosis* sirtuin, the single deacetylase encoded by this organism (48).

***Regulation of non-ACL enzymes of* E. coli.** Non-ACL enzymes from *E. coli* are modified by protein acetyltransferases. The *E. coli* homolog of *Se*Pat, *Ec*Pka, acetylates Ribonuclease R (RNase R) (36, 37). Lysine acetylation destabilizes RNase R by rendering it susceptible to proteolytic degradation. Acetylation of RNase R is not reversible by deacetylation with the *E. coli* sirtuin CobB.

Incubation of a *E. coli* proteome microarray with *Se*PatA (92% identical to *Ec*Pka) and radiolabeled [1-^{14}C]acetyl-CoA revealed that *Se*Pat acetylates a bacterial transcription factor (59). *Se*Pat acetylated *E. coli* RcsB, a transcription factor involved in multiple cellular pathways including flagellum synthesis, capsule synthesis and cell division (22, 39). Acetylation of RcsB decreases its affinity for DNA *in vitro*, and RcsB can be deacetylated by the *E. coli* sirtuin CobB. The role of RcsB acetylation *in vivo* has not been established.

***Acetylation of DNA-binding proteins in* S. solfataricus.** The archaeon *S. solfataricus* encodes a single-domain GNAT *Ss*Pat that acetylates the DNA-binding protein Alba (1, 40). Alba binds DNA *in vitro*, and binding inhibits transcription (1). Alba residue Lys16 is thought to make direct contacts with DNA (65), and acetylation of K16 by Pat decreases the affinity of Alba for DNA by ~3-fold (40). *S. solfataricus* sirtuin Sir2 deacetylation of K16 and increases the affinity of Alba for DNA (1).

STRUCTURAL, THERMODYNAMIC, AND MECHANISTIC INSIGHTS INTO BACTERIAL AND ARCHAEAL PROTEIN ACETYLTRANSFERASES

***GNAT protein acetyltransferase mechanism*.** All bacterial and archaeal protein acetyltranferases characterized have a conserved predicted active site glutamate, characteristic of the GNAT family (described above). The catalytic glutamate acts as a general base, facilitating a water-mediated proton abstraction from the target lysine (2, 63). The amine group then carries out a nucleophilic attack on the carbonyl carbon of the acetyl moiety of CoA. The GNAT then releases products CoA and acetylated target protein. Single-domain GNAT protein aceytltransferases have been identified in bacteria (e.g. *Rp*KatA, *Ss*Pat, *Bs*AcuA). The GNAT proteins aceyltransferase activity of *S. enterica* Pat homologs and mycobacterial Pat enzymes is thought to be regulated by secondary accessory domains.

***Regulation of acetyltransferase activity in* Se*Pat and its ortholog* Sl*PatA*.** *Se*Pat homologs contain a domain of unknown function fused to the GNAT domain that may be critical for acetyltransferase activity. Though this domain is homologous to the NDP-forming family of CoA ligases, *Se*Pat does not have this enzymatic activity (56). Insights into the role of this large domain were gained with *Se*Pat was found to bind two molecules of Ac-CoA in isolation. The GNAT domain binds a single Ac-CoA molecule, and a second Ac-CoA molecule is thought to bind outside of the GNAT domain (60). The *Se*Pat GNAT domain was inactive in isolation. These data suggest that the large domain of *Se*Pat and its homologs *Rp*Pat, *Ec*Pka, and *Sl*PatA may be involved in regulating activity of the GNAT domain.

The large domain in the *S. lividans* homolog of *Se*Pat, *Sl*PatA, may play a different role in this enzyme. The domain organization of *Sl*PatA is reversed with respect to *Se*Pat, *Ec*Pka, and *Rp*Pat (i.e. the GNAT domain is *N*-terminal to the large domain of unknown function) (62). The *Sl*PatA large domain contains a proline-rich region that includes a degenerate collagen-like G-P-S sequence. Collagenase cleaves *Sl*PatA but not an active *Sl*PatA variant that lacks the collagen-like motif (*Sl*PatAΔPro, unpublished results, A. C. Tucker and J.C. Escalante-Semerena). *Sl*PatAΔPro also appears to have increased catalytic activity compared to *Sl*PatA. The location and features of the *Sl*PatA enzyme suggest that regulation of this enzyme may be more complex than other *Se*Pat homologs.

Cyclic AMP binding regulates activity of mycobacterial protein acetyltransferases. Mycobacterial protein acetyltranferases are allosterically regulated by the binding of cAMP to a cAMP-binding domain at the *N*-terminus of the GNAT domain (described above). Structural studies of *Mt*PatA demonstrate that this enzyme can exist in an active or autoinhibited state (35). When autoinhibited, a lid blocks entrance of the protein substrate in the GNAT domain active site cleft. In the active state, the cAMP-binding domain is rotated 40° relative to the GNAT domain, the lid swings away from the GNAT domain, exposing the active site cleft. Binding of cAMP stabilizes the active form of *Mt*PatA, enabling lysine acetylation of protein targets.

INTERACTIONS OF GNAT WITH PROTEIN SUBSTRATES

Recognition of a* N-*terminal histone tail by Gcn5. The structure of Gcn5, the founding GNAT, bound to CoA and a representative histone H3 11 amino acid peptide revealed key features of GNAT interactions with protein substrates (50). Binding of CoA to Gcn5 induces structural changes in Gcn5 that are thought to facilitate interactions with the protein substrate. The conserved, catalytic Glu122 in the *Tetrahymena* Gcn5 (tGCN5) active site was poised to abstract a proton from the target lysine (K14) via an ordered water molecule that bridges Glu122 and K14 in the histone H3 peptide (Fig 1.3). Hydrophobic residues in tGcn5 interact with the methlyene groups in the K14 side chain as the ε-amine group. tGcn5 interacts with the backbone and side chain residues of the peptide substrate. Gcn5 is thought to recognize a G-**K**-X-P recognition sequence. The glycine residue preceding K14 interacts with Gcn5 through van der Waals forces is thought to be important for introducing flexibility into the substrate protein near K14. The conserved Pro in the Gcn5 acetylation motif contacts CoA and makes hydrophobic interactions with Gcn5.

Substrate specificity of* Rp*Pat for ACL enzymes. Characterization of *Rp*Pat substrate specificity for ACL enzymes demonstrated that GNATs make contacts with substrate proteins at regions far from the target lysine. A crystal structure of *Rp*Pat has not been determined, but *Rp*Pat substrates identified to-date share a relatively conserved acetylation sequence P-X_4-G-**K** (9). Conservation of the Gly residue directly preceding the target lysine suggests that this may be a common feature of GNAT substrates since it is also seen in the Gcn5 acetylation sequence. However, this acetylation motif is not sufficient to render an ACL enzyme a target of acetylation. Crosby *et al.* demonstrated that *R. palustris* ACL enzyme methylmalonyl-CoA synthetase (MatB) is not acetylated by *Rp*Pat, but introduction of residues from a known *Rp*PatA substrate, pimelate CoA synthetase (PimA), make MatB a substrate of *Rp*Pat (10). Notably, residues swapped from PimA into MatB are located ~20Å from the target lysine, demonstrating that *Rp*Pat must recognize additional structural elements in protein substrates in addition to the target lysine.

***Substrate specificity of* Sl*PatA*.** *S. coelicolor* Acs is acetylated *in vivo*, suggesting that Acs from the closely related actinomycete *S. lividans* is also subject to acetylation *in vivo*. *S. lividans* protein

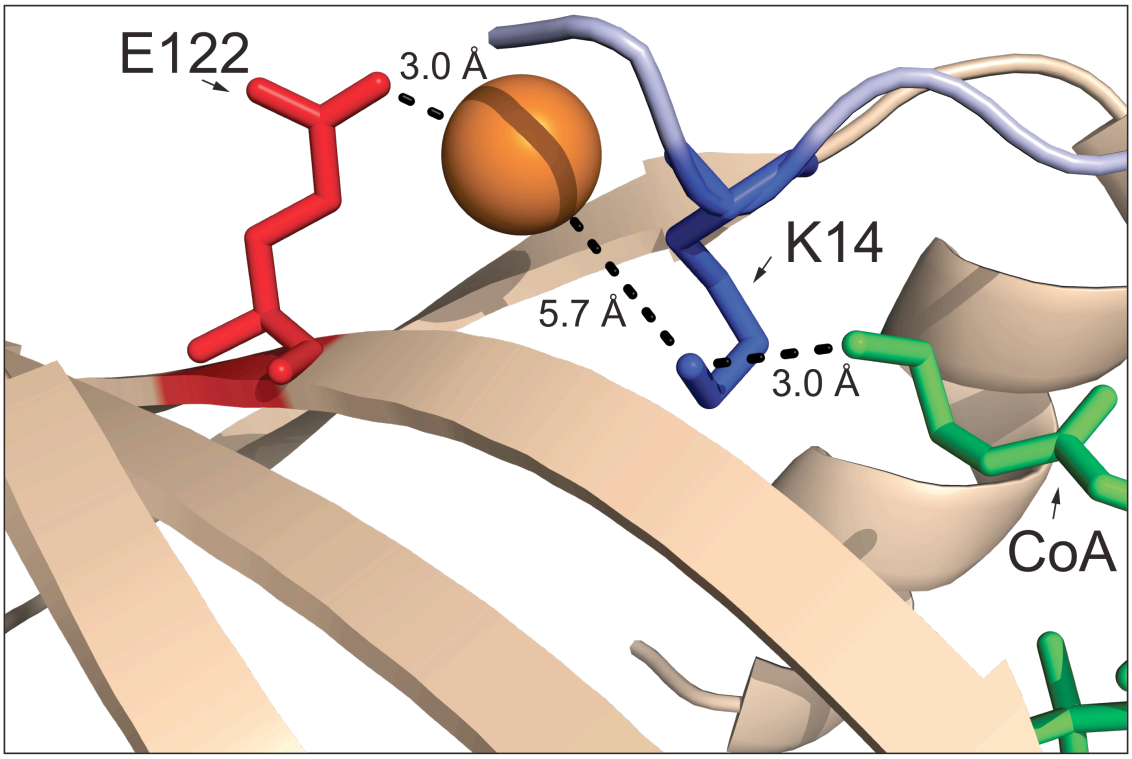


Figure 1.3. Gcn5 residue E122 utilizes a water-mediated bridge to catalyze proton abstraction from the target lysine. *Tetrahymena* Gcn5 (PDB: 1QSN, shown in beige) residue E122 (red) acts as a general base to abstract a proton from the target lysine (K14, dark blue) via a bridging water molecule (orange sphere). Image was generated in PyMol (14). CoA is shown in green. Peptide substrate is shown in light blue. Dashed lines represent distance between atoms.

acetyltransferase *Sl*PatA represents the first Pat homolog characterized that does not efficiently acetylate its cognate Acs enzyme *in vitro* (62). *Sl*PatA acetylates the ACL enzyme acetoacetyl-CoA synthetase, suggesting that *Sl*PatA may not be the enzyme responsible for Acs acetylation in *S. lividans*. Alternatively, acetylation of *S. lividans* Acs by *Sl*PatA may require additional factors that are not required by *Se*Pat and *Rp*Pat for acetylation of Acs orthologs from those organisms.

RATIONALE FOR REVERSIBLE LYSINE ACETYLATION AS A REGULATORY MECHANISM

Ac-CoA, cAMP, and NAD$^+$ link reversible lysine acetylation to cellular metabolites. Bacterial protein acetyltransferases use Ac-CoA as a substate linking protein acetylation to Ac-CoA levels. The two-domain Pat homologs *Se*Pat, *Ec*Pka, *Rp*Pat, and *Sl*PatA have large regulatory domains that may bind Ac-CoA to allosterically regulate acetyltransferase activity (60). Ac-CoA is a key metabolite linked to many pathways including glycolysis, the tricarboxylic acid cycle, and fatty acid biosynthesis/degradation. Levels of Ac-CoA are regulated by the metabolic activity of the cell. For example, Acs produces Ac-CoA, and Pat may sense acetyl-CoA levels and acetylate Acs as a negative feedback control mechanism. Additionally, *Rp*Pat acetylates many ACL enzymes whose products are likely to be metabolized to produce acetyl-CoA (9). Acetylation of these *R. palustris* ACL enzymes would control CoA homeostasis by preventing depletion of CoA or build-up of Ac-CoA. Chan *et al.* demonstrated another important aspect of regulating the activity of ACL enzymes. In the absence of Pat, Acs depletes cellular ATP pools, leading to growth arrest of *S. enteric* during growth on acetate (5).

Mycobacterial GNATs *Mt*PatA and *Ms*PatA are allosterically activated by the secondary signaling molecule cAMP (35, 46-48, 70). Mycobacteria encode a large number of adenylate cyclases relative to other bacteria (42), suggesting that cAMP is an important signaling molecule in mycobacteria. Pathogenic *Mycobacterium* sp. use cAMP to reroute host signaling during infection (43), indicating that this molecule is important for pathogenesis. High cAMP levels are observed in non-pathogenic bacteria

(12). Thus, cAMP signal is involved in both basic biology and pathogenesis of *Mycobacterium* spp. The secondary messenger cAMP is consequently integrated into the lysine acetylation system of mycobacteria by allosterically regulating *Mt*PatA activity. Mycobacteria may use cAMP as one signal to adjust flux through ACL enzymes (48), the major targets of *Ms*PatA in *M. smegmatis*.

NAD^+ is required as a co-substrate for sirtuin deacetylases, linking activity of these enzymes to the availability of NAD^+ in the cell. NAD^+ is the oxidized form of NADH, an important electron carrier in oxidative phosphorylation. NAD^+ levels show more variation than NADH levels when bacteria are grown under conditions with differing nutrients and oxygen availability (67), suggesting that intracellular NAD^+ levels are dynamic and important measures of carbon and energy status. For example, aerobic respiratory conditions that result in high NAD^+ levels could trigger deacetylation of ACL enzymes and other sirtuin targets.

Transcriptional regulation of* E. coli *protein acetyltransferase and deacetylases. Gene expression also controls the acetyltransferase and deacetylase activities in a bacteria cell. *E. coli pat* gene expression is activated by cAMP, a molecule that signals low glucose availability (4). cAMP binds master transcriptional regulator CRP to regulate genes involved in utilization of alternative carbon sources (17). *E. coli cobB* is expressed constitutively and is not activated by cAMP, but *E. coli acs* gene expression activate by cAMP *in vivo* (4). Thus activation *pat* and *acs* expression in *E. coli* may represent a regulatory mimic of the direct cAMP activation of the GNAT activity seen in mycobacterial protein acetyltransferases *Mt*PatA and *Ms*PatA, described above.

SUMMARY

Lysine acetylation is a dynamic and efficient mechanism to alter protein function. As proteomic studies reveal new targets of acetylation, the potential roles of lysine acetylation are expanded. Validation of these acetylomes is the next step in understanding the physiological role of this modification in biological systems. Protein acetyltransferases acetylate specific residues of target proteins using Ac-CoA as the acetyl donor, and the physiological role of many of these acetylation events is understood. Protein

acetyltransferases and their cognate protein deacetylases have been identified in bacteria, archaea, and eukaryotes. Because acetyltransferases and sirtuin-type deacetylases use Ac-CoA and NAD^+, respectively, as substrates, reversible lysine acetylation is tied to cellular metabolism and energy status. GNATs are the single acetyltransferase family found in all domains of life. Current studies suggest that a major role of bacterial GNAT protein acetyltransferases is to regulate ACL family enzymes. Biochemical and structural studies of GNATs and their substrates demonstrate that GNATs have interactions with multiple amino acids and regions of amino acids in their protein substrates. Importantly, these studies suggest that while there may be some absolutely conserved features of acetylation targets, other features may be dependent on the context of the target lysine and the identity of the cognate acetyltransferase.

Identifying acetyltransferases, deacetylases, and their targets will provide understanding of the logic behind the proteins and pathways targeted by lysine acetylation. Structural studies of GNATs and these substrates will define how GNATs can recognize diverse protein substrates and distinguish between closely related ones.

DISSERTATION OUTLINE

This work focuses on the machinery and targets that compose the reversible lysine acetylation in the actinomycete *S. lividans* and the γ-Proteobacterium *S. enterica*. I identify and characterize a new isoform of the *S. enterica* sirtuin deacetylase CobB. Previous work in *S. enterica*, *E. coli*, *R. palustris*, and mycobacteria has described 2-domain protein acetyltransferases and their substrates in these organisms. I demonstrate that *Sl*PatA, a related GNAT protein acetyltransferase from *S. lividans*, regulates previously unidentified ACL targets in this organism. I also probe the substrate specificity of *Sl*PatA by examining the differences between strong and poor substrates of *Sl*PatA.

Chapter 2 describes the discovery and characterization of dual start codons that give rise to two isoforms of the CobB sirtuin deacetylase in *S. enterica*. We demonstrate that both CobB isoforms are active, and we identify potential transcriptional start sites for each isoform in *S. enterica*. The in-frame

start codons are conserved within *Enterobacteriaceae*, suggesting that this isoform may have a conserved role in this family of bacteria.

In Chapter 3, I describe members of the reversible lysine acetylation system in *S. lividans*. We identified a 2-domain protein acetyltransferase, *Sl*PatA, with a unique domain organization compared to known protein acetyltransferases. *Sl*PatA is shown to acetylate acetoacetyl-CoA synthetase (AacS), an ACL enzyme.

In Chapter 4, I used a proteomic-based approach to expand the known acetylation targets under control of *Sl*PatA in *S. lividans*. Four new targets of protein acetylation were identified, and all are members of the ACL family of enzymes. These data suggested that regulation of ACL enzymes may the major role for *Sl*PatA, as was suggested for *Rp*Pat.

In Chapter 5, I demonstrated that *Sl*PatA does not efficiently acetylate *S. lividans* Acs *in vitro*, even though it efficiently acetylates the *S. enterica* Acs ortholog. By introducing amino acid sequences from *Se*Acs into *Sl*Acs, I generated a *Sl*Acs-*Se*Acs chimeric protein that was efficiently acetylated by *Sl*PatA *in vitro* and in the heterologous host *S. enterica*. Regions *N*-terminal and *C*-terminal to the target lysine were identified as important for *Sl*PatA recognition of *Se*Acs *in vitro*. The work suggested that regulation of *Sl*Acs by *Sl*PatA in *S. lividans* might be more complex than is observed in *S. enterica* and *R. palustris*.

Chapter 6 is a structural and mechanistic study of how GNAT domains recognize their protein substrates. I identified charged residues on the surface of the *Sl*PatA GNAT domain that are critical for interaction with the protein substrate *in vitro* and *in vivo*. I also showed that charged residues on the surface of the protein substrate are important for recognition by Pat homolog. I proposed the first model GNAT interactions with a protein substrate and validated this model by crosslinking the GNAT and protein substrate and confirming interaction.

In chapter 7, I provide conclusions, closing remarks, and future directions for work related to these studies.

REFERENCES

1. **Bell, S. D., C. H. Botting, B. N. Wardleworth, S. P. Jackson, and M. F. White.** 2002. The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. *Science* **296:**148-151.

2. **Berndsen, C. E., and J. M. Denu.** 2008. Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol.* **18:**682-689.

3. **Cain, J. A., N. Solis, and S. J. Cordwell.** 2013. Beyond gene expression: The impact of protein post-translational modifications in bacteria. *J. Proteomics*.

4. **Castano-Cerezo, S., V. Bernal, J. Blanco-Catala, J. L. Iborra, and M. Canovas.** 2011. cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in *Escherichia coli*. *Mol. Microbiol.* **82:**1110-1128.

5. **Chan, C. H., J. Garrity, H. A. Crosby, and J. C. Escalante-Semerena.** 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.

6. **Chater, K. F.** 2006. *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361:**761-768.

7. **Courtois, S., C. M. Cappellano, M. Ball, F. X. Francou, P. Normand, G. Helynck, A. Martinez, S. J. Kolvek, J. Hopke, M. S. Osburne, P. R. August, R. Nalin, M. Guerineau, P. Jeannin, P. Simonet, and J. L. Pernodet.** 2003. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol.* **69:**49-55.

8. **Crosby, H. A., E. K. Heiniger, C. S. Harwood, and J. C. Escalante-Semerena.** 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.

9. **Crosby, H. A., D. A. Pelletier, G. B. Hurst, and J. C. Escalante-Semerena.** 2012. System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J. Biol. Chem.* **287:**15590-15601.

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

11. **Cumberlidge, A. G., and K. Isono.** 1979. Ribosomal protein modification in Escherichia coli. I. A mutant lacking the N-terminal acetylation of protein S5 exhibits thermosensitivity. *J. Mol. Biol.* **131:**169-89.

12. **Dass, B. K., R. Sharma, A. R. Shenoy, R. Mattoo, and S. S. Visweswariah.** 2008. Cyclic AMP in mycobacteria: characterization and functional role of the Rv1647 ortholog in *Mycobacterium smegmatis*. *J. Bacteriol.* **190:**3824-3834.

13. **Davies, J., and G. D. Wright.** 1997. Bacterial resistance to aminoglycoside antibiotics. *Trends Microbiol.* **5:**234-240.

14. **Delano, W. L.** 2002. The Pymol Molecular Graphics System.

15. **Denu, J. M.** 2005. The Sir2 family of protein deacetylases. *Curr. Op. Chem. Biol.* **9:**431-440.

16. **Glozak, M. A., N. Sengupta, X. Zhang, and E. Seto.** 2005. Acetylation and deacetylation of non-histone proteins. *Gene* **363:**15-23.

17. **Gosset, G., Z. Zhang, S. Nayyar, W. A. Cuevas, and M. H. Saier, Jr.** 2004. Transcriptome analysis of Crp-dependent catabolite control of gene expression in *Escherichia coli*. *J. Bacteriol.* **186:**3516-3524.

18. **Guan, K. L., and Y. Xiong.** 2011. Regulation of intermediary metabolism by protein acetylation. *Trends Biochem. Sci.* **36:**108-116.

19. **Gulick, A. M.** 2009. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* **4:**811-827.

20. **Hallows, W. C., S. Lee, and J. M. Denu.** 2006. Sirtuins deacetylate and activate mammalian acetyl-CoA synthetases. *Proc. Natl. Acad. Sci. U S A* **103:**10230-10235.

21. **Hu, L. I., B. P. Lima, and A. J. Wolfe.** 2010. Bacterial protein acetylation: the dawning of a new age. *Mol. Microbiol.* **77:**15-21.

22. **Huang, Y. H., L. Ferrieres, and D. J. Clarke.** 2006. The role of the Rcs phosphorelay in Enterobacteriaceae. *Res. Microbiol.* **157:**206-212.

23. **Hubbert, C., A. Guardiola, R. Shao, Y. Kawaguchi, A. Ito, A. Nixon, M. Yoshida, X. F. Wang, and T. P. Yao.** 2002. HDAC6 is a microtubule-associated deacetylase. *Nature* **417:**455-458.

24. **Ikeuchi, Y., K. Kitahara, and T. Suzuki.** 2008. The RNA acetyltransferase driven by ATP hydrolysis synthesizes N4-acetylcytidine of tRNA anticodon. *Embo J.* **27:**2194-2203.

25. **Isono, K., and S. Isono.** 1980. Ribosomal protein modification in Escherichia coli. II. Studies of a mutant lacking the N-terminal acetylation of protein S18. *Mol. Gen. Genet.* **177:**645-51.

26. **Iwabata, H., M. Yoshida, and Y. Komatsu.** 2005. Proteomic analysis of organ-specific post-translational lysine-acetylation and -methylation in mice by use of anti-acetyllysine and -methyllysine mouse monoclonal antibodies. *Proteomics* **5:**4653-4664.

27. **Khorasanizadeh, S.** 2004. The nucleosome: from genomic organization to genomic regulation. *Cell* **116:**259-272.

28. **Kim, D., B. J. Yu, J. A. Kim, Y. J. Lee, S. G. Choi, S. Kang, and J. G. Pan.** The acetylproteome of Gram-positive model bacterium *Bacillus subtilis*. *Proteomics* **13:**1726-1736.

29. **Kim, G. W., and X. J. Yang.** 2011. Comprehensive lysine acetylomes emerging from bacteria to humans. *Trends. Biochem. Sci.* **36:**211-220.

30. **Kim, S. C., R. Sprung, Y. Chen, Y. Xu, H. Ball, J. Pei, T. Cheng, Y. Kho, H. Xiao, L. Xiao, N. V. Grishin, M. White, X. J. Yang, and Y. Zhao.** 2006. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell.* **23:**607-618.

31. **Kouzarides, T.** 2000. Acetylation: a regulatory modification to rival phosphorylation? *Embo J* **19:**1176-1179.

32. **Kouzarides, T.** 2007. Chromatin modifications and their function. *Cell* **128:**693-705.

33. **Kuo, Y. M., and A. J. Andrews.** 2013. Quantitating the specificity and selectivity of Gcn5-mediated acetylation of histone H3. *PLoS One* **8:**e54896.

34. **Lee, D. W., D. Kim, Y. J. Lee, J. A. Kim, J. Y. Choi, S. Kang, and J. G. Pan.** Proteomic analysis of acetylation in thermophilic *Geobacillus kaustophilus*. *Proteomics* **13:**2278-2282.

35. **Lee, H. J., P. T. Lang, S. M. Fortune, C. M. Sassetti, and T. Alber.** 2012. Cyclic AMP regulation of protein lysine acetylation in *Mycobacterium tuberculosis*. *Nat. Struct. Mol. Biol.*

36. **Liang, W., and M. P. Deutscher.** 2010. A novel mechanism for ribonuclease regulation: transfer-messenger RNA (tmRNA) and its associated protein SmpB regulate the stability of RNase R. *J. Biol. Chem.* **285:**29054-8.

37. **Liang, W., and M. P. Deutscher.** 2012. Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). *RNA* **18:**37-41.

38. **Liu, X., L. Wang, K. Zhao, P. R. Thompson, Y. Hwang, R. Marmorstein, and P. A. Cole.** 2008. The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* **451:**846-850.

39. **Majdalani, N., and S. Gottesman.** 2005. The Rcs phosphorelay: a complex signal transduction system. *Annu. Rev. Microbiol*. **59:**379-405.

40. **Marsh, V. L., S. Y. Peak-Chew, and S. D. Bell.** 2005. Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba. *J. Biol. Chem*. **280:**21122-21228.

41. **Marvil, D. K., and T. Leisinger.** 1977. N-acetylglutamate synthase of *Escherichia coli*: purification, characterization, and molecular properties. *J. Biol. Chem*. **252:**3295-32303.

42. **McCue, L. A., K. A. McDonough, and C. E. Lawrence.** 2000. Functional classification of cNMP-binding proteins and nucleotide cyclases with implications for novel regulatory pathways in *Mycobacterium tuberculosis*. *Genome. Res.* **10:**204-219.

43. **McDonough, K. A., and A. Rodriguez.** 2012. The myriad roles of cyclic AMP in microbial pathogens: from signal to sword. *Nat. Rev. Microbiol.* **10:**27-38.

44. **McMahon, M. D., C. Guan, J. Handelsman, and M. G. Thomas.** 2012. Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. *Appl. Environ. Microbiol.* **78:**3622-3629.

45. **Mikulik, K., J. Felsberg, E. Kudrnacova, S. Bezouskova, D. Setinova, E. Stodulkova, J. Zidkova, and V. Zidek.** 2012. CobB1 deacetylase activity in *Streptomyces coelicolor*. *Biochem. Cell Biol.* **90:**179-187.

46. **Nambi, S., S. Badireddy, S. S. Visweswariah, and G. S. Anand.** 2012. Cyclic AMP-induced conformational changes in mycobacterial protein acetyltransferases. *J. Biol. Chem.* **287:**18115-18129.

47. **Nambi, S., N. Basu, and S. S. Visweswariah.** 2010. Cyclic AMP-regulated protein lysine acetylases in mycobacteria. *J. Biol. Chem.* **285:**24313-24323.

48. **Nambi, S., K. Gupta, M. Bhattacharya, P. Ramakrishnan, V. Ravikumar, N. Siddiqui, A. T. Thomas, and S. S. Visweswariah.** 2013. Cyclic AMP-dependent protein lysine acylation in mycobacteria regulates fatty acid and propionate metabolism. *J. Biol. Chem.* **288**.

49. **Phillips, D. M.** 1963. The presence of acetyl groups of histones. *Biochem. J.* **87:**258-263.

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

51. **Schwer, B., M. Eckersdorff, Y. Li, J. C. Silva, D. Fermin, M. V. Kurtev, C. Giallourakis, M. J. Comb, F. W. Alt, and D. B. Lombard.** 2009. Calorie restriction alters mitochondrial protein acetylation. *Aging Cell* **8:**604-606.

52. **Seow, K. T., G. Meurer, M. Gerlitz, E. Wendt-Pienkowski, C. R. Hutchinson, and J. Davies.** 1997. A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J. Bacteriol.* **179:**7360-7368.

53. **Soppa, J.** 2010. Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea* **2010:**pii: 820681.

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

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

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

57. **Starai, V. J., H. Takahashi, J. D. Boeke, and J. C. Escalante-Semerena.** 2003. Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae*. *Genetics* **163:**545-555.

58. **Sterner, D. E., and S. L. Berger.** 2000. Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* **64:**435-459.

59. **Thao, S., C. S. Chen, H. Zhu, and J. C. Escalante-Semerena.** 2010. N(epsilon)-Lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. *PLoS ONE* **5:**e15123.

60. **Thao, S., and J. C. Escalante-Semerena.** 2011. Biochemical and thermodynamic analyses of *Salmonella enterica* Pat, a multidomain, multimeric N(epsilon)-lysine acetyltransferase involved in carbon and energy metabolism. *mBio* **2:**e00216-11.

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

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

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

64. **Wang, Q., Y. Zhang, C. Yang, H. Xiong, Y. Lin, J. Yao, H. Li, L. Xie, W. Zhao, Y. Yao, Z. B. Ning, R. Zeng, Y. Xiong, K. L. Guan, S. Zhao, and G. P. Zhao.** 2010. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327:**1004-1007.

65. **Wardleworth, B. N., R. J. Russell, S. D. Bell, G. L. Taylor, and M. F. White.** 2002. Structure of Alba: an archaeal chromatin protein modulated by acetylation. *Embo J.* **21:**4654-4662.

66. **Weinert, B. T., V. Iesmantavicius, S. A. Wagner, C. Scholz, B. Gummesson, P. Beli, T. Nystrom, and C. Choudhary.** 2013. Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Mol. Cell* **51:**265-272.

67. **Wimpenny, J. W., and A. Firth.** 1972. Levels of nicotinamide adenine dinucleotide and reduced nicotinamide adenine dinucleotide in facultative bacteria and the effect of oxygen. *J. Bacteriol.* **111:**24-32.

68. **Wu, H., N. Moshkina, J. Min, H. Zeng, J. Joshua, M. M. Zhou, and A. N. Plotnikov.** Structural basis for substrate specificity and catalysis of human histone acetyltransferase 1. *Proc. Natl. Acad. Sci.* **109:**8925-8930.

69. **Wu, X., A. Vellaichamy, D. Wang, L. Zamdborg, N. L. Kelleher, S. C. Huber, and Y. Zhao.** Differential lysine acetylation profiles of *Erwinia amylovora* strains revealed by proteomics. *J. Proteomics* **79:**60-71.

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

71. **Yan, Y., N. A. Barlev, R. H. Haley, S. L. Berger, and R. Marmorstein.** 2000. Crystal structure of yeast Esa1 suggests a unified mechanism for catalysis and substrate binding by histone acetyltransferases. *Mol. Cell* **6:**1195-1205.

72. **Yan, Y., S. Harper, D. W. Speicher, and R. Marmorstein.** 2002. The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. *Nat. Struct. Biol.* **9:**862-869.

73. **Yu, B. J., J. A. Kim, J. H. Moon, S. E. Ryu, and J. G. Pan.** 2008. The diversity of lysine-acetylated proteins in *Escherichia coli*. *J. Microbiol. Biotechnol.* **18:**1529-1536.

74. **Yuan, H., and R. Marmorstein.** 2013. Histone acetyltransferases: Rising ancient counterparts to protein kinases. *Biopolymers* **99:**98-111.

75. **Zhang, J., R. Sprung, J. Pei, X. Tan, S. Kim, H. Zhu, C. F. Liu, N. V. Grishin, and Y. Zhao.** 2009. Lysine acetylation is a highly abundant and evolutionarily conserved modification in Escherichia coli. *Mol. Cell. Proteomics* **8:**215-25.

76. **Zhao, S., W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li, Y. Li, J. Shi, W. An, S. M. Hancock, F. He, L. Qin, J. Chin, P. Yang, X. Chen, Q. Lei, Y. Xiong, and K. L. Guan.** 2010. Regulation of cellular metabolism by protein lysine acetylation. *Science* **327:**1000-1004.

CHAPTER 2

BIOLOGICALLY ACTIVE ISOFORMS OF COBB SIRTUIN DEACETYLASE IN *SALMONELLA ENTERICA* AND *ERWINIA AMYLOVORA*[1]

[1] Tucker A.C. and J.C. Escalante-Semerena. 2010. *J. Bacteriol*. 192:6200-6208.
Reprinted here with permission from the publisher.

ABSTRACT

Sirtuins are NAD^+-dependent protein deacylases that are conserved in all domains of life, and are involved in diverse cellular processes, including control of gene expression and central metabolism. Eukaryotic sirtuins have *N*-terminal extensions that have been linked to protein multimerization and cellular localization. Here the first evidence of sirtuin isoforms in bacteria is reported. The enterobacterium *Salmonella enterica* synthesizes two isoforms of CobB sirtuin, a shorter 236-amino acid isoform (hereafter $CobB_S$), and a longer 273-amino acid isoform (hereafter $CobB_L$). The *N*-terminal 37-amino acid extension of $CobB_L$ is amphipathic, containing 18 basic amino acids (12 of which are Arg), and 13 hydrophobic ones; both isoforms were active *in vivo* and *in vitro*. Northern blot and transcription start site analyses revealed that *cobB* is primarily expressed as two monocistronic *cobB* mRNAs from two transcription start sites, one of which was mapped within the neighboring *ycfX* gene, and the other transcription start site was located within *cobB*. Additionally, a low-abundance *ycfX-cobB* bicistronic mRNA was observed, which could encode up to three proteins (YcfX, $CobB_L$, and $CobB_S$). $CobB_L$ isoforms are common within the family Enterobacteriaceae, but species of the genus *Erwinia* (including the plant pathogen *E. amylovora*) encode only the $CobB_L$ isoform. The $CobB_L$ isoform from *E. amylovora* restored growth of a *S. enterica cobB* strain on low acetate.

INTRODUCTION

Homologues of the yeast silence information regulator 2 protein (ySir2p, a.k.a. sirtuin) are Zn(II)-containing, NAD^+-dependent deacylases present in cells of all domains of life (21, 24, 35). In eukaryotes, sirtuin function impacts gene expression, metabolism, cancer, stress responses, and life-span control (5, 18, 22, 29, 38). Ongoing work on the regulation of sirtuin activity level has shown that sirtuins are regulated at the transcriptional, posttranscriptional, and posttranslational levels (22, 34).

In eukaryotes, some sirtuins contain *N*-terminal extensions of the enzymatic core, and these regions can affect binding of interacting partners (11, 20), facilitate oligomerization of sirtuins (46), mediate interactions with other sirtuin forms (8, 27), and direct cellular localization (8, 31). Thus, these

regions seem to contribute to the regulation of sirtuin function in the eukaryotic cells. It is not known whether sirtuin isoforms exist in bacteria or archaea, and if they do, what their functions may be.

Bacterial sirtuins play roles in short-chain fatty acid metabolism in *S. enterica* sv Typhimurium LT2 (hereafter *S. enterica*), *Bacillus subtilis* (16, 17, 38), and in the catabolism of aromatic and allicyclic acids in *Rhodopseudomona palustris* (10). In *S. enterica*, the protein acetyltransferase Pat acetylates acetyl-CoA synthetase (Acs), and propionylates propionyl-CoA synthetase (PrpE) (17, 37). In both cases, Pat acylates a conserved active-site lysine within the adenylation domain, effectively inactivating the enzyme by preventing the formation of the acetyl-AMP or propionyl-AMP intermediates (17, 36). Deacylation of propionylated PrpE ($PrpE^{Pr}$) or acetylated Acs (Acs^{Ac}) by the CobB sirtuin of *S. enterica* reactivates both enzymes, allowing cells to grow on propionate or low levels of acetate, respectively (17, 38).

The role of sirtuins in *S. enterica* central metabolism was recently expanded by data that suggest that a large number of central metabolic enzymes in *S. enterica* may be modified by reversible lysine acetylation (44). In addition to cellular metabolism, sirtuins have been linked to bacterial chemotaxis through its ability to deacetylate the response regulator CheY in *E. coli* (26).

Though the role of bacterial sirtuins is expanding, little is known about the transcriptional regulation of the genes encoding these enzymes. Wang *et al*. reported the transcriptional profile of the *cobB* gene of *S. enterica* (44), but the signals that induce *cobB* expression, and the location of the *cobB* promoter elements were not identified.

Based on the annotated *cobB* sequence of *S. enterica*, we sought to determine whether the CobB sirtuin contained an *N*-terminal extension to the sirtuin catalytic core as predicted by the bioinformatics analysis of the region. We also determined the transcription start site from which *cobB* mRNA is synthesized. Here, we provide evidence that *S. enterica* synthesizes two biologically active isoforms of CobB, one form of which contains a cationic 37-residue *N*-terminal extension. We used *in vitro* and *in vivo* approaches to demonstrate that each CobB isoform was active. We provide evidence that *cobB* is

transcribed from multiple promoters, and propose that these alternate transcripts may allow *S. enterica* to differentially regulate CobB isoform synthesis.

EXPERIMENTAL PROCEDURES

Bacteria strains and growth conditions. All bacterial strains and plasmids used in this study are listed in Table 2.1. Strains were grown in nutrient broth (NB), no-carbon essential (NCE) (3), or 3-(*N*-morpholino)propanesulfonic acid (MOPS) minimal medium (28). For western blots cells were grown using NCE minimal medium. For growth curves for which acetate was the sole carbon source, MOPS minimal medium was supplemented with $MgSO_4$ (1 mM), L-methionine (0.5 mM), and acetate (10 mM, pH 7.0). For growth on glycerol, MOPS minimal medium was supplemented with $MgSO_4$ (1 mM), trace minerals (1X), dicyanocobinamide (15 nM), 5,6-dimethylbenzamidazole (125 μM), and glycerol (22 mM). When used, ampicillin was at 100 μg/ml. Cells were grown at 37°C with shaking. For growth curves, cell density was measured using a microtiter plate and microtiter plate reader (Bio-Tek Instruments). Doubling times were calculated using the GraphPad Prism 4 software package (La Jolla, CA).

Strain construction. Gene deletions in *S. enterica* were constructed as using the phage lambda Red recombinase system as described (12). Notably, the *ycfX* deletion was specifically designed to delete the 5' portion (nt 4-300) of the *ycfX* gene without disrupting putative *cobB* regulatory elements that exist in the 3' region of the *ycfX* open reading frame.

Molecular techniques. DNA manipulations were performed using standard techniques (4, 14). Restriction endonucleases were purchased from Fermentas. DNA was amplified using PfuUltra II Fusion DNA polymerase (Stratagene), and site-directed mutagenesis was performed using the Quikchange™ Site Directed Mutagenesis kit (Stratagene). 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 sequencing

Table 2.1. Strains and plasmids used in this study.

Strain or Plasmid	Relevant Genotype	Reference[2]
S. enterica sv Typhimurium LT2 strains		
JE6583	*metE205 ara-9*	Lab collection
JE7088	Δ*metE2702 ara-9*	Lab collection
Derivatives of JE6583		
JE10806	Δ*ycfX128*	
JE11228	Δ*cobB1375*	
Derivatives of JE7088		
JE12939	Δ*cobB1375* Δ*cobT1380*	Lab collection
JE12972	Δ*cobB1375* Δ*cobT1380* / pBAD30	
JE12973	Δ*cobB1375* Δ*cobT1380* / pCOBB8 (or $pCobB_{L/S}$)	
JE12974	Δ*cobB1375* Δ*cobT1380* / pCOBB19 (or $pCobB_L$)	
JE12975	Δ*cobB1375* Δ*cobT1380* / pCOBB24 (or $pCobB_S$)	
JE13302	Δ*cobB1375* Δ*cobT1380* / pCOBB93 (or p*Ea*$CobB_L$)	
E. coli strains		
JE6090	C41λ(DE3)	Lab collection
Plasmids		
pBAD30	P_{araBAD} expression vector, bla^+	(19)
pTEV6	N-terminal, rTEV-cleavable MBP-fusion overexpression vector, bla^+	(30)
pCOBB8 (or $pCobB_{L/S}$)	$cobB^+$ cloned into pBAD30, bla^+	(38)
pCOBB19 (or $pCobB_L$)	*cobB1372* allele ($CobB^{M37A,M38A}$) in pBAD30, bla^+	
pCOBB24 (or $pCobB_S$)	*cobB1373* allele ($CobB^{M1A}$) in pBAD30, *bla*	
pCOBB93 (or p*Ea*$CobB_L$)	*E. amylovora* $cobB^+$ allele in pBAD30, bla^+	
pCOBB71	$CobB_S$ ORF in pTEV6	
pCOBB72	$CobB_L$ ORF in pTEV6	

[2] Unless otherwise indicated, all strains were constructed during the course of this work.

reactions were resolved at the University of Wisconsin-Madison Biotechnology Center. Oligonucleotide primer sequences are listed in Table 2.2.

***Cloning of* S. enterica *and* Erwinia amylovora cobB *genes*.** Cloning of *S. enterica cobB* has been described (38). Site-directed mutagenesis of start codons M1 or M37/M38 yielded *cobB* alleles *cobB1372* (encoding $CobB^{M37A,M38A}$ or $CobB_L$) on plasmid pCOBB19 (also referred to as $pCobB_L$), and *cobB1373* (encoding $CobB^{M1A}$ or $CobB_S$) on plasmid pCOBB24 (also referred to as $pCobB_S$).

The *E. amylovora cobB* gene was amplified from purified genomic DNA (a kind gift from N. Perna). The 1-kb fragment containing *Ea cobB*$^+$ was digested with *Sac*I and *Sph*I and ligated into pBAD30 (19), which had been digested with the same enzymes. The resulting 6-kb plasmid was called pCOBB93 or p*Ea*$CobB_L$.

Protein purification. *S. enterica cobB* was amplified from JE6583 genomic DNA using primers to add a 5' *Kpn*I site and a 3' *Sal*I site to the amplification product. For $CobB_L$ overproduction, nt 1-822 of the *cobB* allele were amplified and for $CobB_S$ nucleotides 109-822 of the *cobB* allele were amplified. Amplified fragments were digested with *Kpn*I and *Sal*I and subsequently ligated into plasmid pTEV6 (30) with the same enzymes. The resulting plasmids pCOBB71 and pCOBB72 direct synthesis of $CobB_S$ or $CobB_L$, respectively, with recombinant tobacco etch virus (rTEV) protease-cleavable N-terminal maltose-binding protein (MBP) tags.

Overproduction and purification methods were the same for both $CobB_L$ and $CobB_S$. The plasmids pCOBB71 and pCOBB72 were moved into E. coli strain C41(λDE3). The resulting strains were grown overnight and sub-cultured 1:100 (v/v) into 2 liters of super broth (SB) (13), containing ampicillin (100 μg/ml). Cultures were grown with shaking at 37 °C to $A_{600} \sim 0.7$, and MBP-H_6-CobB synthesis was induced with isopropyl- β-D-1-thiogalactopyranoside (IPTG, 0.5 mM). The culture was grown overnight at 37°C. Cells were harvested at 8000 x g for 15 minutes at 4°C in a Beckman Coulter Avanti J-2 XPI centrifuge fitted with rotor JLA-8.1000. Cell pellets were re-suspended in 30 ml cold His-Bind buffer (Buffer A) (sodium phosphate buffer (20 mM, pH 7.5), NaCl (500 mM), and imidazole (20 mM)). Cells

Table 2.2. Primers and DNA probe used in this study

Name	Sequence[3]
cobB deletion primers	
cobB(L) Wanner 5'	GGTGCTGCTTTTTTACATCTTACCGACTAATCAAA AAAAGAGGTTGTTATGGTGTAGGCTGGAGCTGCT TC
cobB(L) Wanner 3'	ACGTAACGTGAAATGTAGGCCGGATAAGGCGTTA CCGGGCAAACAGCACTACATATGAATATCCTCCTTAG
ycfX deletion primers	
ycfX1 Wanner 5'	GAGTATTGAGCGGCCAGTAATCGCGGAATATAAAAACAAGG AGCGGTGATGGTGTAGGCTGGAGCTGCTTC
ycfX1 Wanner 3'	ATCATCCCAGGCTTCGGACAAGGCAAAACAGTTAGCGTCATT GTCCAGACGCATATGAATATCCTCCTTAG
Cloning primers	
SacI+5'cobB_Eamy	CAT**GAGCTC**CGCTTTGCTAAGGCGCG
SphI+3'cobB_Eamy	ACT**GCATGC**TAAAGCGGGTTGGCCGTCG
KpnI + CobBL 5' pTEV6	GTA**GGTACC**ATGCAGTCGCGTCGGTTTCATCG
KpnI + CobBS 5' pTEV6	ATA**GGTACC**ATGGAAAACCCAAGAGTATTAGTCC
SalI + CobB 3' pTEV6	GAT**GTCGAC**CAGCACTACAGCCCTTTCAGG
Site-directed mutagenesis primers	
SeCobB M1A F	CAAAAAAAGAGGTTGTTGCGCAGTCGCGTCG
SeCobB M1A R	CGACGCGACTGCGCAACAACCTCTTTTTTTG
SeCobB M37A, M38A F	GACAGAGTGGTGCCGGAAGCGGCGGAAAACCCAAGAG
SeCobB M37A, M38A R	CTCTTGGGTTTTCCGCCGCTTCCGGCACCACTCTGTC
5' RACE Analysis primers	
CobB_499_RT	CGCCGTTCCATTCCAGAATCTGG
CobB_435_RT	ATGAATGATGTTGCGATTGC
CobB_325_RT	GATAAGCTTCCGCATTGGGTTGTATTTCC
Northern Blot Probes	
cobB_Northern_98-mer	CGGATAGACATGACCGGACGTGCCAATGGCGATAAAAA TATCCGCCATCGACAGCGCCATATAAATTTCATCCATGCCAA GCGGCATCTCGCCAAACC
ycfX_Northern_96-mer	TCTAAAAAGGCGCTATAGCTGGTATGGGGCGTGGGAACCCG TTTTTCCGAGTGCAGCCGCCGCGTTGAGTCAAATACGCCTAA CGCAATGTTTGTT

[3] Bold typeface indicates restriction sites

were placed on ice and lysed by sonication for 2 min (2-s pulse followed by 2 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*. Clarified cell extract was loaded onto a 5 ml HisTrap FF column (GE Healthcare) connected to a computer-controlled ÄKTA fast-protein liquid chromatography (FPLC) system. Unbound proteins were eluted off the column by extensive washing with buffer A. A 65-ml wash step with 95% buffer A, and 5% buffer B (sodium phosphate buffer (20 mM, pH 7.5), NaCl (500 mM), and imidazole (500 mM)) was applied to the column a prior to a 65-ml linear gradient 5-100% buffer B. All fractions containing MBP-H_6-CobB were combined. rTEV protease (33) was added to MBP-H_6-CobB and samples were dialyzed at room temperature for 3 h in buffer C (sodium phosphate buffer (20 mM, pH 7.5), NaCl (500 mM)) containing 1 mM dithiotreitol (DTT). CobB/rTEV mixtures were then dialyzed at 4°C against buffer C for 3 h, and again against buffer C containing imidazole (20 mM) for 12 h. After cleavage and dialysis, protein mixtures were passed over the 5-ml HisTrap column following the protocol described above. Protein that did not bind to the column was analyzed by sodium dodecyl sulfate polyacrylamide electrophoresis (SDS-PAGE) (23). Fractions containing $CobB_L$ or $CobB_S$ were pooled. $CobB_L$ and $CobB_S$ were stored in HEPES buffer (20 mM) containing NaCl (100 mM) and 10% glycerol. CobB concentrations were determined by measuring absorbance at 280 nm. The molar extinction coefficient used to calculate $CobB_L$ and $CobB_S$ concentration was 22,460 M^{-1} cm^{-1}.

Western blot analysis. A 200-ml overnight culture grown in NB was washed twice with 1X NCE medium. Cells were then inoculated into either NB or NCE minimal medium containing L-methionine (0.5 mM), $MgSO_4$ (1 mM) and supplemented with acetate (50 mM), glycerol (50 mM) or glucose (50 mM) to an OD_{650} of 0.1. For Western blots of citrate-grown cultures, overnight cultures grown in NB were inoculated to an OD_{650} of 0.025 into NCE minimal medium containing L-methionine (0.5 mM), $MgSO_4$ (1 mM) and supplemented with citrate (50 mM) without washing. When cultures reached mid-log or stationary phase, cultures were placed on ice for 10 min to arrest growth, followed by centrifugation at 8,000 x *g* for 15 min at 4°C in a Beckman Coulter Avanti J-25I centrifuge fitted with rotor JLA-16.250. Cells were re-suspended in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (50 mM,

pH 7.5) containing phenylmethylsulfonyl fluoride (PMSF, 1 mM) and ethylenediaminetetraacetic acid (EDTA, 10 mM). Cells were lysed by sonication on ice using a microtip at 50% duty, for 1 min. Cell lysates were spun at 16,000 x *g* for 10 min at 4°C in an Eppendorf 5415D microcentrifuge to remove unbroken cells and cellular debris. Proteins in cellular extracts were quantified by absorbance at a wavelength of 280 nm. Fifty or 400 mg of cellular protein were resolved using SDS-PAGE, and transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore). Rabbit polyclonal CobB antiserum (Laboratory Animal Resources, University of Wisconsin, Madison, WI) was used to detect isoforms of CobB.

CobB isoform abundance was quantified as described (43). Briefly, 2-fold dilutions of total cellular lysates (12.5 mg–1.6 mg) prepared as described above were resolved separated by SDS-PAGE using 1.5-mm thick, 12% polyacrylamide gels. Proteins were transferred and detected as described above. Band intensity was quantified using the TotalLab 1D gel analysis software v2003 (Nonlinear Dynamics, Durham, NC), and values within the linear range of detection were used to determine the relative CobB isoform levels. Values were reported as arbitrary units.

In vitro *deceacetylation assays*. Conditions for *in vitro* acetylation and deacetylation of Acs have been described (37). In the assays described here, we used homogeneous $CobB_L$, $CobB_S$, Pat, and Acs from *S. enterica* (17, 36, 37). Acs was radiolabeled using Pat protein acetyltransferase (37) and [1-^{14}C]acetyl-CoA in a final volume of 1 ml. Excess [1-^{14}C]acetyl-CoA was removed by buffer exchange; NAD^+ was added to the Acs^{Ac} reaction mixture to a final concentration of 1 mM, and deacetylation reactions were initiated by the addition of $CobB_L$ or $CobB_S$. Samples were periodically removed from the reaction mixture and quenched with SDS-PAGE loading buffer. Samples were resolved by SDS-PAGE, and subjected to phosphorimaging analysis to assess the acetylation state of Acs after incubation with $CobB_S$ or $CobB_L$.

Oligomeric state analysis of $CobB_L$ and $CobB_S$. Gel filtration was performed using a HiPrep 26/60 Sephacryl S-100 High Resolution Column (GE Healthcare) connected to a computer-controlled ÄKTA FPLC system. The column was equilibrated with sodium phosphate buffer (50 mM, pH 7.5)

containing NaCl (150 mM). $CobB_L$ (0.8 mg) or $CobB_S$ (1-2 mg) was applied to the column, which was developed isocratically at a rate of 2 ml min^{-1}. Molecular weight calibrations were performed using ovalbumin (44 kDa), myoglobin (17 kDa), and vitamin B_{12} (1.35 kDa) components of the Bio-Rad gel filtration standards. The above standards were supplemented with bovine serum albumin (66.4 kDa, Promega) and DNaseI (31 kDa, Sigma).

RNA isolation. For Northern blot analysis, RNA was isolated as described (1). *S. enterica* cultures were grown in NB to an OD_{650} of 0.8 with shaking at 37°C. 520 μl cells were added to 8X lysis buffer and phenol:water (3.75:1, Invitrogen) and were incubated at 65°C with shaking at 500 rpm for 20 min. Lysates were centrifuged at 16,000 x g for 10 min at room temperature in an Eppendorf 5415D microcentrifuge and the aqueous phase was extracted twice with 700 μl phenol:chloroform:isoamyl alcohol (25:24:1, Ambion). RNA was precipitated with 1.3 ml ethanol (100%) at -80°C for at least 30 min. RNA was pelleted by centrifugation at 4°C for 10 min as described above. Pellet was washed with 500 μl ethanol (75%) and centrifuged again for 10 min at 4°C. RNA pellets were allowed to dry at room temperature, and RNA was resuspended in tris(hydroxymethyl)aminomethane (Tris)-ethylenediaminetetraacetic acid (EDTA) buffer, pH 8.0 (Ambion).

To perform 5' rapid amplification of cDNA ends (5' RACE) analysis, RNA was isolated using the RNA Protect Reagent and the RNeasy Mini kit (Qiagen). Cells were grown to an OD_{650} of 0.7 and RNA was isolated according to the manufacturer's protocols. RNA was resuspended in RNase-free water.

Northern blot analysis. RNA (21 μg in 5.5 μl) was mixed with 1 μl MOPS buffer (1X), 3.5 μl formaldehyde, and 10 μl formamide (100%). RNA samples were incubated at 65°C for 15 min to denature the RNA. Samples were incubated on ice for 5 min, and 3 μl RNA agarose gel loading buffer (glycerol (50%), EDTA (1 mM), bromophenol blue (0.25%), xylene cyanol (0.25%)) was added. Samples were separated on a denaturing agarose gel (MOPS (1X), formaldehyde (2.2 M), agarose (1%)) with a Millennium marker (Ambion) run at 45 V for approximately 3 h. Gels were prepared for transfer as described (6).

RNA was transferred to 0.2 micron Hybond N^+ membrane (GE Healthcare) by capillary transfer for 16 h. RNA was linked to the membrane by baking 2 h at 80°C. Membranes were stained for 45 s for total RNA with methylene blue (0.03%) in sodium acetate (0.3 M, pH 5.2). Membranes were then de-stained for 2 min with DPEC-treated water. Membranes were pre-hybridized with Ultrahyb (Ambion) at 42°C for at least 30 min. The membrane was probed overnight with a [^{32}P]-radiolabeled single-stranded DNA probe (Table 2). Membranes were washed as described (6), and subjected to phosphorimage analysis to visualize relative mobility of the *cobB* and *ycfX* transcripts. The phosphorimage and menthylene blue-stained membrane image were overlayed using Adobe Photoshop CS2 (Adobe Systems, Inc.) to determine relative mobility of the *cobB and ycfX* transcripts.

5' RACE protocol. 5' RACE was performed using the Version 2, 5' RACE system from Invitrogen, according to the manufacturer's instructions. Nested, *cobB*-specific primers were used to synthesize the cDNA and for the subsequent amplification reactions. After the second round of amplification, the PCR products were separated by agarose gel electrophoresis and gel purified using the Wizard SV Gel and PCR Clean up System (Promega). Fragments were sequenced using a *cobB*-specific primer to determine the 5' end of the *cobB* transcript using the BigDye Terminator® v3.1 Cycle Sequencing Kit (Applied Biosystems) according to the manufacturer's protocol. Reaction mixtures were resolved and analyzed at the University of Wisconsin-Madison Biotechnology Facility.

RESULTS AND DISCUSSION

Multiple start codons of sirtuins are conserved in **Enterobacteriaceae.** An alignment of *cobB* homologues from representative members of the family *Enterobacteriaceae* suggested that most of these bacteria might synthesize two isoforms of CobB, a short ($CobB_S$) and a long ($CobB_L$) isoform (Fig. 2.1). The one exception we found was the plant pathogen *Erwinia amylovora*, whose genome sequence appeared to direct the synthesis of only one form, $CobB_L$; this feature was conserved within all sequenced *Erwinia* species (data not shown). Notably, all the *N*-terminal extensions of the putative $CobB_L$ isoforms synthesized by enterobacteria were amphipathic in nature, with many conserved basic residues (Arg, Lys

highlighted in gray), and in several instances, additional Arg residues were present in the extension, but were not conserved in all genera. The extended peptide of the *S. enterica* $CobB_L$ isoform contained 17 basic [12 Arg, 2 Lys, 2 Gln, and 1 Asn; 17/37 = 46%], and 11 hydrophobic amino acids (4 Phe, 4 Leu, 2 Val, and 1 Ile; 11/37 = 30%).

***S. enterica* synthesizes short and long isoforms of CobB sirtuin.** The *cobB* gene of *S. enterica* contains two putative translation initiation sites within the first 38 codons (M1 and M38, respectively; Fig. 2.1). Although another potential start codon, M37, lies juxtaposed to M38, we considered M38 to be the biologically relevant start codon due to its conservation in other enterobacteria (Fig. 2.1). To determine whether different CobB isoforms were synthesized from the putative M1 and M38 start codons, we grew *S. enterica* in rich medium (NB) or in NCE minimal medium with acetate (50 mM) as the sole source of carbon and energy. Under all of these conditions, *S. enterica* expressed two isoforms of CobB, which matched the predicted size of purified CobB standards (Fig. 2.2A). Samples were serially diluted and densitometry was used to quantify the relative abundance of each isoform on Western blots. As shown in figure 2.2B, the shorter, 236-amino acid $CobB_S$ isoform (start at M38) was expressed at levels approximately 7- or 11-fold higher than the longer 273-amino acid $CobB_L$ isoform (start at M1) on NB and acetate, respectively. Previous work demonstrated that *cobB* is translated most highly at mid-log and late-log stages of growth (44). Therefore, we assessed relative $CobB_L$ and $CobB_S$ levels in *S. enterica* during mid-log phase when grown in NB or minimal medium supplemented with acetate, citrate, glucose, or glycerol (50 mM). Under these conditions $CobB_S$ levels were about 5- to 20-fold higher than $CobB_L$ levels (Fig. 2.2B, C). The expression of these two CobB isoforms at differing levels suggested a previously unknown level of regulation of the CobB sirtuin in *S. enterica*.

***The $CobB_L$ and $CobB_S$ isoforms deacetylate Acs^{Ac}* in vitro.** CobB sirtuin is the only annotated NAD^+-dependent deacetylase of *S. enterica*. Previous *in vitro* work with $CobB_S$ demonstrated that this isoform efficiently deacetylates acetylated acetyl-CoA synthetase (Acs^{Ac}) *in vitro* (36). However, prior to this work, the ability of $CobB_L$ to perform this reaction had not been assessed. To assess the deacetylase activity of $CobB_L$, *S. enterica* Acs was acetylated with radiolabeled acetyl-CoA, and the product of the

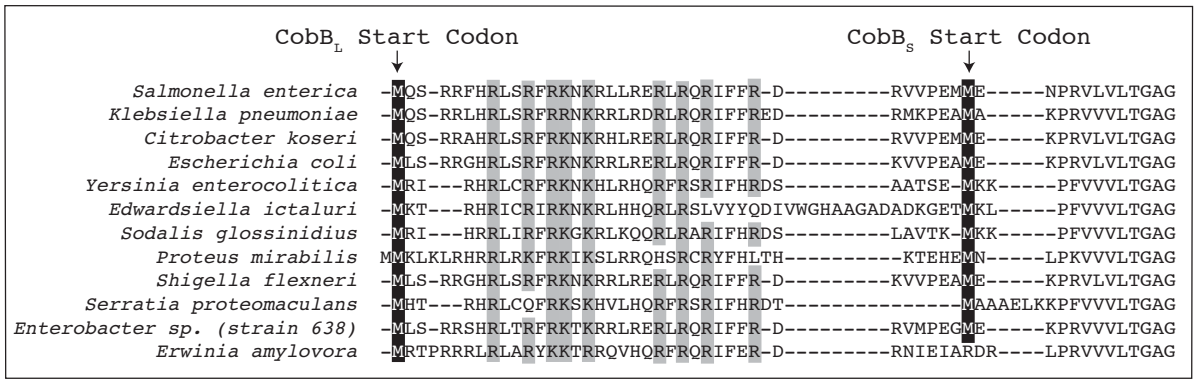


Figure 2.1. CobB$_L$ and the positively charged N-terminal region are conserved among the family *Enterobacteriaceae*. The *N*-terminal regions of CobB homologues from the family *Enterobacteriaceae* were aligned. CobB$_L$ and CobB$_S$ starting methionines are highlighted in black. Conserved, positively charged residues (Arg and Lys) are highlighted in grey.

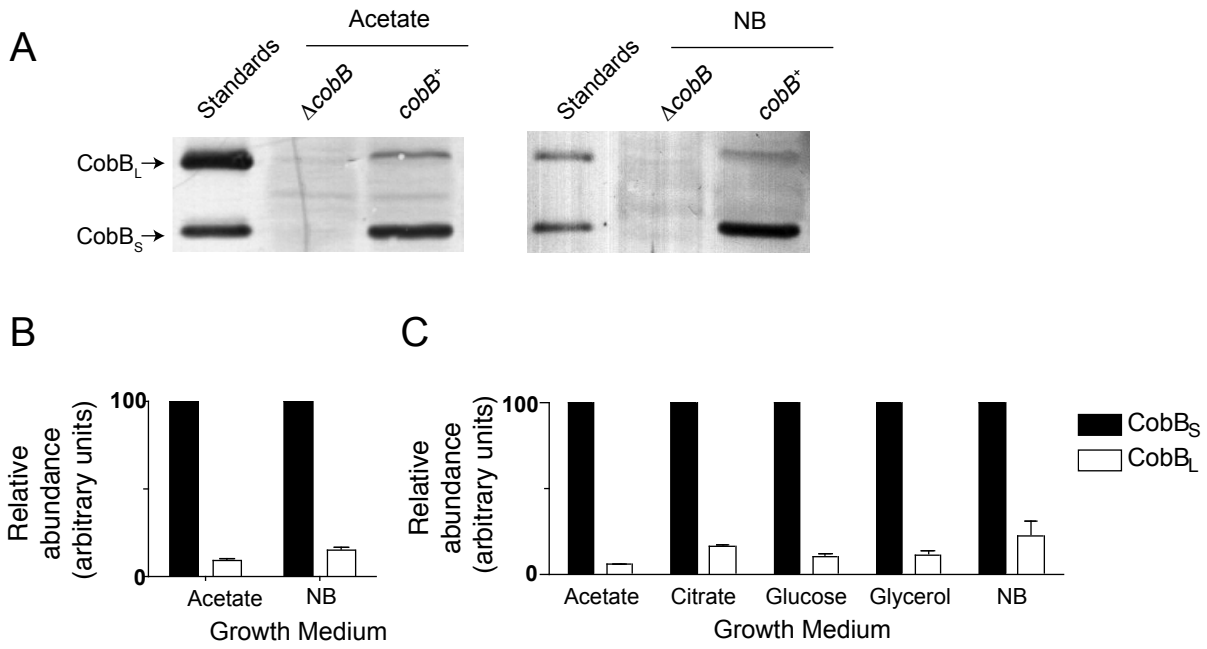


Figure 2.2. *S. enterica* synthesizes two isoforms of CobB. Western blot analysis performed using rabbit α-CobB antiserum on whole-cell lysates of wild type and *ΔcobB* strains grown to stationary phase on NB and 50 mM acetate (A). Purified *Se*CobBL (31.1 kDa) and *Se*CobBS (26.2 kDa) proteins were included as molecular weight standards. Relative quantities of *Se*CobB$_L$ and *Se*CobB$_S$ were determined for cells grown to stationary phase (B) or mid-log phase (C) in NB or minimum medium supplemented with acetate, citrate, glucose, or glycerol (50 mM each). Quantities were normalized to *Se*CobB$_S$ levels, which were assigned a value of 100 arbitrary units.

reaction was used as the substrate for *in vitro* deacetylation activity assays. The decreasing amount of radiolabel associated with Acs over time showed that both $CobB_L$ and $CobB_S$ deacetylated Acs^{Ac} *in vitro* in a NAD^+-dependent manner (Fig. 2.3).

***$CobB_L$ and $CobB_S$ deacetylate Acs^{Ac}* in vivo.** To determine whether $CobB_L$ and $CobB_S$ were active *in vivo*, we used a sirtuin-deficient *S. enterica* strain JE12939 (Δ*cobB* Δ*cobT*). To assess the activity of the two CobB isoforms in vivo, we relied on the role of CobB during growth on 10 mM acetate (38). Under such conditions, acetate activation depends on Acs activity, which is regulated by reversible N^e-Lys acetylation (36). The lack of *cobB* function in strain JE12939 prevents deacetylation and re-activation of Acs^{Ac} (36, 38). Therefore, growth of strain JE12939 on 10 mM acetate would only occur if $CobB_L$ or $CobB_S$ were active, that is, if Acs^{Ac} were deacetylated.

To assess $CobB_S$ and $CobB_L$ activity in vivo, we introduced plasmids encoding arabinose-inducible alleles *cobB1372* ($CobB^{M37A,M38A}$ = $CobB_L$) and *cobB1373* ($CobB^{M1A}$ = $CobB_S$) into strain JE12939 (Δ*metE* Δ*cobB* Δ*cobT*). In figure 2.4A, we show that $CobB_L$ and $CobB_S$ restored growth of strain JE12939 on 10 mM acetate equally well. Western blot analysis confirmed that each plasmid expressed the expected CobB isoform (Fig. 2.4B). As expected, strain JE12939 failed to grow in the absence of CobB, a phenotype that was corrected by the introduction of the $cobB^+$ allele (positive control). Together, the data support the conclusion that $CobB_L$ and $CobB_S$ are functional *in vivo*.

To determine whether the ability of $CobB_L$ to deacetylate Acs^{Ac} was conserved among the *Enterobacteriaceae*, we asked whether the *E. amylovora cobB* gene could restore growth of *S. enterica* strain JE12939 on 10 mM acetate. As shown in figure 2.4, the *E. amylovora* $CobB_L$ (*Ea*$CobB_L$) protein supported growth of the *S. enterica* JE12939 strain on 10 mM acetate.

The $CobB_L$ and $CobB_S$ isoforms have phosphoribosyltransferase activity. It has been known for some time that CobB compensates for the lack of the nicotinate mononucleotide (NaMN):5,6-dimethylbenzimidazole (DMB) phosphoribosyltransferase (CobT) enzyme during the synthesis of coenzyme B_{12} from its precursor cobinamide (40, 41). It was not known, however, whether both isoforms of CobB could compensate for CobT. Thus, we determined whether *Se*$CobB_L$, *Se*$CobB_S$, and *Ea*$CobB_L$

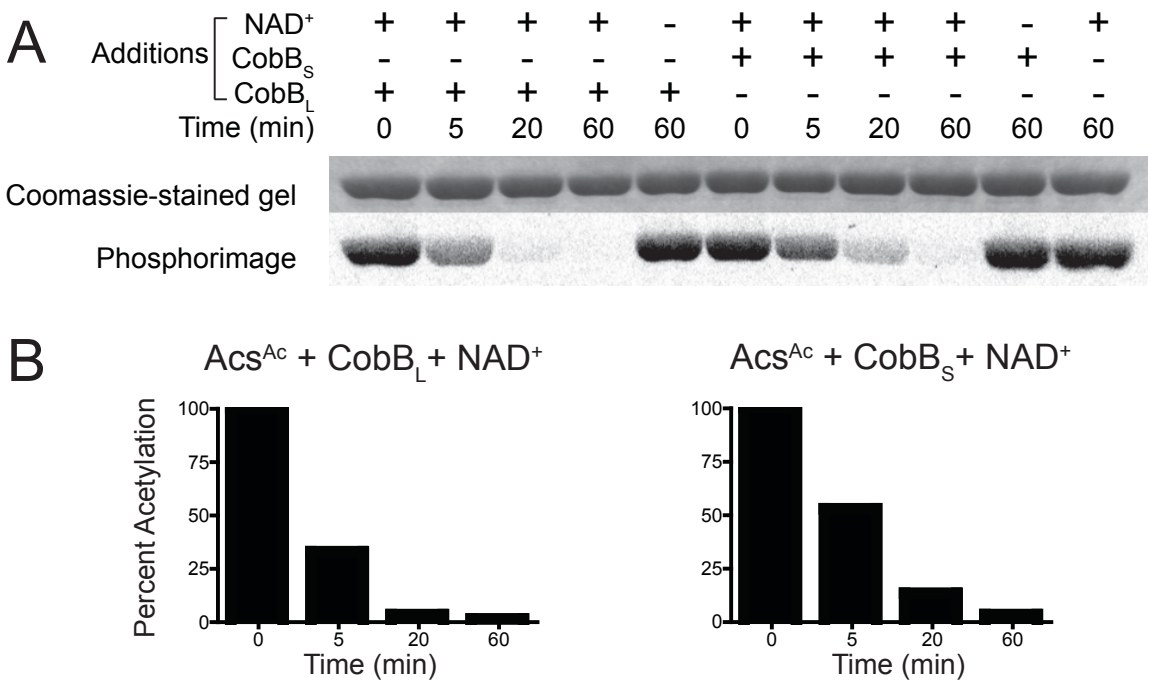


Figure 2.3. *Se*CobB$_L$ and *Se*CobB$_S$ deacetylate Acs^{Ac} in vitro. *Se*CobB$_L$ or *Se*CobB$_S$ was incubated with NAD$^+$ and Acs^{Ac} that was previously acetylated by Pat and [14C-1]acetyl-CoA. Samples were removed as a function of time, and were quenched with gel-loading buffer. Proteins were resolved by SDS-PAGE, and analyzed using phosphorimaging (A). The amount of radiolabeled AcsAc remaining at each time point was quantified by densitometry (B).

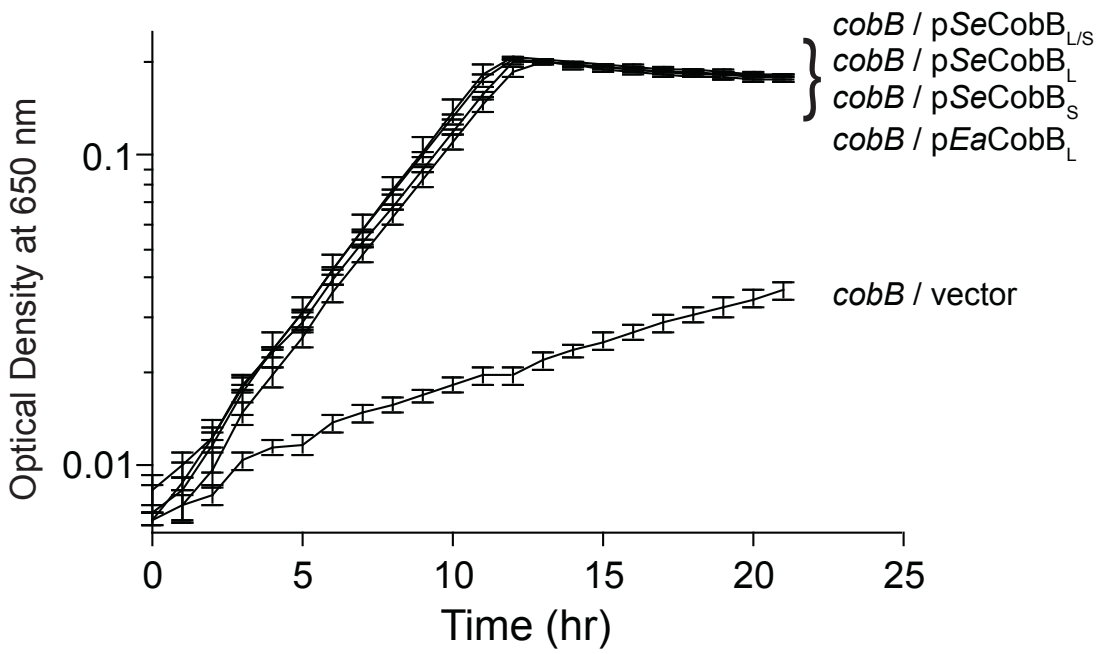


Figure 2.4. *Se*CobB$_L$, *Se*CobB$_S$, and *Ea*CobB$_L$ proteins restore growth of a *S. enterica cobB* strain on low acetate. Growth behavior of *S. enterica* on MOPS minimal medium supplemented with acetate (10 mM). Growth experiments were performed at 37°C using a microtiter plate and microtiter plate reader (Bio-Tek Instruments).

could restore adenosylcobalamin biosynthesis in a strain that carried a chromosomal deletion of the *cobT* gene. To ensure that the plasmid was the only source of CobB protein, we also deleted the chromosomal copy of *cobB*. Thus, the *S. enterica* JE12939 strain (Δ*metE* Δ*cobB* Δ*cobT*) used in this analysis was a cobalamin auxotroph, because it cannot synthesize adenosylcobalamin from the precursors cobinamide and DMB. Hence, growth of strain JE12939 on glycerol + cobinamide + DMB would only occur if *Se*CobB$_L$, *Se*CobB$_S$, or *Ea*CobB$_L$ could activate DMB. Derivatives of strain JE12929 that synthesized *Se*CobB$_L$, *Se*CobB$_S$, or *Ea*CobB$_L$ grew on minimal medium supplemented with glycerol (22 mM), dicyanocobinamide (15 nM), and DMB (125 mM) at very similar doubling times (h), namely: JE12939 / p*Se*CobB$_L$, 1.2 $\pm$ 0.05; JE12939 / p*Se*CobB$_B$, 1.1 $\pm$ 0.12; JE12939 / p*Ea*CobB$_L$, 1.5 $\pm$ 0.04; and JE12939 / vector, 13.3 $\pm$ 0.04. The above data show that all forms of CobB tested activated DMB, and restored adenosylcobalamin biosynthesis in a *S. enterica* strain lacking the CobT enzyme. Results from Western blot analysis confirmed that plasmids used in this experiment expressed the expected CobB isoform (Fig. 2.5).

***Initial characterization of the CobB$_L$* N-*terminus*.** Previous work with the yeast sirtuin Hst2 demonstrated that an N-terminal extension to the core sirtuin domain could mediate sirtuin oligomerization (46). Gel filtration analysis of purified *Se*CobB$_L$ and *Se*CobB$_S$ demonstrated that both isoforms exist as monomers under the conditions tested (Fig. 2.6). Thus the N-terminal extension of *Se*CobB$_L$ does not mediate sirtuin oligomerization under the conditions tested.

Amphipathic helices have been demonstrated to allow passage of small peptides through membranes (15). Because the N-terminus of CobB is amphipathic in composition, we used secondary structure prediction software to determine if this structure was capable of forming an amphipathic helix. The software package JPred3 (9) predicts that residues 6-28 form an alpha helix. However, a helical wheel projection of the predicted helix (Fig. 2.7) suggested that this region of *Se*CobB$_L$ had charged residues evenly dispersed around the central axis of the helix, and would not exist as a canonical amphipathic helix in which one face along the axis is hydrophobic while the other face is hydrophilic (32, 39).

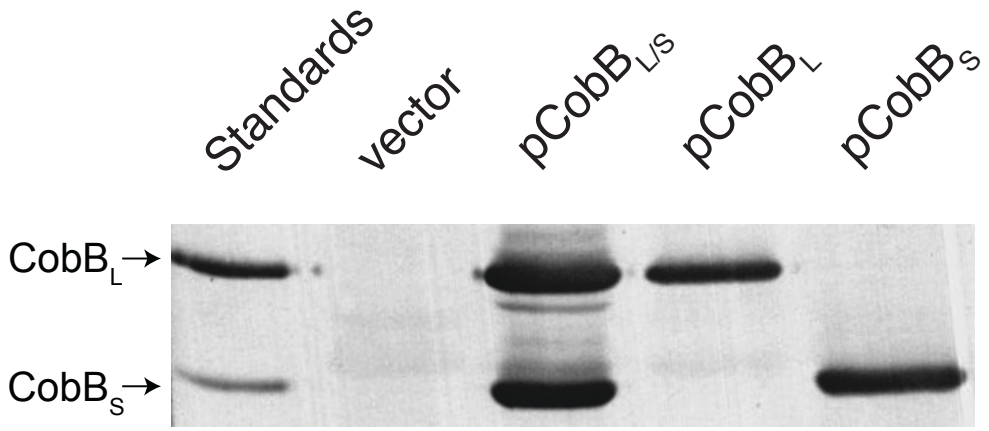


Figure 2.5. *Se*CobB$_{L}$ and *Se*CobB$_{S}$ are synthesized from the pCobB$_{L}$ and pCobB$_{S}$ plasmids when cells are grown on glycerol + dicyanocobinamide + DMB. Western blot analysis of *S. enterica cobB* strains overproducing CobB$_{L}$, CobB$_{S}$, or CobB$_{L/S}$ during growth on glycerol (22 mM), dicyanocobinamide (15 nM), and DMB (125 μM). CobB$_{L}$ (CobB^{M37A,M38A}, encoded by *cobB1372*); CobB$_{S}$ (CobB^{M1A}, encoded by *cobB1373*).

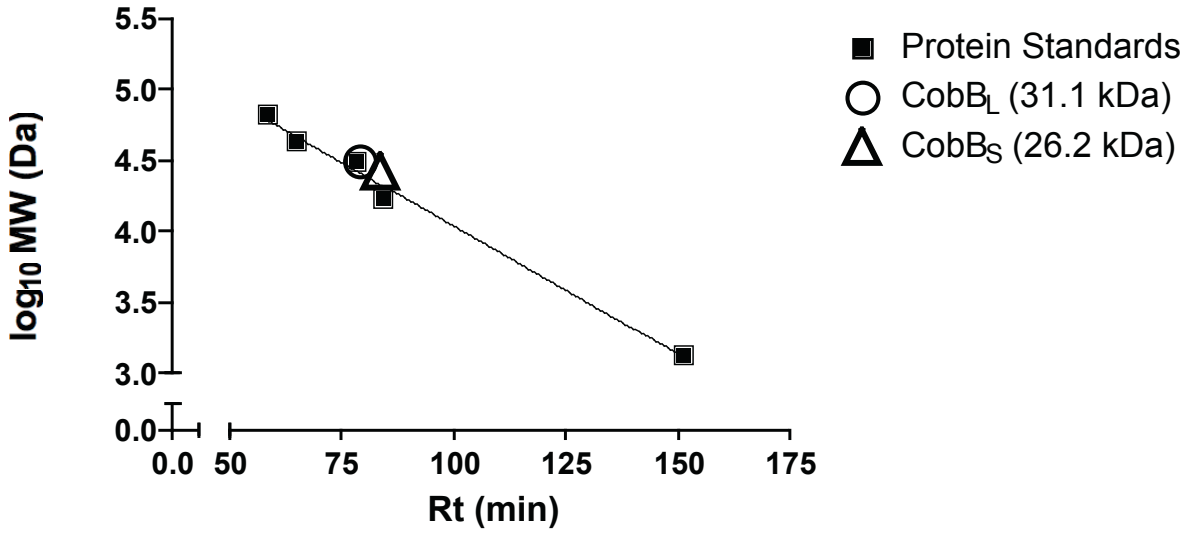


Figure 2.6. Gel filtration analysis of *Se*CobB$_L$ and *Se*CobB$_S$. Samples were applied to a HiPrep 26/60 Sephacryl S-100 High Resolution gel filtration column using isocratic elution with sodium phosphate (50 mMM, pH 7.4) containing 150 mM NaCl. Calibration was performed with Bio-Rad gel filtration standards supplemented with BSA and DNaseI with linear regression to build the standard curve. Micration of *Se*CobB$_L$ and *Se*CobB$_S$ through the matrix was analyzed and the retention time was consistent with an apparent mass of monomeric enzyme states.

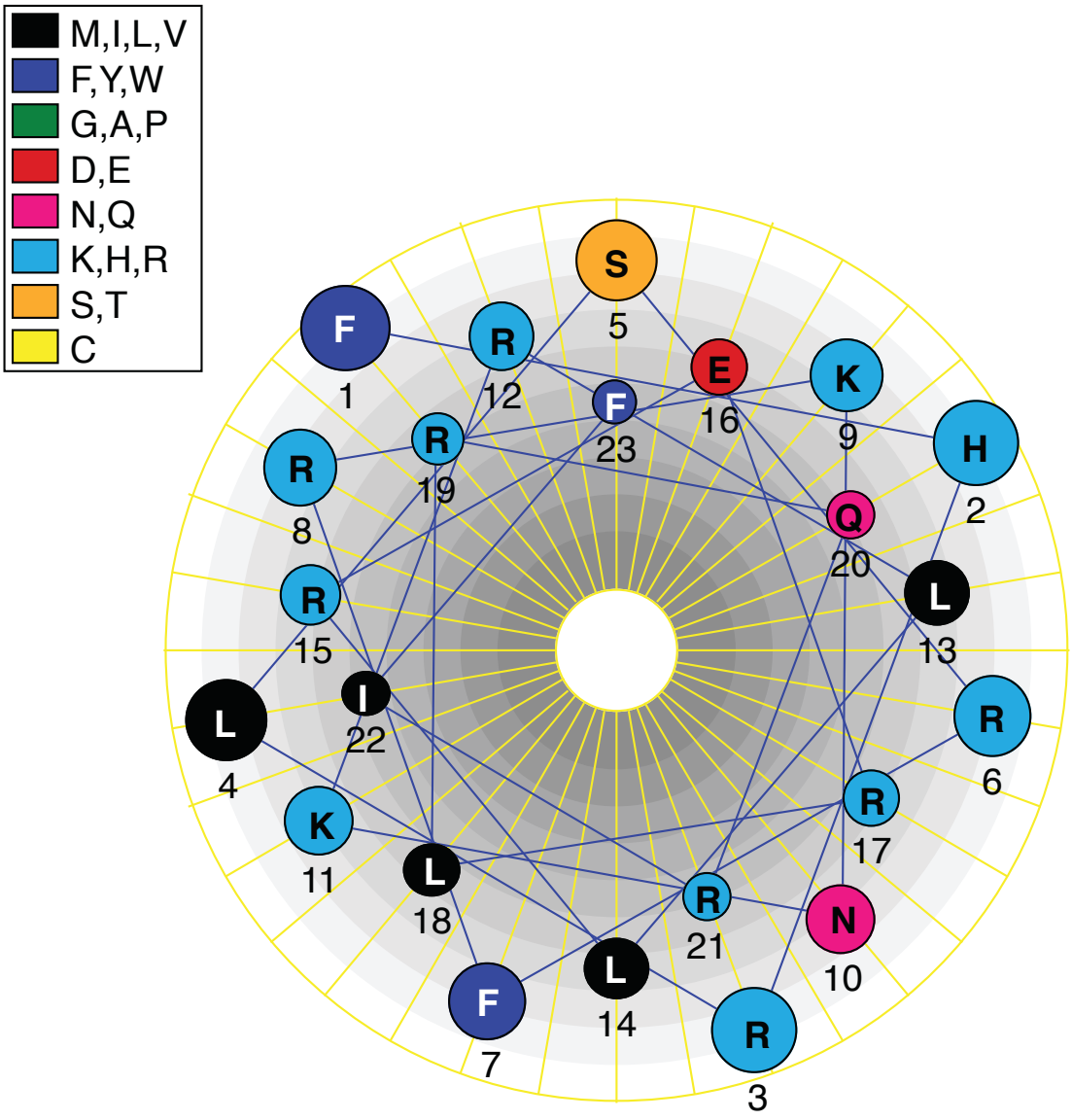


Figure 2.7. Helical wheel projection of $CobB_L$ N-terminal residues. Residues 6-28 (N-FHRLSRFRKNKRLLRERLRQRIF-C) of the $CobB_L$ N-terminus were subjected to helical wheel analysis using the DNAStar Protean program (Madison, WI).

In *S. enterica*, *cobB* is transcribed from two promoters independent of the neighboring *ycfX* gene. To gain insights into the mechanisms that control the synthesis of CobB isoforms in *S. enterica*, we investigated whether the cell synthesized ≥1 *cobB* mRNA. We note that, as currently ascribed, *cobB* is located 18 nucleotides 3' of *ycfX*, a homolog of the *nagK* gene that in *E. coli* encodes *N*-acetylglucosamine kinase (42). The proximity of *cobB* to *ycfX* suggested that *cobB* and *ycfX* might constitute an operon. In fact, information available online ascribes *nagK (ycfX) cobB* as an operon (http://regulondb.ccg.unam.mx). Although Uehara & Park suggested that a promoter for *cobB* lied within *nagK*, the idea was not pursued (42).

We used Northern blot analysis to determine the length of the *cobB* transcript in *S. enterica*. If *cobB* (822 bp) and *ycfX* (912 bp) were co-transcribed, we expected a >1.7-kb transcript. Using a *cobB*-specific probe, we identified a dominant transcript approximately 1000 nt in size, which suggested that the transcript spanned only *cobB* (Fig. 2.8A). Noteworthy was the small amount of signal appearing around 1800 nt that was detected using *cobB*-specific and *ycfX*-specific probes (Fig. 2.8A, B), but was absent in either the *ΔycfX* or *ΔcobB* strains, which suggested that a small amount of CobB was synthesized from a *ycfX-cobB* transcript.

We used 5' RACE analysis to determine the transcription start sites of *cobB*. PCR amplification of the 5' RACE products revealed two transcripts of different lengths. The sequences of these 5' RACE products revealed a transcription site 5' of each translation start site (Fig. 2.8C). One transcription start site was located within *ycfX* (nt 731), whilst the other was located within *cobB* (nt 44). The precise *cobB* transcription start site located within *ycfX* was somewhat ambiguous because the 5' RACE protocol resulted in a poly-G tail on the 5' end of the transcript. It is possible that the G at the -1 site relative to the upstream *cobB* promoter could be the +1 nucleotide of the upstream *cobB* transcription start site.

A schematic of the *cobB* promoters is shown in figure 2.8D. Promoters were assigned based on the Northern blot analysis (for the *ycfX-cobB* promoter, P1), or Northern blot and 5' RACE analysis (for the two promoters initiating monocistronic *cobB* mRNAs at P2 and P3). A small amount of *cobB* mRNA is transcribed from P1, and the P1-derived transcript contains the entire *cobB* open reading frame. The P2-

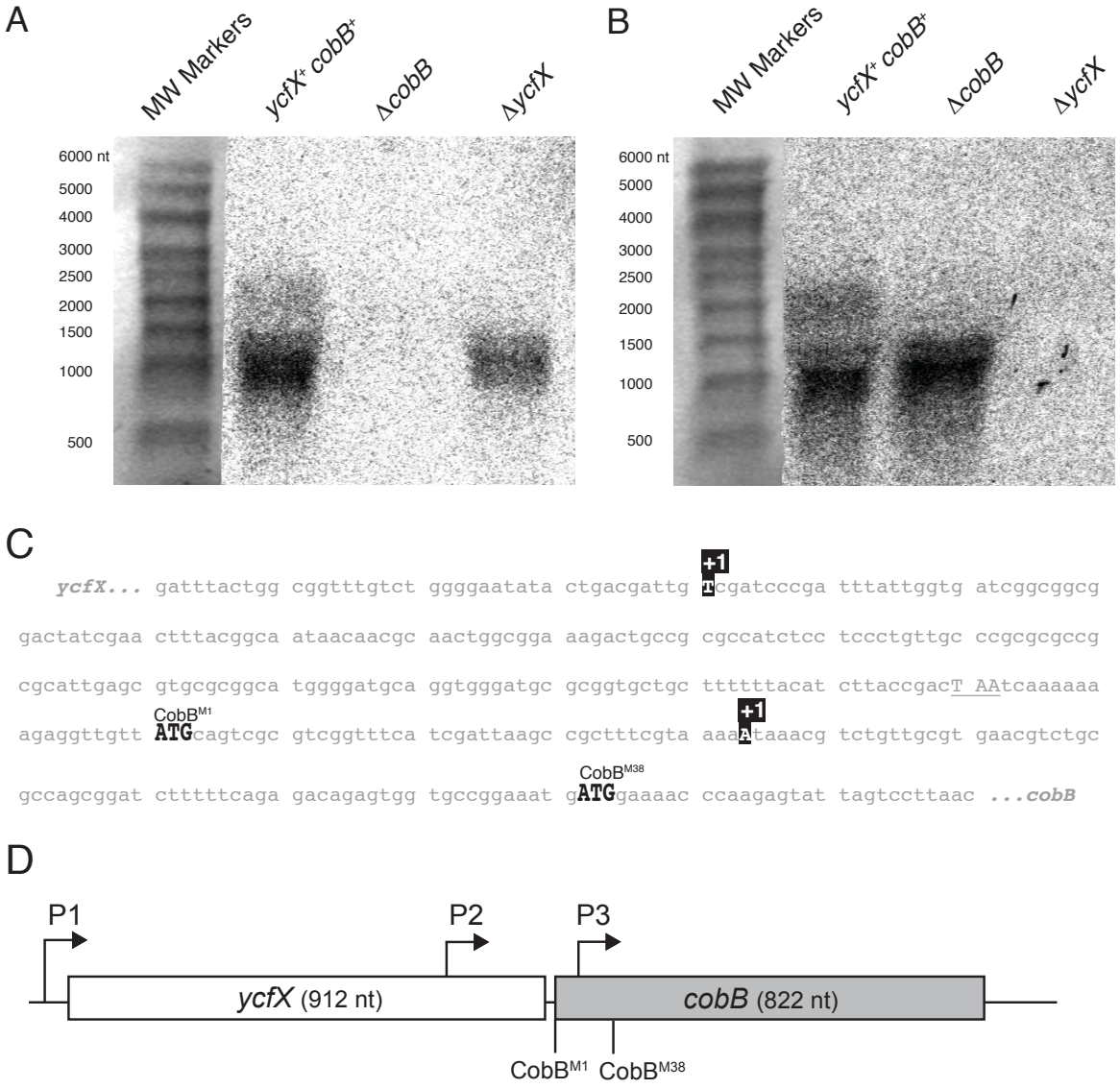


Figure 2.8. The *cobB* gene has two promoters independent of the *ycfX* promoter. Northern blots of wildtype (*cobB*$^{+}$ *ycfX*$^{+}$), *ΔcobB*, and *ΔycfX* strains probing for *cobB* (A) or *ycfX* (B) transcripts. RNA Millenium® markers were used as molecular weight standards. The phosphorimage was overlayed over the total RNA-stained membrane to align the lanes. C. Identity of the *cobB* transcription start sites within the *ycfX* and *cobB* open reading frames. The transcription start sites are denoted by +1. The CobB translation start sites are highlighted in gray. The *ycfX* stop codon is underlined. D. Model of the *cobB* promoters P1, P2, and P3 (drawn to scale). Assignment of promoters was based on Northern blot analysis (P1), or Northern blot analysis and transcription start site data (P2 and P3). $CobB_L$ start codon is noted with $CobB^{M1}$. $CobB_S$ start codon is noted with $CobB^{M38}$.

derived transcript encompasses both the open reading frames encoding $CobB_L$ and $CobB_S$. $CobB_L$ is likely synthesized from the P2-derived transcript, but it is unclear whether or not $CobB_S$ is synthesized from this transcript. Promoter P3 exists within the *cobB* open reading frame. The P3-derived transcript only spans the open reading frame encoding $CobB_S$ suggesting that only $CobB_S$ can be synthesized from this transcript. By separating the expression of each *cobB* isoform with independent transcription start sites, *S. enterica* could differentially regulate each transcription start site and, thus each CobB isoform.

Notably, when the *cobB* open reading frame is expressed from the P_{BAD} promoter, levels of $CobB_L$ and $CobB_S$ seem to be equalized (Fig. 2.4B), relative to when the *cobB* allele is expressed from the chromosome (Fig. 2.2A). The *cobB* complementation vector $pCobB_{WT}$ lacks *cobB* P1 and P2 promoters, but does contain the vector-based P_{BAD} promoter and *cobB* P3. Therefore $CobB_L/CobB_S$ ratios seem to be regulated, in part, by copy number of the *cobB* allele and/or the identity of the promoters driving *cobB* expression.

Contributions from this work. This is the first report of sirtuin isoforms in bacteria or archaea. The existence of sirtuin isoforms in a single bacterium suggests that these enzymes have more than one physiological role. We have shown that *S. enterica* synthesizes two biologically active isoforms of the CobB sirtuin deacetylase, which are synthesized from two sets of in-frame start codons. Long and short CobB sirtuin isoforms appear to be common among the *Enterobacteriaceae*, with the notable exception of members of the genus *Erwinia*, which appear to synthesize only the long sirtuin isoform. Our data show that the $CobB_L$ isoform of *E. amylovora* is biologically active (Fig. 2.4). In *S. enterica*, the two CobB isoforms do not appear to be the result of processing of $CobB_L$ into $CobB_S$, instead, the isoforms are synthesized from alternative start codons. Examination of *cobB* gene expression revealed an additional level of regulation in which multiple *cobB* mRNAs are synthesized from transcription start sites located 5' of each *cobB* translation initiation site (Fig. 2.8).

***What roles could multiple isoforms of CobB be playing in* S. enterica?** Unique roles for $CobB_L$ and $CobB_S$ are not immediately obvious. It is striking that, among the bacterial $CobB_L$ isoforms we analyzed (Fig. 2.1), on average 37% of the residues in the $CobB_L$ extension are Lys or Arg, with 78% of

the positively charged residues being Arg. The structure of CobB from the enterobacterium *E. coli* has been determined (PDB: 1S5P) (47), however, since the structure was solved using $CobB_S$, it is not clear what structure of the *N*-terminus of $CobB_L$ is or how it could be interacting with other macromolecules in the bacterial cell. Since *E. amylovora* is only capable of synthesizing $CobB_L$, it seem that this isoform is the biologically relevant isoform for this organism and is capable of carrying out all necessary $CobB_L$ functions in *E. amylovora*.

Although both isoforms of CobB can deacetylate Acs^{Ac}, Acs^{Ac} may not be the only physiologically relevant substrate for both isoforms. Recent work suggests that approximately 191 proteins in *S. enterica* are modified by acetylation (44), and putative protein partners of CobB in *E. coli* have been identified using a multiprotein complex purification procedure (7). The *N*-terminal region of CobB may be important for mediating interaction with some of these putative acetylated protein substrates. Alternatively, the positively charged *N*-terminus of CobB may also be needed to interact with other molecules or structures in the cell. Non-ribosomal, Arg-rich peptides with amino acid composition similar to the *N*-terminus of $CobB_L$ have been implicated in protein-RNA interactions (2, 25, 45). Thus, the *N*-terminus of $CobB_L$ could bind nucleic acids, or target the isoform to another negatively charged molecule of the cell. Thus, many questions remain unanswered. For example: i) Why are the levels of $CobB_L$ and $CobB_S$ different? ii) Does the positively charged *N*-terminus of $CobB_L$ affect its activity, localization, or substrate specificity? iii) Are there any physiological conditions under which either isoform is essential? These questions need to be address as we broaden our understanding of the physiological role of bacterial sirtuins.

REFERENCES

1. **Aiba, H., S. Adhya, and B. de Crombrugghe.** 1981. Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. *J. Biol.Chem.* **256:**11905-11910.

2. **Bayer, T. S., L. N. Booth, S. M. Knudsen, and A. D. Ellington.** 2005. Arginine-rich motifs present multiple interfaces for specific binding by RNA. *RNA* **11:**1848-1857.

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

4. **Bloch, K. D., and B. Grossman.** 1995. Digestion of DNA with restriction endonucleases, p. Unit 3.1.1-3.1.21. *In* F. M. Ausubel, R. E. R. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, a. Smith, and K. Struhl (ed.), *Current protocols in molecular biology.* Greene Publishing Associates & Wiley Interscience, New York, N.Y.

5. **Brachmann, C. B., J. M. Sherman, S. E. Devine, E. E. Cameron, L. Pillus, and J. D. Boeke.** 1995. The SIR2 gene family, conserved from bacteria to humans, functions in silencing, cell cycle progression, and chromosome stability. *Genes Dev.* **9:**2888-2902.

6. **Brown, T., K. Mackey, and T. Du.** 2004. Analysis of RNA by Northern and slot blot hybridization, p. Unit 4.9.1-4.1.19. *In* F. M. Ausubel, R. E. R. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, a. Smith, and K. Struhl (ed.), *Current protocols in molecular biology.* Greene Publishing Associates & Wiley Interscience, New York, N.Y.

7. **Butland, G., J. M. Peregrin-Alvarez, J. Li, W. Yang, X. Yang, V. Canadien, A. Starostine, D. Richards, B. Beattie, N. Krogan, M. Davey, J. Parkinson, J. Greenblatt, and A. Emili.** 2005. Interaction network containing conserved and essential protein complexes in *Escherichia coli. Nature* **433:**531-537.

8. **Cockell, M. M., S. Perrod, and S. M. Gasser.** 2000. Analysis of Sir2p domains required for rDNA and telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* **154:**1069-1083.

9. **Cole, C., J. D. Barber, and G. J. Barton.** 2008. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res*. **36:**W197-201.

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

11. **Cuperus, G., R. Shafaatian, and D. Shore.** 2000. Locus specificity determinants in the multifunctional yeast silencing protein Sir2. *Embo J* **19:**2641-2651.

12. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97:**6640-6645.

13. **Elbing, K., and R. Brent.** 2002. Media preparation and bacteriological tools, p. Unit 1.1.3. *In* F. M. Ausubel, R. E. R. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, A. Smith, and K. Struhl (ed.), *Curr. Prot. Molec. Biol.* Wiley Interscience, New York.

14. **Elion, E. A., P. Marina, and L. Yu.** 2007. Constructing recombinant DNA molecules by PCR, p. Unit 3.17.1-3.17.12. *In* F. M. Ausubel, R. E. R. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, a. Smith, and K. Struhl (ed.), *Current protocols in molecular biology.* Greene Publishing Associates & Wiley Interscience, New York, N.Y.

15. **Futaki, S., T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda, and Y. Sugiura.** 2001. Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. *J. Biol. Chem*. **276:**5836-58340.

16. **Gardner, J. G., and J. C. Escalante-Semerena.** 2009. In *Bacillus subtilis*, the sirtuin protein deacetylase encoded by the *srtN* gene (formerly *yhdZ*), and functions encoded by the *acuABC* genes control the activity of acetyl-CoA synthetase. *J. Bacteriol*. **191:**1749-1755.

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

18. **Guarente, L.** 2007. Sirtuins in aging and disease. Cold Spring Harb. Symp. Quant. Biol. **72:**483-488.

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

20. **Hou, Z., J. R. Danzer, L. Mendoza, M. E. Bose, U. Muller, B. Williams, and C. A. Fox.** 2009. Phylogenetic conservation and homology modeling help reveal a novel domain within the budding yeast heterochromatin protein Sir1. *Mol. Cell. Biol.* **29:**687-702.

21. **Imai, S., C. M. Armstrong, M. Kaeberlein, and L. Guarente.** 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* **403:**795-800.

22. **Kwon, H. S., and M. Ott.** 2008. The ups and downs of SIRT1. *Trends Biochem. Sci*. **33:**517-525.

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

24. **Landry, J., J. T. Slama, and R. Sternglanz.** 2000. Role of NAD(+) in the deacetylase activity of the SIR2-like proteins. *Biochem. Biophys. Res. Commun*. **278:**685-690.

25. **Lazinski, D., E. Grzadzielska, and A. Das.** 1989. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* **59:**207-218.

26. **Li, R., J. Gu, Y. Y. Chen, C. L. Xiao, L. W. Wang, Z. P. Zhang, L. J. Bi, H. P. Wei, X. D. Wang, J. Y. Deng, and X. E. Zhang.** 2010. CobB regulates *Escherichia coli* chemotaxis by deacetylating the response regulator CheY. *Mol. Microbiol.* **76:**1162-1174.

27. **Moazed, D., A. Kistler, A. Axelrod, J. Rine, and A. D. Johnson.** 1997. Silent information regulator protein complexes in *Saccharomyces cerevisiae*: a SIR2/SIR4 complex and evidence for a regulatory domain in SIR4 that inhibits its interaction with SIR3. *Proc. Natl. Acad. Sci. U S A* **94:**2186-2191.

28. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. *J. Bacteriol.* **119:**736-747.

29. **North, B. J., B. L. Marshall, M. T. Borra, J. M. Denu, and E. Verdin.** 2003. The human Sir2 ortholog, SIRT2, is an NAD(+)-dependent tubulin deacetylase. *Mol. Cell* **11:**437-444.

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

31. **Schwer, B., B. J. North, R. A. Frye, M. Ott, and E. Verdin.** 2002. The human silent information regulator (Sir)2 homologue hSIRT3 is a mitochondrial nicotinamide adenine dinucleotide-dependent deacetylase. *J. Cell Biol.* **158:**647-57.

32. **Segrest, J. P., H. De Loof, J. G. Dohlman, C. G. Brouillette, and G. M. Anantharamaiah.** 1990. Amphipathic helix motif: classes and properties. *Proteins* **8:**103-117.

33. **Shih, Y. P., H. C. Wu, S. M. Hu, T. F. Wang, and A. H. Wang.** 2005. Self-cleavage of fusion protein in vivo using TEV protease to yield native protein. *Protein Sci.* **14:**936-941.

34. **Smith, B. C., W. C. Hallows, and J. M. Denu.** 2008. Mechanisms and molecular probes of sirtuins. *Chem. Biol. Interact.* **15:**1002-1013.

35. **Smith, J. S., C. B. Brachmann, I. Celic, M. A. Kenna, S. Muhammad, V. J. Starai, J. L. Avalos, J. C. Escalante-Semerena, C. Grubmeyer, C. Wolberger, and J. D. Boeke.** 2000. A phylogenetically conserved NAD+-dependent protein deacetylase activity in the Sir2 protein family. *Proc. Natl. Acad. Sci. U S A* **97:**6658-6663.

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

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

38. **Starai, V. J., H. Takahashi, J. D. Boeke, and J. C. Escalante-Semerena.** 2003. Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae*. *Genetics* **163:**545-555.

39. **Tossi, A., L. Sandri, and A. Giangaspero.** 2000. Amphipathic, alpha-helical antimicrobial peptides. *Biopolymers* **55:**4-30.

40. **Trzebiatowski, J. R., G. A. O'Toole, and J. C. Escalante-Semerena.** 1994. The *cobT* gene of *Salmonella typhimurium* encodes the NaMN: 5,6-dimethylbenzimidazole phosphoribosyltransferase responsible for the synthesis of N^1-(5-phospho-alpha-D-ribosyl)-5,6-dimethylbenzimidazole, an intermediate in the synthesis of the nucleotide loop of cobalamin. *J. Bacteriol.* **176:**3568-3575.

41. **Tsang, A. W., and J. C. Escalante-Semerena.** 1998. CobB, a new member of the SIR2 family of eucaryotic regulatory proteins, is required to compensate for the lack of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase activity in *cobT* mutants during cobalamin biosynthesis in *Salmonella typhimurium* LT2. *J. Biol. Chem*. **273:**31788-31794.

42. **Uehara, T., and J. T. Park.** 2004. The N-acetyl-D-glucosamine kinase of *Escherichia* coli and its role in murein recycling. *J. Bacteriol*. **186:**7273-7279.

43. **Venembre, P., N. Seta, A. Boutten, M. Dehoux, M. Aubier, and G. Durand.** 1994. Comparison of enhanced chemiluminescence and colorimetric techniques for the immuno-detection of alpha 1-antitrypsin. *Clin. Chim. Acta* **227:**175-184.

44. **Wang, Q., Y. Zhang, C. Yang, H. Xiong, Y. Lin, J. Yao, H. Li, L. Xie, W. Zhao, Y. Yao, Z. B. Ning, R. Zeng, Y. Xiong, K. L. Guan, S. Zhao, and G. P. Zhao.** 2010. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327:**1004-1007.

45. **Weiss, M. A., and N. Narayana.** 1998. RNA recognition by arginine-rich peptide motifs. *Biopolymers* **48:**167-180.

46. **Zhao, K., X. Chai, A. Clements, and R. Marmorstein.** 2003. Structure and autoregulation of the yeast Hst2 homolog of Sir2. *Nat. Struct. Biol.* **10:**864-71.

47. **Zhao, K., X. Chai, and R. Marmorstein.** 2004. Structure and substrate binding properties of CobB, a Sir2 homolog protein deacetylase from *Escherichia coli*. *J. Mol. Biol.* **337:**731-741.

CHAPTER 3

ACETOACETYL-COA SYNTHETASE ACTIVITY IS CONTROLLED BY A PROTEIN ACETYLTRANSFERASE WITH UNIQUE DOMAIN ORGANIZATION IN *STREPTOMYCES LIVIDANS*[1]

[1] Tucker A.C. and J.C. Escalante-Semerena. 2013. *Mol. Micro*. 87:152-167.
Reprinted here with permission from the publisher.

ABSTRACT

GCN5-type *N*-acetyltransferases (GNATs) are enzymes that catalyze the transfer of the acetyl group from acetyl-CoA to a primary amine. GNATs are conserved in all domains of life. Some members of this family of enzymes acetylate the side chain of specific lysine residues in proteins of diverse function. In bacteria, GNAT-catalyzed protein acetylation regulates carbon metabolism, RNA metabolism, and transcriptional regulation. Metabolic regulation in *Streptomyces* species is of interest due to the role of these organisms in natural product synthesis. Here we identify *Sl*PatA, a GNAT in *S. lividans* with unique domain organization, and a new acetylation target, namely acetoacetyl-CoA synthetase (*Sl*AacS). The latter has homologues in all domains of life. *In vitro* and *in vivo* evidence show that *Sl*AacS is a bona fide acetoacetyl-CoA synthetase. *Sl*PatA acetylates *Sl*AacS more efficiently than it does acetyl-CoA synthetase, an enzyme known to be under acetylation control. *Sl*PatA acetylates *Sl*AacS at the active site residue Lys617 and acetylation inactivates *Sl*AacS. Acetylated *Sl*AacS was deacetylated by a sirtuin-type protein deacetylase. *Sl*AacS acetylation/deacetylation may represent a conserved mechanism for regulation of acetoacetyl-CoA synthetase activity in all domains of life.

INTRODUCTION

GCN5-related *N*-acetyltransferase (GNAT) enzymes catalyze the transfer of an acetyl moiety from acetyl-coenzyme A (Ac-CoA) to a primary amine of small molecules and proteins. GNATs, named for the homology to the yeast GCN5 protein (yGCN5p), are identified by signature sequence motifs and structurally homology (53). GNATs are conserved in all domains of life, and represent one of the largest proteins super families (66). GNATs are involved in acetylation of antibiotics, hormones, tRNA, histones, metabolic enzymes, and transcription factors, implicating GNATs in a wide variety of cellular processes (29, 56, 61, 63, 66). Protein lysine acetylation by GNATs occurs on the epsilon amino group of the lysine side chain. Acetylation neutralizes the positive charge of the lysine side chain altering the side-chain chemistry. Protein acetylation has been documented in all domains of life (55), but recent proteomic approaches performed in eukaryotes (12, 39, 68, 74) and bacteria (15, 72, 73) have expanded our view of

the potential role of this posttranslational modification in cell physiology. As new protein acetylation targets and the concomitant protein acetyltransferases are identified, the role of acetylation on protein and cellular functions is broadened.

GNATs are the only class of protein acetyltransferases identified in bacteria. The Gram-negative γ-proteobacterium *Salmonella enterica* encodes a protein acetyltransferase *Se*Pat that contains a large *N*-terminal domain (~700 residues) with homology to NDP-forming CoA synthetases and a *C*-terminal GNAT domain (~200 residues) (59). *Se*Pat lacks NDP-forming Ac-CoA synthetase activity, and the role of the domain remains unclear. Biochemical studies indicate that the *N*-terminal domain of *Se*Pat may be required for acetyltransferase activity, subunit interactions, and positive cooperativity (62). *Se*Pat uses acetyl-CoA to acetylate acetyl-CoA synthetase (*Se*Acs) and propionyl-CoA to propionylate propionyl-CoA synthetase (PrpE) (22, 60). In each case, acylation occurs at the active site lysine of the CoA synthetase and inactivates the enzyme. Specifically, lysine acylation prevents the first half reaction that activates the fatty acid to the corresponding fatty-acyl-AMP intermediate. In each case, deacylation reactivates the enzyme (22, 58).

In the photoheterotrophic purple non-sulfur α-proteobacterium *Rhodopseudomonas palustris*, a *Se*Pat homologue (*Rp*Pat), and a single-domain GNAT, *Rp*KatA, regulate organic acid degradation by acetylating and inactivating seven acyl-CoA synthetases and three aryl-CoA synthetases (14, 15). Acetylation of the acyl- or aryl- CoA synthetases has the same effect as seen in *S. enterica*. Notably, the protein deacetylation system of *R. palustris* is more complex than the one in *S. enterica*. *R. palustris* encodes a sirtuin-type, NAD^+-dependent deacetylase *Rp*SrtN, and a Zn(II)-dependent histone deacetylase homologue *Rp*LdaA. *Rp*LdaA hydrolyzes the acetyl group from acetyllysine releasing free acetate and reactivating acyl-CoA synthetases. At present, the only protein with acetyllysine deacetylase activity in *S. enterica* is a sirtuin encoded by the *cobB* gene (64, 65). Protein acetyltransferases have also been described in Gram-positive bacteria. *Bacillus subtilis* encodes a single domain GNAT, AcuA, which acetylates acetyl-CoA synthetase (*Bs*AcsA) (21). *Mycobacterium tuberculosis* and *M. smegmatis* encode protein acetyltransferases, *Mt*PatA and *Ms*Pat, respectively, with a cyclic-AMP binding domain fused to

the *N*-terminus of a GNAT domain (46). Binding of cAMP to *Mt*PatA activates the enzyme by exposing the catalytic site (41). *Ms*Pat acetylates a universal stress protein and Ac-CoA synthetase in *M. smegmatis* (46, 69).

Recently, Mikulik et al. demonstrated that Ac-CoA synthetase from the actinomycete *Streptomyces coelicolor* is acetylated *in vivo* (45), but the acetyltransferase responsible for this modification remains unidentified. *S. coelicolor* encodes 77 putative GNAT acetyltransferases (Pfam00583), whilst the genome of the closely related species *S. lividans* (32) encodes 72 putative GNATs. However, none of the putative GNATs share end-to-end homology with the known bacterial protein acetyltransferases *Se*Pat, *Ec*Pka, *Rp*Pat, *Rp*KatA, *Bs*AcuA, *Ms*Pat, or *Mt*PatA. Further understanding of the regulation of metabolism in *Streptomyces* species is of interest because of the diversity of natural products synthesized by these organisms (11, 13, 44, 52).

Here, we identify a GNAT in *S. lividans* that has protein acetyltransferase activity and contains regions of homology to both the *N*- and *C*-terminal domains of *Se*Pat. The *S. lividans* Pat homologue (*Sl*PatA) can acetylate native Ac-CoA synthetase (*Sl*Acs), albeit to a limited extent, but acetylates acetoacetyl-CoA synthetase (*Sl*AacS) substantially better. Prior to this work, the regulation of *Sl*AacS activity by lysine acetylation/deacetylation systems was unknown. Consistent with the effect of acetylation on other AMP-forming acyl-CoA synthetases, acetylation of *Sl*AacS inactivated the enzyme, whilst deacetylation reactivated it. We also report evidence that *Sl*PatA acetylates *Sl*AacS *in vivo* in *S. lividans*. Further, we used *E. coli* as a heterologous host to demonstrate *Sl*AacS regulation by *Sl*PatA *in vivo*. In summary, we show that a *Se*Pat homologue with reversed domain organization has bona fide acetyltransferase activity. Our findings raise the possibility that acetoacetyl-CoA synthetase homologues in all domains of life may be also controlled by lysine acetylation/deacetylation systems.

EXPERIMENTAL PROCEDURES

Bacterial strains and growth conditions. All strains and plasmids used in this study are listed in Tables 3.1 and 3.2, respectively. *Streptomyces* strains are derivatives of *Streptomyces lividans* TK24. ISP-2 medium (54) or R2YE medium (38) was used to culture *S. lividans* on solid medium. Liquid cultures of

Table 3.1. Strains used in this study.

Strain or Plasmid	Relevant Genotype	Reference[2]
E. coli strains		
MG1655	Wild type	
HB101	F$^-$, *thi-1*, *hsdS20* (rB$^-$, mB$^-$), *supE44*, *recA13*, *ara-14*, *leuB6*, *proA*2, *lacY1*, *galK2*, *rpsL20* (*str*^{r}), *xyl*-5, *mtl-1*	Promega
Stellar	F$^-$, *endA1*, *supE44*, *thi-1*, *recA1*, *relA1*, *gyrA96*, *phoA*, *Φ80d lacZ* Δ*M15*, Δ(*lacZYA - argF*) *U169,* Δ*(mrr – hsdRMS - mcrBC),* Δ*mcrA,* λ$^-$	Clonetech
JE6090	C41λ(DE3)	Lab Collection
JE9314	C41λ(DE3) *pka12*::*kan*$^+$	Lab Collection
Derivatives of MG1655		
JE15076	Δ*pka* Δ*cobB* Δ*atoDA*	
JE17126	Δ*pka* Δ*atoDA*	
Derivatives of JE15076		
JE15976	Δ*pka* Δ*cobB* Δ*atoDA* / pBAD30 *bla*$^+$	
JE15977	Δ*pka* Δ*cobB* Δ*atoDA* / pATO1 *bla*$^+$	
JE16068	Δ*pka* Δ*cobB* Δ*atoDA* / p*Sl*AacS4 *bla*$^+$	
JE16968	Δ*pka* Δ*cobB* Δ*atoDA* / p*Sl*AacS6 pSRK-Km *bla*$^+$ *kan*$^+$	
JE16969	Δ*pka* Δ*cobB* Δ*atoDA* / p*Sl*AacS6 p*Sl*PatA9 *bla*$^+$ *kan*$^+$	
Derivatives of JE17126		
JE17146	Δ*atoDA* Δ*pka* / p*Sl*AacS6 pSRK-Km *bla*$^+$ *kan*$^+$	
JE17147	Δ*atoDA* Δ*pka* / p*Sl*AacS6 *Sl*PatA9 *bla*$^+$ *kan*$^+$	
Derivatives of HB101		
JE16095	*F*$^-$, *thi-1*, *hsdS20* (rB$^-$, mB$^-$), *supE44*, *recA13*, *ara-14*, *leuB6*, *proA2*, *lacY1*, *galK2*, *rpsL20* (*str*^{r}), *xyl-5*, *mtl-1* / pRK2013 *kan*$^+$	(19)
Derivatives of Stellar		

[2] Unless otherwise indicated, all strains were constructed during the course of this work.

Table 3.1. (cont'd) Strains used in this study.

S. lividans strains		
TK24	Wild type	Lab Collection
Derivatives of S. lividans TK24		
JE16707	Δ*patA*	
JE16752	ΔEFD65580 ΔEFD68590 ΔEFD71509	
JE16731	TK24 / pSlAacS5 tsr+	
JE16758	Δ*aacS* (EFD70521)	
Derivative of JE16707		
JE16821	Δ*patA* / p*Sl*AacS5 tsr+	

Table 3.2. Plasmids used in this study.

Plasmids	Relevant Genotype	Reference[3]
pBAD30	P_{araBAD} expression vector, *bla*$^+$	Lab Collection
pTEV5	N-terminal, rTEV-cleavable His_6-tag overexpression vector, *bla*$^+$	(50)
pSRK-Km	*lacI*^{q}*-lac* promoter-operator expression vector, *kan*$^+$	(35)
pSE34	PermE constitutive expression vector, *bla*$^+$ *tsr*$^+$	(67)
pRK2013	self-transmissible helper plasmid, *mob*$^+$ *tra*$^+$ *kan*$^+$	(19)
pKC1139	Temperature-sensitive shuttle vector, *apr*$^+$	(5)
pACS10	*S. enterica acs*$^+$ allele in pTYB1	(58)
pATO1	*E. coli atoDA* allele in pBAD30, *bla*$^+$	
p*Sl*Acs1	*S. lividans acs*$^+$ allele (EFD68454) in pTEV5, *bla*$^+$	
p*Sl*AacS1	*S. lividans aacS*$^+$ allele (EFD70521) in pTEV5, *bla*$^+$	
p*Sl*AacS4	*S. lividans aacS*$^+$ allele (EFD70521) in pBAD30, *bla*$^+$	
p*Sl*AacS5	*S. lividans aacS*$^+$ allele (EFD70521) in pSE34, *bla*$^+$ *tsr*$^+$	
p*Sl*AacS6	*E. coli* codon-optimized *S. lividans aacS* allele (EFD70521) in pBAD30, *bla*$^+$	
p*Sl*PatA1	*S. lividans patA*$^+$ allele (EFD66247) in pTEV5, *bla*$^+$	
p*Sl*PatA9	*S. lividans patA*$^+$ allele (EFD66247) in pSRK-Km, *bla*$^+$	
pKC1139-Δ*patA*	deletion construct of *S. lividans* *patA*$^+$ allele (EFD66247) in pKC1139, *apr*$^+$	

[3] Unless otherwise indicated, all strains were constructed during the course of this work.

S. lividans were grown in either yeast extract-malt extract (YEME) rich medium or NMMP medium supplemented with lithium acetoacetate (10 mM) (38). *S. lividans* liquid cultures were grown in baffled flasks (YEME) or with marine-grade stainless steel springs (NMMP) to aid in cell dispersion. Strains were cultured 3 to 5 days at 30°C or 42°C. When necessary, antibiotics were used at the following concentrations: apramycin, 50 μg ml^{-1} (YEME, ISP-2); thiostrepton, 10 μg ml^{-1} (YEME), 5 μg ml^{-1} (ISP2, R2YE, NMMP).

Unless noted otherwise, all *E. coli* strains used were derivatives of *E. coli* MG1655. *E. coli* strains were grown at 37°C in lysogeny broth (LB, Difco) (4) or no-carbon essential (NCE) minimal medium (3) supplemented with lithium acetoacetate (30 mM), $MgSO_4$ (1 mM), and ampicillin (100 μg ml^{-1}). When necessary, antibiotics were used at the following concentrations: ampicillin, 100 μg ml^{-1}; apramycin, 50 μg ml^{-1}; chloramphenicol, 12.5 μg ml^{-1}. L-(+)-arabinose was added at a final concentration of 50 μM to induce the expression of the *E. coli* codon-optimized *S. lividans aacS* (EFD70521) from the P_{araBAD} promoter. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 50 μM IPTG to induce expression of *S. lividans patA* (EFD66247).

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 sequencing reactions were resolved at the University of Wisconsin-Madison Biotechnology Center. Oligonucleotide primer sequences are listed in Table 3.3.

Plasmids used for protein overproduction. The *S. lividans patA* (EFD66247), *aacS* (EFD70521), and *acs* (EFD68454) genes were amplified from purified *S. lividans* TK24 genomic DNA with the primers listed in Table 3.3. The first codon of *Sl*Acs (GTG) and *Sl*PatA (TTG) were changed to

Table 3.3. Primers used in this study.

Name	Sequence[4,5]
atoDA deletion primers	
atoDA Wanner 5’	ACCCACAACGGTGTATGCAAGAGGGATAAAAAATGAAAA CGTGTAGGCTGGAGCTGCTTC
atoDA Wanner 3’	CGCGATATGCGACCAATCATAAATCACCCCGTTGCGTATT CATATGAATATCCTCCTTAG
Cloning primers	
atoDA+ EcoRI 5’	GCA**GAATTC**ACGGTGTATGCAAGAGGGAT
atoDA + KpnI 3’	GCC**GGTACC**CCGGCTGACAAAACGCGTC
SlAcs pTEV 5 5’	GTA**GCTAGC**ATGAGCAACGAATCCTTGGCCAAC
SlAcs pTEV5 3’	ACT**GAATTC**GCCTGTGGCCGGCTCAGTC
SlAacS pTEV5 3’	TCG**GAATTC**GCTGCTCAGGAGCGCTTG
SlAacS pBAD30 5’	GGA**GAATTC**AGGAGGACAGCTATGTCGACCGAGAACCCA C
SlAacS pBAD30 3’	TCG**AAGCTT**GCTGCTCAGGAGCGCTTG
SlAacS (opt) 5’	GGA**GAATTC**AGGAGGACAGCTATGTCGACCGAAAACCCG C
SlAacS (opt) 3’	ACT**GGTACC**TCAAGAGCGTTTGCGTGCC
SlAacS 5’ XbaI pSE34	GCC**TCTAGA**CCGGGAGCCGCCCCATGCACCACCACCACCA CCACATGTCGACCGAGAACCCAC
SlAacS 3’ HindIII pSE34	TCG**AAGCTT**GCTGCTCAGGAGCGCTTG
PatA 5’ pTEV5 NheI	GTA**GCTAGC**ATGTCGTACGCGAGCCGTACTCTGG
PatA 3’ pTEV5 EcorI	ACT**GAATTC**GGTCAGTAGGCCGGCAGGGTC
PatA 5’ pSRK-Km NdeI	GCT**CATATG**TCGTACGCGAGCCGTACTC
PatA 3’ pSRK-Km KpnI	ACT**GGTACC**TCAGTAGGCCGGCAGGGTC
SlREVPat 5’ EcoRV IF	CTATGACATGATTACGAATTC**GATATC**GGCCTAGAGTCGC TGAAAGC
SlREVPat R linker BamHI	CGTGGCGGGTCAGTAGGCC**GGATCC**CCTGCTCGTAGAAGC TGACC
SlREVPat F linker BamHI	GGTCAGCTTCTACGAGCAGG**GGATCC**GGCCTACTGACCCG CCACG
SlREVPat 3’ HindIII IF	TGTAAAACGACGGCCAGTGCC**AAGCTT**GTCCTCGTCGAAA TCGTCCGGTG

[4] Bold typeface indicates restriction sites
[5] Ribosome binding sites are underlined

the more common ATG start codon. The DNA fragments were digested with NheI and EcoRI and ligated into pTEV5 (50) cut with the same enzymes. The resulting plasmids p*Sl*PatA1 and p*Sl*AacS1, and p*Sl*Acs1 direct synthesis of *Sl*PatA, *Sl*AacS, and *Sl*Acs, respectively, with *N*-terminal H_6 tags cleavable by recombinant tobacco etch virus (rTEV) protease prepared as described (6). Plasmids directing synthesis of *Sl*AacS^{K617A} and *Sl*Acs^{K610A} variants were generated from the p*Sl*AacS1 and p*Sl*Acs1 plasmids using site-directed mutagenesis

Construction of other plasmids. E. coli *complementation plasmids.* The *S. lividans aacS* (EFD70521) was amplified from purified *S. lividans* TK24 genomic DNA or from the *aacS* allele whose codon usage was optimized for *E. coli* (Genscript) with the primers that included an optimized ribosome-binding site (Table 3.3). The native *S. lividans aacS* DNA fragment was cut with HindIII and EcoRI and ligated into pBAD30, cut with the same enzymes. The *E. coli* codon-optimized *S. lividans aacS* DNA fragment was cut with EcoRI and KpnI and ligated into pBAD30 (25) cut with the same enzymes. The resulting plasmids p*Sl*AacS4 and p*Sl*AacS6 express the native *aacS* or optimized *aacS* genes under the control of the arabinose-inducible P_{araBAD}.

S. lividans patA (EFD66247), was amplified from purified *S. lividans* TK24 genomic DNA with the primers listed in Table 3.3. The *patA* DNA fragment was digested with NdeI and KpnII and ligated into pSRK-Km (35) cut with the same enzymes. The resulting plasmid p*Sl*PatA9 express the *S. lividans patA* gene under the control of the *lacI*^{q}-*lac* promoter-operator system.

S. lividans *H_6*-Sl*AacS plasmid.* The *S. lividans aacS* gene (EFD70521) was amplified from purified *S. lividans* TK24 genomic DNA with the primers listed in Table 3.3. The *S. lividans aacS* DNA fragment was cut with XbaI and HindIII and ligated into pSE34 (67) cut with the same enzymes. The resulting plasmids p*Sl*AacS5 expresses *S. lividans aacS* with an *N*-terminal H_6 tag under the control of the *erm* promoter.

***Construction of gene deletion in* E. coli.** An in-frame deletion of *atoDA* genes in *E. coli* was constructed as using the phage lambda Red recombinase system as previously described (16).

***Construction of gene deletions in* S. lividans.** An in-frame deletion of *S. lividans patA* (EFD66247) was generated using described protocols (43). DNA fragments of 1.5 kb in length were amplified from the DNA upstream and downstream of *S. lividans patA* using purified *S. lividans* TK24 genomic DNA. The resulting fragments were cloned into pKC1139 (5) using the In-Fusion ND Cloning Kit (Clontech) and transformed into *E. coli* Stellar competent cells (Clontech). The resulting plasmid (pKC1139-Δ*patA*) was conjugated into *S. lividans* using the helper strain HB101 harboring pRK2013 (19) on mannitol soya (MS) agar as previously described (37). Plates were flooded with apramycin (50 μg ml^{-1} final concentration) 16 h after plating the conjugation mixtures to select for *S. lividans* carrying pKC1130-Δ*patA*. Apramycin-resistant strains were inoculated into 25 ml YEME + apramycin medium and grown in baffled flasks at 30°C for 4 days. Strains were then plated on ISP-2 + apramycin and incubated at 42°C for 3 days to select for a strain in which the plasmid had integrated into the chromosome. Apramycin-resistant strains that grew at 42°C were inoculated into 25 ml YEME in baffled flasks and incubated at 30°C for 4 days and subsequently plated on ISP-2 at 30°C to promote loss of the integrated plasmid. Isolated colonies were screened on ISP-2 for apramycin resistance. Apramycin-sensitive strains were screened by PCR for deletion of the *patA* gene.

Protein purification. Sl*PatA purification.* Plasmid p*Sl*PatA1 was transformed into *E. coli* strain C41λ(DE3)/ pLysSRARE2 (EMD Millipore). The resulting strains were grown overnight and sub-cultured 1:100 (v/v) into 8 liters of superbroth containing ampicillin (100 μg ml^{-1}) and chloramphenicol (12.5 μg ml^{-1}). The cultures were grown shaking at 37 °C to A_{600} ~ 0.7 and H_6-*Sl*PatA synthesis was induced with IPTG (0.5 mM). Upon induction, the cultures were grown overnight at 30°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 resuspended in 30 ml cold His-Bind buffer (buffer A) [tris(hydroxymethyl)aminomethane-HCl (Tris-HCl) buffer (50 mM, pH 8), NaCl (500 mM)], and imidazole (5 mM)) containing phenylmethanesulfonylfluoride (PMSF, 1 mM). Cells were placed on ice and lysed by sonication for 2 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*. Clarified cell extract was loaded onto a 1 ml HisTrap FF column (GE Healthcare) connected to a computer-controlled ÄKTA fast protein liquid chromatography (FPLC) system. Unbound proteins were eluted off the column by extensive washing with buffer A. A 10-ml wash step with 90% buffer A and 10% buffer B [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (250 mM)] was applied to the column a prior to a 10-ml linear gradient 10-100% Buffer B. All fractions containing H_6-*Sl*PatA were combined. rTEV protease was added to H_6-*Sl*PatA and the *Sl*PatA/rTEV mixture was incubated at room temperature for 3 h. PMSF was added to the protein mixture and incubated 15 min at room temperature. The *Sl*PatA/rTEV mixture was dialyzed at 4°C against buffer C (Tris-HCl (50 mM, pH 8), NaCl (500 mM)) twice for 3 hours 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. Cleaved *Sl*PatA was desorbed from the resin using a 50 ml wash step with 97% buffer A and 3% Buffer B prior to a 25-ml linear gradient 3-100% Buffer B. Purified *Sl*PatA was analyzed by SDS-PAGE. Fractions containing *Sl*PatA were pooled together. *Sl*PatA was stored in Tris-Cl buffer (50 mM, pH 8.0) containing NaCl (100 mM) and glycerol (20%, v/v). *Sl*PatA concentration was determined by measuring absorbance at 280 nm. The molar extinction coefficient used to calculate *Sl*PatA concentration was 57,760 $M^{-1}cm^{-1}$.

Sl*Acs and* Sl*AacS purification*. Plasmids containing *S. lividans acs* or *aacS* were transformed with pRARE2 (EMD Millipore) into a Δ*pat* derivative of C41λ(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 LB containing ampicillin (100 μg ml^{-1}) and chloramphenicol (12.5 μg ml^{-1}). The cultures were grown shaking at 37°C to A_{600} ~ 0.7 and protein synthesis was induced with IPTG (0.25 mM). Upon induction, the cultures were grown overnight at 30°C. *Sl*Acs and *Sl*AacS proteins were purified and stored as described above with modifications. During the first purification step, a 10-ml wash step with 94% buffer A and 6% buffer B was applied to the column a prior to a 10-ml linear gradient 6-100% buffer B. *Sl*Acs and *Sl*AacS proteins did not adsorb to the column and were present in the flow-through fractions.

The molar extinction coefficient used to calculate protein concentrations were 135,455 cm^{-1} M^{-1} for *Sl*Acs and 142,320 cm^{-1} M^{-1} for *Sl*AacS.

Se*Acs purification.* Plasmid pACS10 was transformed into a Δ*pat* derivative of C41λ(DE3) (JE9314). The resulting strain was grown overnight and sub-cultured 1:100 (v/v) into 2 liters of LB containing ampicillin (100 μg ml^{-1}). The culture was grown shaking at 37 °C to A_{600} ~ 0.7 and protein synthesis was induced with IPTG (0.25 mM). Upon induction, the cultures were grown overnight at 30°C. *Se*Acs was purified and stored as described (58).

In vitro *acyl-CoA synthetase assays.* *Sl*AacS activity was measured using an NADH-consuming assay (14, 20, 22, 30). *R*eactions (100 μl total volume) contained HEPES buffer (50 mM, pH 7.5), TCEP (1 mM), ATP (2.5 mM) CoA (0.5 mM), $MgCl_2$ (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 either acetoacetate, β-hydroxybutyrate, butyrate, isovalarate, malonate, crotonate, isobutyrate, succinate, propionate, or acetate (0.2 mM). Reactions were started by the addition of *Sl*AacS (15 nM). The absorbance at 340 nm was monitored in a 96-well plate using the Spectramax Plus UV-visible spectrophotometer (Molecular Devices). Enzyme activities were calculated as described (22). Specific activity data are presented with standard deviations from triplicate experiments.

HPLC separation and mass spectral analysis of* Sl*AacS reaction product. To determine the molecular mass of the *Sl*AacS product, *Sl*AacS

was incubated in the presence of CoA, ATP, and acetoacetate, described above, in a final volume of 1 mL. *Sl*AacS was removed from reactions by filtration using Amicon Ultracel centrifugal filters (3-KDa molecular mass size exclusion). Filtrates containing reaction products and reagents were separated by analytical by reverse-phase ion-pair HPLC using a Beckman Coulter System Gold 126 system equipped with a Phenomenex Synergi Hydro-RP (150 by 4.5 mm, 5 micron particle size, 2.5 ml column volume) column at a flow rate of 1 mL min^{-1}. The solvent system was previously described (27, 33) Briefly, the column was equilibrated with buffer A (100 mM KH2P04, 2 mM tetra-butyl ammonium

bromide (TBAB), 15% CH_3CN, pH 3.3). Sample was injected onto the column and washed with buffer A for 5 min. Acetoacetyl-CoA was eluted by a 10-min linear gradient from buffer A to buffer B (100 mM KH2P04, 2 mM TBAB, 35% CH_3CN, pH 3.3). TCEP was added to each sample to a final concentration of 2 mM before injection. Compounds eluted with the following retention times: CoA, 4.7 min; ATP, 2.6 min; AMP, 2.6 min, Acetoacetyl-CoA, 7.3 min. The sample containing acetoacetyl-CoA that eluted from 7 to 9 min was dried under vacuum using an Eppendorf Vacufuge plus concentrator operating at room temperature. The dried sample was resuspended in dH_2O and applied onto a C_{18} Sep-Pak (Waters) previously conditioned with 10 mL of 100% methanol followed by 10 mL of dH_2O. Acetoacetyl-CoA that bound to the column was washed with 10 mL of dH_2O before elution with 100% methanol. Acetoacetyl-CoA dissolved in methanol was concentrated under vacuum. The pellet was resuspended in dH_2O form mass determination by mass spectrometry. The sample as analyzed on a Bruker Autoflex Matrix Assisted Laser Desorption/Ionization-Time-of-Flight mass spectrometer scanning 80-2275 m/z using 300 shots of nitrogen laser in delayed extraction and positive reflectron mode summed together to form the mass spectrum. Matrix for the analysis was dihydroxybenzoate (15 mg/ml, Sigma) in a 1:1 solution of water: 0.1% TFA, CH_3CN (v/v).

In vitro *protein acetylation assay*. Protein acetylation was observed using radiolabeled Ac-CoA as described (14, 59, 65). Acetylation reactions contained 2(Bis(2-hydroxyethyl)imino)-2-(hydroxymethyl)-1,3-propanediol (Bis-Tris-HCl) buffer (50 mM, pH 6.0), [1-^{14}C]-Ac-CoA (20 μM), acyl-CoA synthetase (3 μM), glycerol (10%, v/v), and *Sl*PatA (1 μM). Reactions (10 μl total volume) were incubated for min at 30°C. Samples (5 μl) were resolved using SDS-PAGE (40) and proteins were visualize by Coomassie Blue staining. Gels were dried and exposed 16 h to a multipurpose phosphor screen (Packard). Labeled proteins were visualized using a Typhoon FLA 9000 Variable Mode Imager (GE Healthcare) equipped with ImageQuant TL software (GE Healthcare).

The effect of acetylation on activity of *Sl*AacS, *Sl*Acs, and *Se*Acs activity was determined as described (14) with modifications. *Sl*AacS (3μM) was incubated with *Sl*PatA (1 μM) and 50 μM Ac-CoA

for 90 minutes at 30°C using the buffer system described above. At 0, 15, 30, 60, and 90 min time points, reactions were diluted 1:20 into 50 mM HEPES buffer pH 7.5 at 4°C. *Sl*AacS, *Sl*Acs, and *Se*Acs activity were measured as described above using the appropriate organic acid substrate.

In vitro *deacetylation assays*. Acetylated *Sl*AacS was deacetylated with *Salmonella enterica* $CobB_S$ as described (65). *In vitro* acetylated *Sl*AacS (3 μM, radiolabeled or non-radiolabeled) was incubated with *Se*$CobB_S$ (3 μM) in deacetylation buffer containing 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 50 mM, pH 7.0), NAD^+ (1 mM) for 60 min at 37°C (25 μl reaction volume). Reaction mixture samples (5 μl) were resolved by SDS-PAGE, and subjected to phosphor imaging analysis to assess the acetylation state of *Sl*AacS after incubation with *Se*$CobB_S$. *Sl*AacS activity of the non-radiolabeled reactions was measured using the CoA synthetase assay described above.

***Purification of* Sl*AacS from* S. lividans.** Plasmid p*Sl*AacS5 was introduced into *S. lividans* TK24 and *S. lividans* Δ*patA* (JE16707) by polyethyleneglycol (PEG)-assisted protoplast transformation as described (36, 37) to generate JE16731 and JE16821 (Table 3.1), respectively. Cells were plated on R2YE and grown at 30°C for 16 hours. Plates were flooded with thiostrepton (10 μg ml^{-1} final concentration) and incubated 3 days at 30°C to select for strains harboring p*Sl*AacS5. *S. lividans* strains harboring the p*Sl*AacS5 plasmid encoding H_6-*Sl*AacS were grown in 30 ml YEME + thiostrepton for 5 days. Cells were harvested by centrifugation at 2000 x *g* for 10 min at 4°C. Cell mass was measured and cells were resuspended in an equal volume of NMMP. Approximately 0.1 g of cells were inoculated into 250 ml NMMP + thiostrepton supplemented with acetoacetate (10 mM). Strains were grown for 24 h at 30°C with shaking. Cells were harvested as described above and washed twice with 50 ml wash buffer containing Tris-HCl buffer (50 mM, pH 8.0) and NaCl (500 mM). Cells pellets were resuspended in 30 ml buffer A (described above) supplemented with PMSF (1 mM), Sigma protease inhibitor cocktail for histidine-tagged proteins (100 μl), and lysozyme (1 mg ml^{-1}). Cells were lysed by sonication and cell debris was removed by centrifugation. H_6-*Sl*AacS was purified using 250 μl HisPur Ni-NTA resin

(Pierce). Using the same buffer system described for *Sl*AacS above. *Sl*AacS-containing fractions were combined and dialyzed overnight into HEPES buffer (50 mM, pH 7.5) containing NaCl (150 mM) and glycerol (20%, v/v).

*Sl*AacS of SlAacS activity was assessed using the deacetylation reaction described above. *Sl*AacS (1 μM) purified from JE16731 and JE16821 (Table 3.1) were incubated with or without *Se*CobB$_S$ as described in the above section for 1 h at 37°C. A sample for the deacetylation reaction mixtures was used to assess *Sl*AacS activity. The concentration of *SlAacS* in the CoA synthetase activity assay was 25 nM.

***Western blot analysis of H_6-*SlAacS* purified from* S. lividans.** Purified *Sl*AacS or in vitro acetylated *Sl*AacS were resolved along with H_6-*Sl*AacS purified from *S. lividans* strains JE16731 and JE16821 (Table 3.1) (2 ug each) using SDS-PAGE and transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore). Total protein was visualized by staining 15 seconds in Ponceau-S staining solution (0.1% (w/v) Ponceau S in 5% (v/v) acetic acid) and de-staining with distilled water. After extensive de-staining, the PVDF membranes were probed with polyclonal rabbit α-acetylated lysine antibodies (1:1,700; Calbiochem) as primary antibody and goat α-rabbit immunoglobulin G conjugated to calf intestinal alkaline phosphatase (Pierce) (1:10,000) as secondary antibody to detect *Sl*AacS^{Ac}. Signal was detected using nitro-blue tetrazolium chloride and 5-bromo-4-chloro-3′-indolylphosphate *p*-toluidine salt (NBT-BCIP) 1-Step Solution according to the manufacturer's instructions (Pierce).

Determination of the* Sl*AacS acetylation site by mass spectrometry. To determine the identity of the *in vivo* modification of *Sl*AacS, H_6-*Sl*AacS purified from *S. lividans* was resolved by SDS-PAGE and the band corresponding to *Sl*AacS was excised. "In Gel" digestion and mass spectrometric analysis was done at the Mass Spectrometry Facility [Biotechnology Center, University of Wisconsin-Madison]. The digestion was performed as outlined on the website:

http://www.biotech.wisc.edu/facilities/massspec/protocols/ingelprotocol.

Peptides were analyzed by nanoLC-MS/MS using the Agilent 1100 nanoflow system (Agilent, Palo Alto, CA) connected to a hybrid linear ion trap-orbitrap mass spectrometer (LTQ-Orbitrap, Thermo

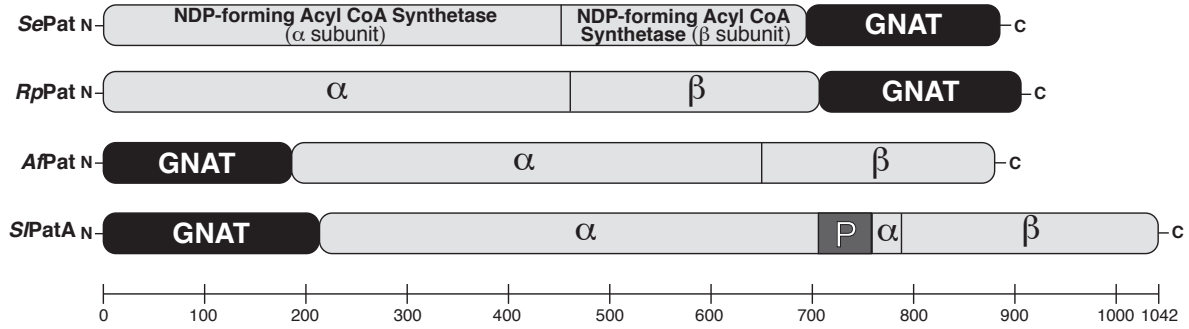
Fisher Scientific, Bremen, Germany) equipped with a nanoelectrospray ion source (Proxeon Biosystems, Odense, Denmark). Chromatography of peptides prior to mass spectral analysis was accomplished using C18 reverse phase HPLC trap column (Zorbax 300SB-C18, 5µM, 5x0.3mm, Agilent) and capillary emitter column (in-house packed with MAGIC C18, 3 µM, 15x0.075mm, Michrom Bioresources, Inc.) onto which 8µl of extracted peptides were automatically loaded. NanoHPLC system delivered solvents A: 0.1% (v/v) formic acid in water, and B: 95% (v/v) acetonitrile, 0.1% (v/v) formic acid at either 10 µL/min, to load sample, or 0.20 µL/min, to elute peptides directly into the nano-electrospray over a 60 minutes 1% (v/v) B to 60% (v/v) B followed by 10 minute 60% (v/v) B to 100% (v/v) B gradient. As peptides eluted from the HPLC-column/electrospray source survey MS scans were acquired in the orbitrap with a resolution of 100 000 and up to 5 most intense peptides per scan were fragmented and detected in the ion trap over the 400 to 2000 m/z; redundancy was limited by dynamic exclusion. Raw MS/MS data were converted to mgf file format using Trans Proteomic Pipeline (Seattle Proteome Center, Seattle, WA). Resulting mgf files were used to search user defined amino acid sequence database using in-house Mascot search engine 2.2.07 (Matrix Science, London, UK) with Cysteine carbamidomethylation as fixed modification and Lysine acetylation, methionine oxidation, and Asparigine/Glutamine deamidation as variable modifications. Peptide mass tolerance was set at 20 ppm and fragment mass at 0.8 Da. Protein annotations and significance of identification was done with help of Scaffold software (version 3.6.1, Proteome Software Inc., Portland, OR). Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm (34). Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (47). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

RESULTS

***Actinomycetes and the archaeon* Archaeoglobus fulgidus *encode a protein acetyltransferase with unique domain organization*.** *Salmonella enterica* and *Rhodopseudomonas palustris* synthesize a protein acetytransferase Pat enzyme with a large (~700 residues) *N*-terminal domain with homology to NDP-forming CoA synthetases and a C-terminal (~200 residues) GNAT domain (14, 59). The EFD66247 locus of the genome of the actinomycete *Streptomyces lividans* encodes a homologue of the *Se*Pat and *Rp*Pat enzymes in which the GNAT and the NDP-forming CoA synthetase-like domains are reversed (Fig. 3.1A). We note that the DNA sequence of *S. lividans* open reading frame (ORF) EFD66247 currently available in databases contains two mistakes, which we discovered upon cloning and sequencing EFD66247. Our sequencing data showed two changes, a transition (T1787C) and a deletion (ΔT1788). BLASTp searches performed with the protein encoded by the corrected DNA sequence identified a full-length NDP-forming CoA synthetase-like domain homologous to the ones found in *Se*Pat and *Rp*Pat. Further bioinformatics analysis of the corrected sequence of ORF EFD66247 revealed that such reverse domain organization of the putative protein was also present in homologues presumably synthesized by other actinomycetes and the archaeon *Archaeoglobus fulgidus*. The predicted primary sequence of the EFD66247 protein contained a proline-rich sequence (P663-P753, 26.4% proline) that included a degenerate G-P-S motif (Fig. 3.1B) (2).

***The EFD66247 protein (hereafter* Sl*PatA) is a functional protein acetyltransferase*.** *Se*Pat and *Rp*Pat have been shown to acetylate the acetyl-CoA synthetase (Acs) enzymes from *S. enterica* and *R. palustris*, respectively (14, 59). *Se*Pat and *Rp*Pat acetylate the epsilon amino group of an active-site lysyl side chain, with the concomitant inactivation of the adenylylation activity of Acs (14, 59). Mikulik et al. recently reported that the Acs homologue from *Streptomyces ceolicolor* (*Sc*Acs) was acetylated *in vivo* (45), suggesting that a protein acetyltransferase was responsible for acetylating Acs in this *Streptomyces* species. We asked whether recombinant *Sl*PatA could acetylate *Sl*Acs *in vitro*. For this purpose, *Sl*PatA was purified to homogeneity and incubated with purified *Sl*Acs (encoded by *S. lividans* locus EFD68454) in the presence of [1-^{14}C]-Ac-CoA. As shown in figure 3.2A, *Sl*PatA acetylated *Sl*Acs, indicating that

A



B *Sl*PatA G-P-S motif

703- LPS **G**AS **G**PS **G**PA DQLPS **G**PS **G**PS **G**PA **G**PS **G**PS **G**PP RPS -740

Figure 3.1. Domain organization of Pat homologues. (A) Pat homologues encode a GNAT domain (black) and large domain that is homologous to NDP-forming acyl-CoA synthetases (light gray). *S. lividans* PatA also contains a proline-rich domain in the large domain (dark gray, denoted by "P"). *Se*Pat, *Salmonella enterica* Pat (NP_461586); *Rp*Pat, *Rhodopseudomonas palustris* Pat (NP_949576); *Af*Pat *Archaeoglobus fulgidus* Pat (NP_070340); *Sl*PatA, *Streptomyces lividans* PatA (ZP_06527997). (B) The degenerate G-P-S motif Leu703-Ser740 in *Sl*PatA.

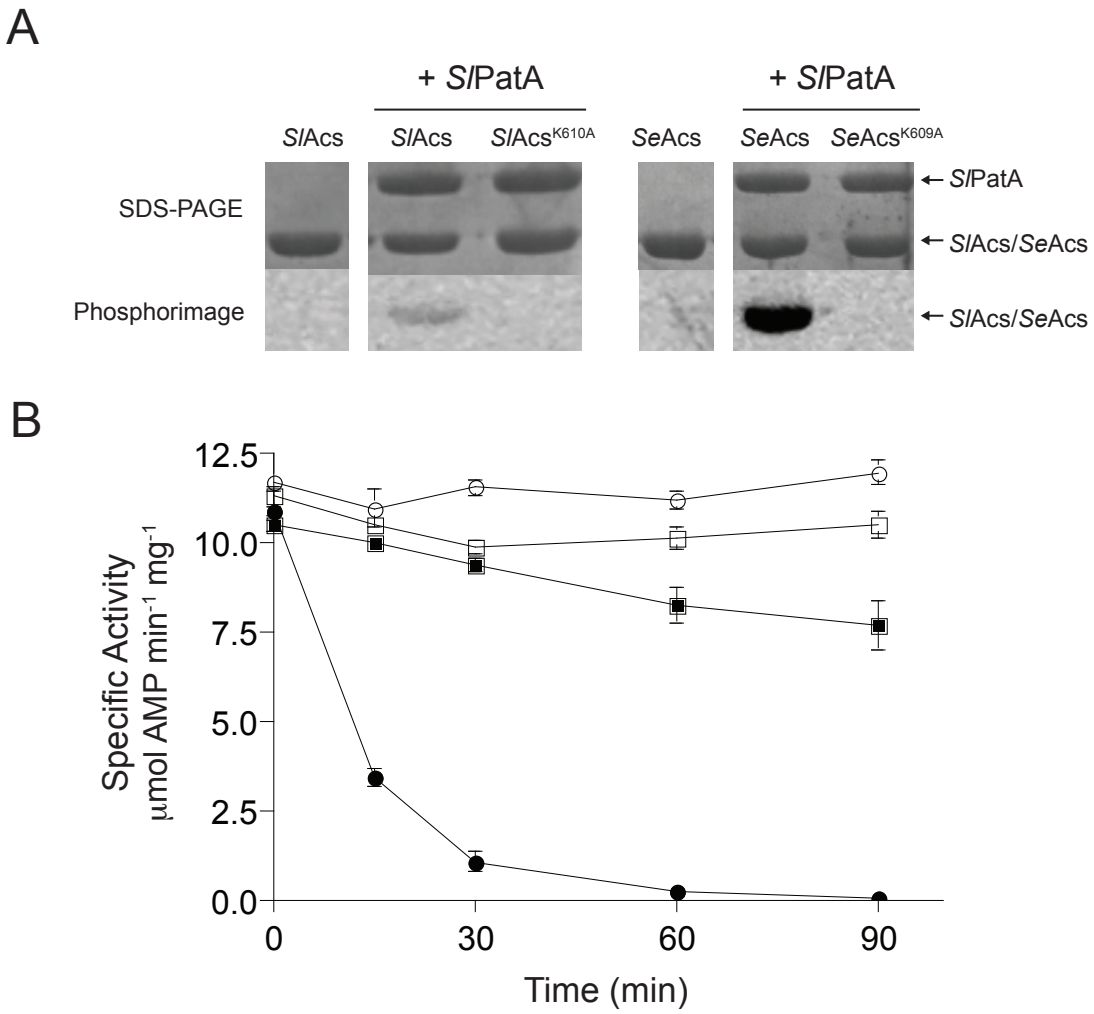


Figure 3.2. *Sl*PatA acetylates *Sl*Acs from *S. lividans*. (A) *Sl*Acs, *Sl*Acs^{K610A}, *Se*Acs, or *Se*Acs^{K609A} was incubated with [1-^{14}C]-acetyl-CoA in the presence or absence of *Sl*PatA. Proteins were separated by SDS-PAGE and stained with Coomassie Blue to visualize proteins. Acetylation was visualized by phosphor imaging. *Sl*PatA was incubated with *Sl*Acs or *Se*Ac in the presence or absence of unlabeled acetyl-CoA (B). Reactions were carried out with 1:3 molar ratios of *Sl*PatA to *Sl*Acs/*Se*Acs. Sample were diluted and assayed to measure *Sl*Acs/*Se*Acs activity at 0, 15, 30, 60, and 90 min after incubation with *Sl*PatA. *Sl*Acs (squares) and *Se*Acs (circles) activities were measured in and NADH-consumption assay and activities are reported for the reactions containing acetyl-CoA (closed symbols) and the control reactions lacking acetyl-CoA (open symbols). Reactions were carried out in triplicate. Error bars represent standard deviations.

*Sl*PatA was a bona fide protein lysine acetyltransferase. To verify the site of acetylation, a *Sl*Acs^{K610A} variant was isolated and tested as a substrate of *Sl*PatA. *Sl*PatA did not modify *Sl*Acs^{K610A}, indicating that the active-site residue K610 was the only site modified by *Sl*PatA. For reference, *Sl*PatA was incubated with *Se*Acs, a known substrate of *Se*Pat (59), and the active site lysine variant *Se*Acs^{K609A}. *Sl*PatA modified *Se*Acs but not *Se*Acs^{K609A} indicating that *Sl*PatA acetylates only the active-site residue K609 of *Se*Acs. Relative acetylation of *Sl*Acs and *Se*Acs was quantified using digital light units. *Sl*PatA modified *Se*Acs approximately 30-fold more efficiently than *Sl*Acs.

To compare the effect of *Sl*PatA-mediated acetylation on *Sl*Acs and *Se*Acs activity, *Sl*Acs or *Se*Acs was incubated with *Sl*PatA in the presence or absence of acetyl-CoA. When *Sl*Acs was incubated with *Sl*PatA in the presence of Ac-CoA, *Sl*Acs retained approximately 75% of its activity relative to the no acetyl-CoA control (Fig. 3.2B). This result suggested that *Sl*PatA acetylation of *Sl*Acs was inefficient. *Sl*PatA inactivated *Se*Acs within 90 minutes of incubation with Ac-CoA, confirming that *Se*Acs was a better substrate for *Sl*PatA than *Sl*Acs under the conditions tested. Additionally, these data confirm that *Sl*PatA acted catalytically. As a point of reference, under the conditions used *Se*Pat fully inactivated *Se*Acs (59).

Sl*PatA acetylates an acetoacetyl-CoA synthetase homologue from* S. lividans. Because *Sl*PatA acetylation of *Sl*Acs was inefficient, we looked for additional *Sl*PatA substrates in *S. lividans*. *Se*Pat homologues have been shown to acylate members of the AMP-forming acyl-CoA synthetase family of enzymes (14, 15, 22, 59). The closest homologue of *Sl*Acs we found in *S. lividans* was a putative acetoacetyl-CoA synthetase (hereafter *Sl*AacS, encoded by locus EFD70521), whose primary sequence was 39% identical to that of the acetoacetyl-CoA synthetase from rat liver (30). *Sl*AacS homologues contained the catalytic residue Lys617 (Table 3.4), which raised the possibility that *Sl*AacS could be acetylated. To test this possibility, *Sl*PatA and *Sl*AacS were incubated in the presence of [1-^{14}C]-Ac-CoA. As shown in figure 3.3, *Sl*PatA efficiently used [1-^{14}C]-Ac-CoA to modify *Sl*AacS, and residue K617 was the only site of acetylation in *Sl*AacS. Under the same assay conditions, *Sl*PatA acetylated 15-fold more *Sl*AacS than *Sl*Acs, indicating that *Sl*AacS was a substantially better substrate than *Sl*Acs for *Sl*PatA.

Table 3.4. Comparison of *Sl*AacS active site to acetylation substrates and *Sl*AacS homologues.

Species	Gene Name/ Locus Tag	Accession Number	Percent amino acid identity to *Sl*AacS	Active site motif
S. lividans	*aacS*	ZP_06532271	N/A	610-IPHTLTG**K**RIEVPVK-624
S. lividans	*acs*	ZP_06530204	N/A	603-LPKTRSG**K**IMRRLLR-617
S. enterica	*acs*	NP_463140	N/A	602-LPKTRSG**K**IMRRILR-616
S. enterica	*prpE*	NP_459366	N/A	585-LPKTRSG**K**MLRRTIQ-599
R. palustris	*badA*	NP_946014	N/A	505-LPKTATG**K**IQRFKLR-519
R. palustris	*aliA*	NP_946004	N/A	525-MPATPSG**K**IQKFRLR-539
R. palustris	*prpE*	NP_949838	N/A	591-LPKTRSG**K**ILRGTIK-605
R. palustris	*acs*	NP_945564	N/A	599-LPKTRSG**K**IMRRILR-613
Pseudomonas aeruginosa PAO1	PA1997	NP_250687	46	606-IPRTLSG**K**IVELAVR-620
Archaeoglobus fulgidus	*acs-1*	NP_069035	44	608-IPMTLNY**K**KLEVPIK-622
Caenorhabditis elegans	SUR-5	NP_509229	37	658-IPYTSSG**K**KVEVAVK-672
Rattus norvegicus	AACS	NP_075592	39	370-IPYTING**K**KVEVAVK-384
Homo sapiens	AACS, SUR5	NP_076417	38	626-IPYTLNG**K**KVEVAVK-640

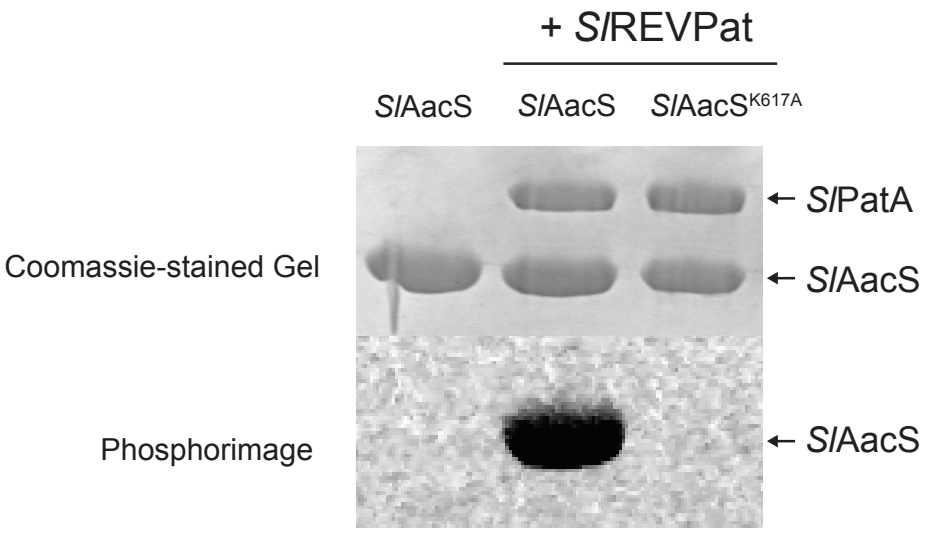


Figure 3.3. *Sl*PatA acetylates *Sl*AacS from *S. lividans* at the conserved active site lysine. Alignment of *Sl*AacS with known acetylation targets from *S. enterica* (*Se*Acs, encoded by locus STM4275; *Se*PrpE, encoded by locus STM0371), *R. palustris* (*Rp*BadA, encoded by locus RPA0661; *Rp*AliA, encoded by locus RPA0651; *Rp*HbaA, encoded by locus RPA0669; *Rp*PrpE, encoded by locus RPA4504; *Rp*Acs, encoded by locus RPA0211), and *S. lividans* indicates a conserved active site lysine (A). Acetylation of *Sl*AacS and *Sl*AacSK617A was assessed after incubation of proteins with [1-^{14}C]-acetyl-CoA in the presence or absence of *Sl*PatA. Proteins were separated by SDS-PAGE and stained with Coomassie Blue to visualize proteins. Acetylation was visualized by phosphor imaging (B).

Sl*AacS is a functional acetoacetyl-CoA synthetase* in vivo. We used strains of *Escherichia coli* to assess *in vivo* the ability of *Sl*AacS to convert acetoacetate to acetoacetyl-CoA (Fig. 3.4A). *E. coli* encodes a protein acetyltransferase (Pka) homologous to *Se*Pat and sirtuin-type protein deacetylase (CobB). Growth experiments were done in a Δ*pka* Δ*cobB E. coli* strain background to avoid potential *Sl*AacS modification by the *E. coli* acetylation machinery. *E. coli* grows on acetoacetate by first converting it to acetoacetyl-CoA using the acetyl-CoA:acetotacetate CoA transferase enzyme (AtoDA, EC 2.8.3.8). In the second step of the pathway, the 3-ketoacyl-CoA thiolase (AtoB, EC 2.3.1.16) enzyme converts acetoacetyl-CoA and CoA to two molecules of acetyl-CoA, which enter central metabolism (Fig. 3.4A) (48). *E. coli* encodes a second 3-ketoacyl-CoA thiolase, FadA, with a substrate preference for mid-length chain fatty acids that is capable of using acetoacetate as a substrate to generate two molecules of Ac-CoA (57). An *E. coli atoDA* strain cannot grow on acetoacetate because it cannot convert acetoacetate to acetoacetyl-CoA (48). An *E. coli atoB fadA* strain cannot grow on acetoacetate due to the lack of acetoacetyl-CoA thiolase activity (31). As shown in figure 3.4B, when *Sl*AacS encoded by the native *S. lividans aacS* allele was ectopically produced in *E. coli*, growth of the Δ*atoDA* Δ*cobB* Δ*pka* strain on acetoacetate was restored. The final culture density was comparable to that reached when wild type alleles of *atoDA* were provided in *trans*. These data suggest that *Sl*AacS converted acetoacetate to acetoacetyl-CoA compensating for the absence of AtoDA in the *atoDA* strain. The difference in the growth rate of the strain complemented with native *S. lividans aacS* alleles may be due to codon bias of the *S. lividans* allele, differences in enzyme kinetic parameters, or enzyme stability in *E. coli* at 37°C.

To determine the substrate specificity for *Sl*AacS, the AMP-forming CoA synthetase activity of *Sl*AacS was measured using a continuous spectrophotometric assay with CoA, ATP, and organic acids structurally similar to, and including, acetoacetate (see *Experimental procedures* for all substrates tested). *Sl*AacS efficiently activated acetoacetate to acetoacetyl-CoA [12 ± 2 μmol AMP min^{-1} mg^{-1}) and β-hydroxybutyrate to β-hydroxybutyryl-CoA (2 ± 0.2 μmol AMP min^{-1} mg^{-1}). *Sl*AacS specific activities for all other compounds tested were <0.1 μmol AMP min^{-1} mg^{-1}. Clearly, acetoacetate was the preferred

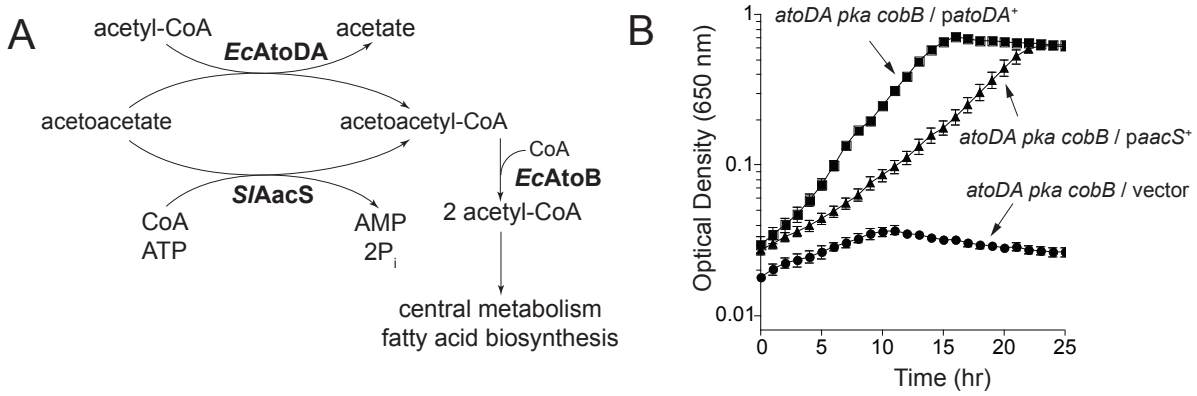


Figure 3.4. *Sl*AacS can substitute for *EcA*toDA in *E. coli* during growth on acetoacetate. Acetoacetate utilization in *E. coli* involves a 2-step conversion from acetoacetate to acetoacetyl CoA and from acetoacetyl-CoA to 2 molecules of acetyl-CoA. *Sl*AacS is predicted by homology to catalyze the conversion of acetatoacetate to acetoacetyl-CoA (A). Growth behavior of *S. enterica* on NCE minimal medium supplemented with acetoacetate (30 mM). Growth experiments were performed at 37°C using a microtiter plate and a microtiter plate reader (Bio-Tek Instruments). Growth experiments were performed in triplicate. Error bars represent standard deviations.

substrate of *Sl*AacS, indicating that *Sl*AacS was an acetoacetyl-CoA synthetase (EC 6.2.1.16). To confirm the identity of the AacS reaction product, the substrates ATP, CoA, and acetoacetate were incubated in the presence or absence of *Sl*AacS. High performance liquid chromatography (HPLC) was used to separate products from reagents. The enzyme-containing reaction produced a single unique peak with a retention time of 7.25 min. Mass spectral analysis (positive mode) of this sample revealed a molecular ion with a *m/z* of 852.3, corresponding to the expected mass of acetoacetyl-CoA (851.6 amu).

***Lys617 is critical for* Sl*AacS activity and acetylation of Lys617 inactivates* Sl*AacS*.** In *Sl*AacS, Lys617 is a conserved active site residue, and the site acetylated by *Sl*PatA (Fig. 3.3A). To assess the role of Lys617 in *SlAacS* activity, site-directed mutagenesis was used to generate *aacS* alleles encoding *Sl*AacS^{K617A} and *Sl*AacS^{K617Q} variants. Alanine served as a catalytically inert substitution and glutamine served as a structural mimic for acetyl-lysine. Relative to the activity of wild-type *Sl*AacS (9.6 ± 0.4 μmol AMP min^{-1} mg^{-1}), both *Sl*AacS^{K617A} and *Sl*AacS^{K617Q} variants were ~20-fold less active (<0.5 μmol AMP min^{-1} mg^{-1}) indicating that Lys617 was critical for activity, lending support to the hypothesis that acetylation of Lys617 inactivated the enzyme.

To test the effect of *Sl*AacS Lys617 acetylation directly, *Sl*AacS was incubated with *Sl*PatA in the presence or absence of acetyl-CoA. In the presence of both acetyl-CoA and *Sl*PatA, *Sl*AacS activity was reduced >97% indicating that *Sl*PatA acetylation of Lys617 effectively decreases *Sl*AacS activity (Fig. 3.5).

Sl*AacS is deacetylated and reactivated by sirtuin deacetylase.* Acetylation of CoA-ligases is reversed by deacetylases (14, 15, 22, 58). The *S. lividans* genome encodes three putative deacetylases, namely two NAD$^+$-dependent sirtuin-type protein deacetylases [EFD65580 (*Sc*CobB2), EFD71509 (*Sc*CobB1)], and one homologue of the *R. palustris* zinc-dependent protein deacetylase *Rp*LdaA (EFD68590). In *S. coelicolor*, *Sc*CobB1 (100% identity to EFD71509) deacetylates *Sc*Acs *in vitro* (45). For unknown reasons, our attempts to purify or enrich for active *S. lividans* deacetylases in *E. coli* extracts were unsuccessful. To circumvent this problem, we used the *Salmonella enterica* CobB (*Se*CobB) sirtuin deacetylase enzyme (58, 65), which has been shown to deacetylate heterologous acyl-

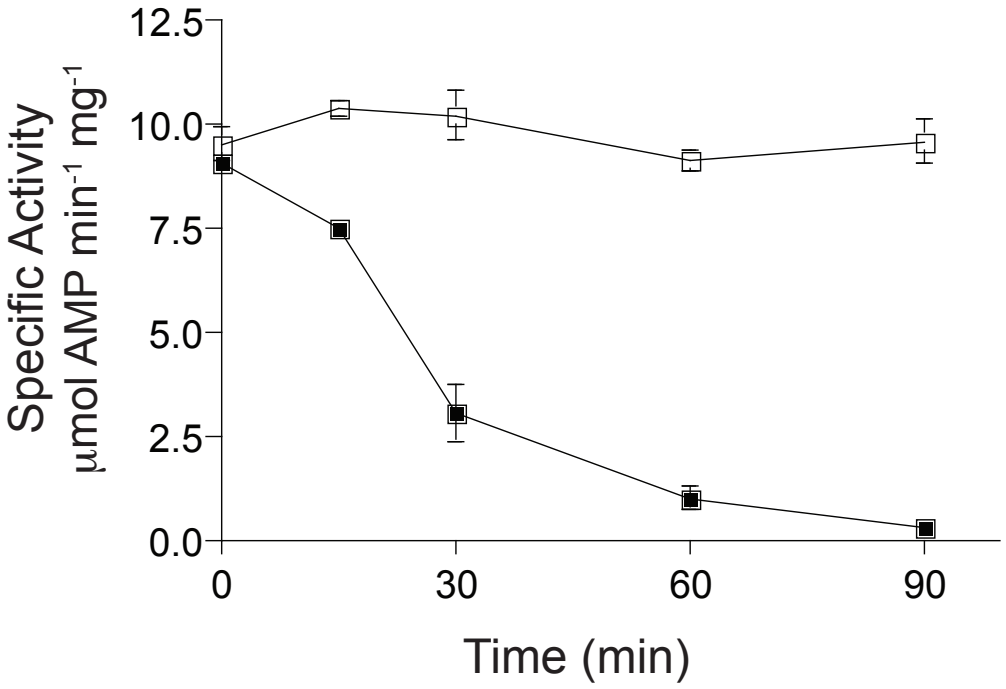


Figure 3.5. *Sl*PatA acetylation inactivates *Sl*AacS. *Sl*AacS was incubated with *Sl*PatA at a 3:1 molar ratio (*Sl*AacS:*Sl*PatA) in the presence (closed squares) or absence (open squares) of acetyl-CoA. Sample were removed, diluted, and assayed to measure *Sl*Acs/*Se*Acs activity at 0, 15, 30, 60, and 90 min after incubation with *Sl*PatA. *Sl*AacS activity was measured in an NADH-consumption assay. Reactions were carried out in triplicate. Error bars represent standard deviations.

and aryl-CoA synthetases (14). We used the short form of *Se*CobB ($SeCobB_S$) sirtuin (65) to demonstrate the reversibility of the acetylation of *Sl*AacS Lys617. *Sl*AacS was acetylated with *Sl*PatA and [^{14}C-1]-Ac-CoA, followed by incubation with *Se*$CobB_S$ in the absence and presence of NAD^+. In the presence of *Se*$CobB_S$ and NAD^+, the amount of radioactivity associated with *Sl*AacS decreased below the limit of detection indicating that *Se*$CobB_S$ deacetylated *Sl*AacS (Fig. 3.6A), demonstrating that *Sl*AacS acetylation was reversible.

We assessed whether deacetylation of *Sl*$AacS^{Ac}$ by *Se*$CobB_s$ would restore *Sl*AacS activity. To do this, we incubated *Sl*PatA with *Sl*AacS and acetyl-CoA. After incubation, buffer was exchanged to remove excess acetyl-CoA. *Se*$CobB_s$ and NAD^+ were added to the reaction mixture prior to a second incubation period, after which *Sl*AacS activity was restored to levels comparable to the unacetylated *Sl*AacS control (Fig. 3.6B). Specific activity of unacetylated *Sl*AacS was reduced when compared to the specific activity measurements reported for the substrate specificity determination. This discrepancy was likely due to instability of the enzyme during the acetylation reaction, buffer exchange, and deacetylation reaction.

Sl*AacS is acetylated* in vivo *during growth of* S. lividans *in the presence of acetoacetate*. We assessed whether or not *Sl*AacS was acetylated *in vivo* in *S. lividans* and whether *Sl*PatA catalyzed the modification of *Sl*AacS. H_6-*Sl*AacS was purified from *S. lividans patA*$^+$ and Δ*patA* strains grown in minimal medium supplemented with acetoacetate. We then assessed the acetylation state of *Sl*AacS by two methods. First, we isolated H_6-*Sl*AacS enzymes from each strain and determined whether or not they were acetylated. To do this, we used Western blot analysis with polyclonal anti-acetyllysine antibodies. Results of control experiments with non-acetylated and *in-vitro* acetylated *Sl*AacS (Fig. 3.7A) showed that acetyllysine was readily detected under the conditions used. A strong signal was detected for acetyllysine in *Sl*AacS isolated from *S. lividans patA*$^+$, whilst the level of acetyllysine in *Sl*AacS isolated from the *S. lividans* Δ*patA* strain was not significantly above background (Fig. 3.7A). Based on these data we inferred that *Sl*AacS was acetylated *in vivo* and that *Sl*PatA was the only protein acetyltransferase that modified *Sl*AacS under the conditions tested. To confirm the modification of

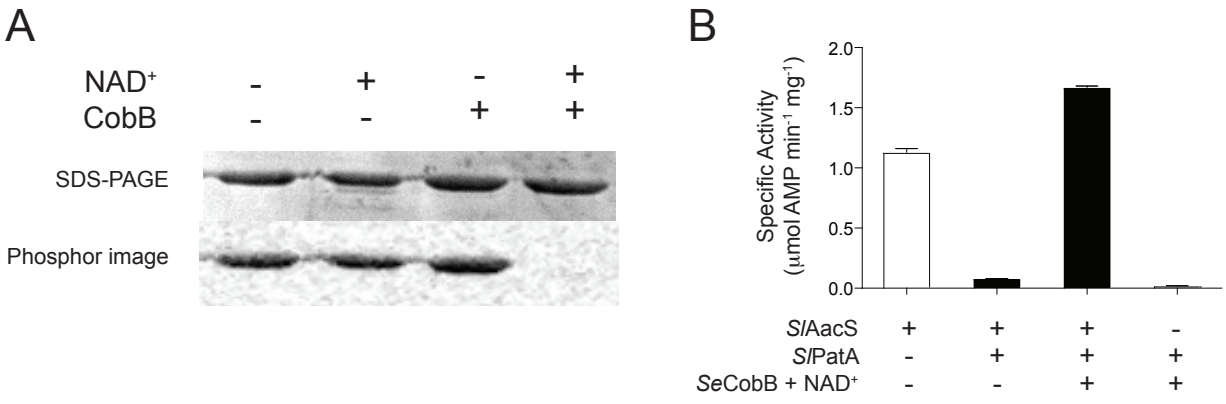


Figure 3.6. *Sl*AacS is deacetylated and reactivated by a heterologous sirtuin. *A*. *Sl*AacS previously acetylated by *Sl*PatA with [1-^{14}C]-acetyl-CoA was incubated with the addition of *Se*CobB and/or NAD$^+$. Proteins were resolved by SDS-PAGE and stained with Coomassie Blue to visualize proteins. Acetylation was visualized by phosphor image. B. *Sl*AacS was incubated in the presence or absence of *Sl*PatA and unlabeled acetyl-CoA and the acetylation reaction was stopped by buffer exchange. *Se*CobB and NAD$^+$ were added to the reactions and *Sl*AacS activity was measured in triplicate. Error bars represent standard deviations.

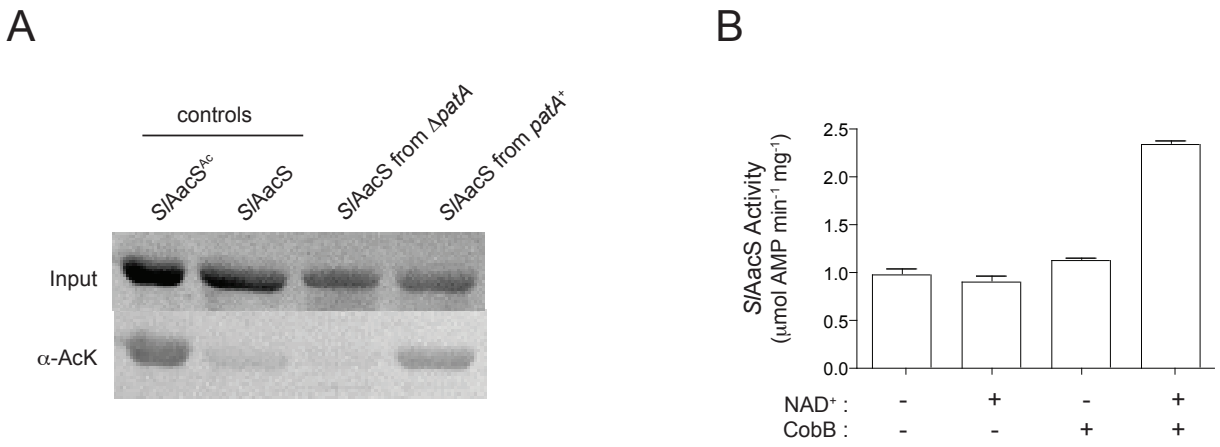


Figure 3.7. *Sl*AacS is acetylated in vivo in *S. lividans*. (A) H_6-*Sl*AacS was isolated from *S. lividans pat*$^+$ and Δ*patA* strains grown on acetoacetate. Acetylation state of the H_6-*Sl*AacS proteins was analyzed by anti-acetyllysine Western blot analysis. *In vitro* acetylated and non-acetylated *Sl*AacS were used as positive and negative controls, respectively. Total protein was visualized by Ponceau S staining prior to Western blot analysis. (B) H_6-*Sl*AacS isolated from *S. lividans pat*$^+$ and *ΔpatA* strains was incubated with *Se*CobB and NAD^+ to assess the affect of deacetylation on H_6-*Sl*AacS activity. For (A) and (B), experiments were conducted on H_6-*Sl*AacS isolated in three independent experiments. Error bars represent standard deviations.

*Sl*AacS, H_6-*Sl*AacS isolated from *S. lividans patA*$^+$ was subjected to trypsin digestion and the resulting peptides were analyzed by LC/MS/MS. Sequence determination of the peptides identified K617 as the single site of acetylation of *Sl*AacS (Figure 3.8).

Secondly, we quantified the fraction of acetylated *Sl*AacS present in the cell under growth the conditions used. To do this, we isolated H_6-*Sl*AacS proteins from *S. lividans patA*$^+$ and *S. lividans* Δ*patA* strains and incubated them with *Se*CobB and NAD^+ *in vitro*. The premise here was that the activity of acetylated H_6-*Sl*AacS should increase upon deacetylation by *Se*CobB, whilst the activity of non-acetylated H_6-*Sl*AacS should remain unchanged after incubation with *Se*CobB and NAD^+. In the presence of the co-substrate NAD^+, *Se*CobB deacetylated H_6-*Sl*AacS isolated from the *S. lividans patA*$^+$ strain, resulting in a 2.5-fold increase in H_6-*Sl*AacS activity (Figure 3.7B). In contrast, incubation of H_6-*Sl*AacS isolated from the *S. lividans* Δ*patA* strain with *Se*CobB and NAD^+ did not change the activity of H_6-*Sl*AacS (data not shown). Collectively, these data indicated that, in *S. lividans*, *Sl*PatA modulated the activity of H_6-*Sl*AacS *in vivo*.

Sl*AacS is required for growth of* S. lividans *on acetoacetate*. As shown above, *Sl*AacS was acetylated in *S. lividans* during growth in minimal medium supplemented with acetoacetate. On the basis of these results we made several predictions. First, we posited that *S. lividans* would not grow if putative genes encoding known protein deacetylases were deleted from the chromosome. This idea assumed that the absence of deacetylase activity would not affect in any way the ability of *Sl*PatA to acetylate *Sl*AacS resulting in a net accumulation of acetylated, inactive *Sl*AacS and should not support growth on acetoacetate. To test this hypothesis, we constructed a strain carrying chromosomal deletions of all three genes encoding putative protein deacetylases, i.e., ΔEFD68590 (Zn(II)-dependent deacetylase), ΔEFD65580 (*Sl*CobB2 sirtuin), and ΔEFD71509 (*Sl*CobB1 sirtuin). Surprisingly, growth of the ΔEFD68590 ΔEFD65580 ΔEFD71509 (JE16752) strain was comparable to that of wild-type *S. lividans* on acetoacetate (Fig. 3.9).

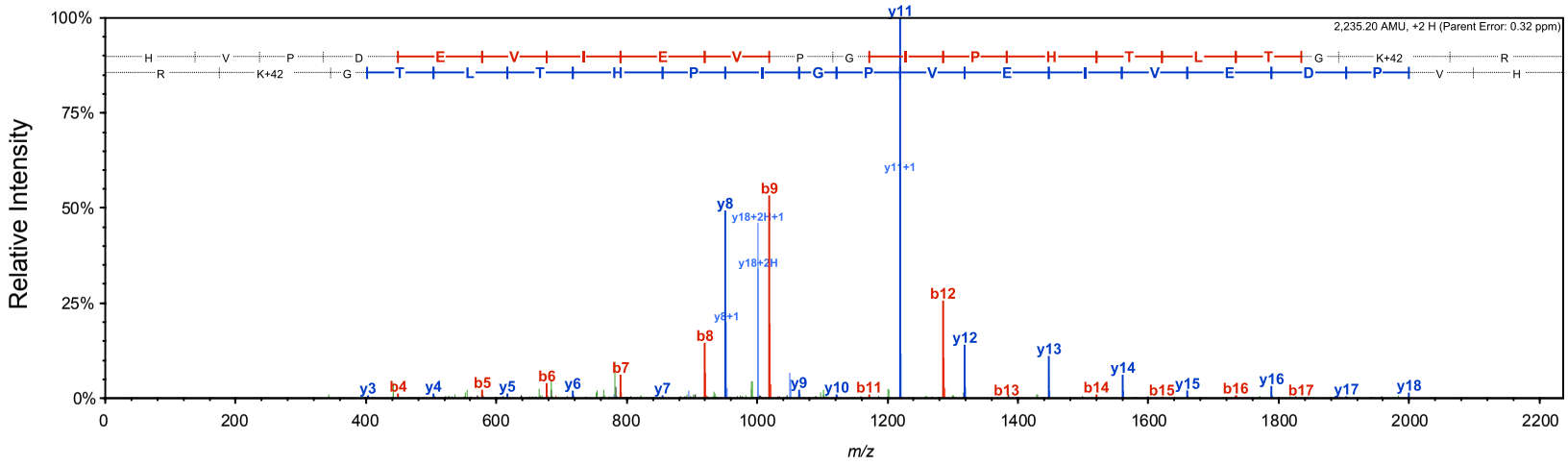


Figure 3.8. Residue K617 is the site of *Sl*AacS acetylation *in vivo*. LC/MS-MS analysis of H6-*Sl*AacS protein isolated from *S. lividans* strain JE16731 (*S. lividans* TK24 / p*Sl*AacS5) grown in NMMP medium supplemented with acetoacetate (10 mM). MS/MS specturm of the 2235.20 amu tryptic peptide, where peaks in red represent the b series m/z (predicted fragment ion masses of HVPDEVIEVPGIPHTLTGKAcR with charge on the N-terminal amino acid), and the blue peaks indicate y series *m/z* (predicted frament ion masses of HVPDEVIEVPGIPHTLTGKAcR) with the charge on the COOH-terminal amino acid). The y (blue) series suggests that the sequence of the presence of a 42 atomic mass unit addition to the HVPDEVIEVPGIPHTLTGKR fragments, specifically within the final three amino acids, GKR.

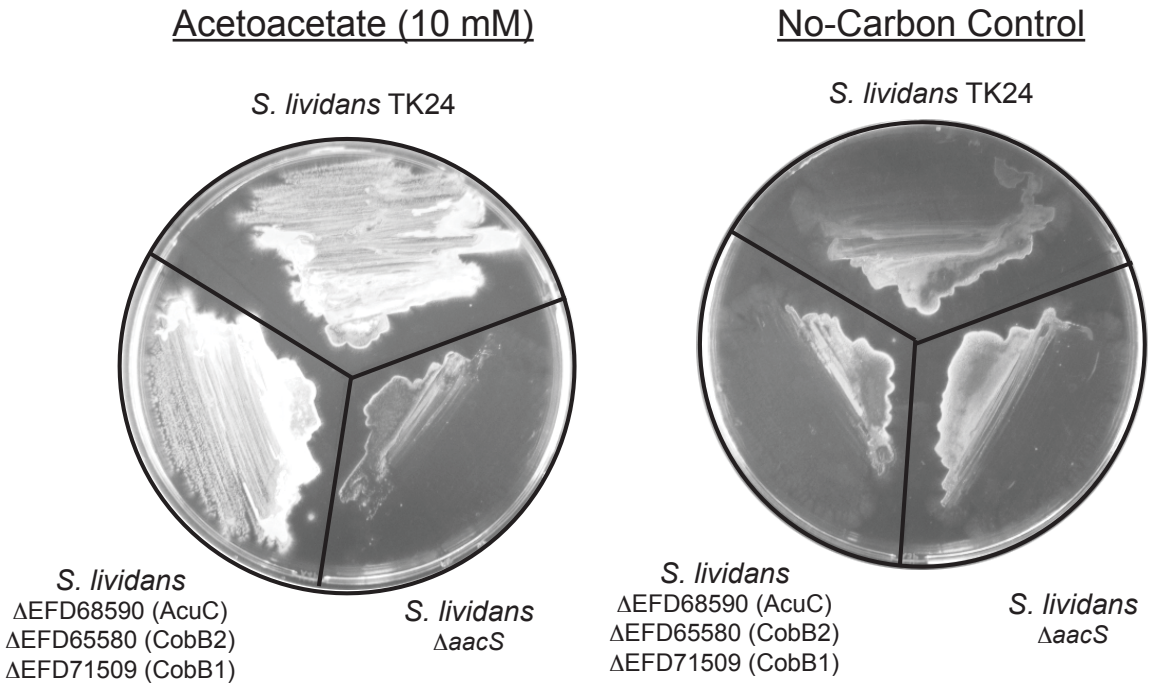


Figure 3.9. *Sl*AacS is required for growth of *S. lividans* on acetoacetate. Spores from *S. lividans* strains TK24 (wild-type), JE16752 (ΔEFD65580 ΔEFD68590 ΔEFD71509), and JE16758 (Δ*aacS*) were streaked on minimal medium containing acetoacetate (10 mM) or no additional carbon source. Plates were incubated at 30°C for 7 days prior to imaging using a Photodyne digital imaging system.

To assess whether or not *Sl*AacS was required for growth of *S. lividans* on acetoacetate, we constructed a *S. lividans* Δ*aacS* strain. Strains were tested for their ability to grow on minimal medium supplemented with acetoacetate. Unlike the wild-type strain, the Δ*aacS* (JE16758) strain grew poorly on acetoacetate (Fig. 3.9), indicating that *Sl*AacS activity was required for growth of *S. lividans* on this carbon and energy source.

In vivo *acetylation control of* Sl*AacS activity in a heterologous system*. We used *E. coli* to demonstrate that *Sl*PatA activity controlled *Sl*AacS function *in vivo*. The *E. coli pka* allele, encoding a *Se*Pat protein acetyltransferase homologue, was disrupted to avoid potential modification of *Sl*AacS by the *E. coli* protein acetylation machinery. As shown in figure 3.4B, an *atoDA pka* strain of *E. coli* can use *Sl*AacS to activate acetoacetate for subsequent use as a source of carbon and energy. Based on these data we predicted that, although *Sl*PatA would inactivate *Sl*AacS in *E. coli*, such negative effect on *Sl*AacS activity would be balanced by the activity of the *E. coli* CobB sirtuin deacetylase enzyme (*Ec*CobB, 92% identical to *Se*CobB). This prediction is illustrated in figure 3.10A. For this purpose, the coding sequence of the *S. lividans aacS* allele was optimized for expression in *E. coli*. We hypothesized that in the absence of *Ec*CobB, the Δ*atoDA* Δ*cobB* Δ*pka* / p*aacS*$^{+}$ p*patA*$^{+}$ strain (JE16969, the letter 'p' preceding a gene denotes that it is plasmid-borne) would not grow on acetoacetate due to the accumulation of inactive *Sl*AacS^{Ac}, a prediction that was experimentally confirmed, as shown in figure 3.10B (solid circles). These data showed that *Sl*PatA acetylation of *Sl*AacS modulated *Sl*AacS activity *in vivo* in a heterologous host. Synthesis of *Sl*PatA in an *E. coli cobB*$^{+}$ strain did not have a deleterious effect during growth on acetoacetate (Fig. 3.10B, open circles), indicating reversibility of *Sl*AacS acetylation by *Ec*CobB sirtuin deacetylation *in vivo*. Results of control experiments demonstrated that *cobB* was not required for growth of *E. coli* Δ*atoDA* Δ*pka* / p*aacS*$^{+}$ in the absence of *Sl*PatA (p*patA*$^{+}$, data not shown).

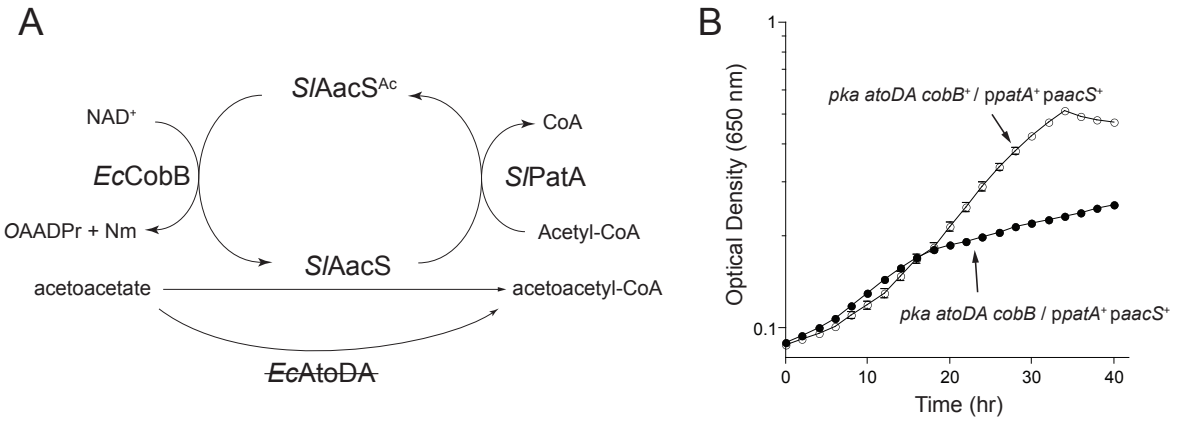


Figure 3.10. *Sl*PatA modulates growth of *E. coli* on acetoacetate using *Sl*AacS. (A) Schematic of predicted regulation of *E. coli atoDA aacS*$^{+}$ growth on acetoacetate. (B) *E. coli pka atoDA aacS*$^{+}$ strains encoding or lacking *cobB* were grown on NCE minimal medium supplemented with acetoacetate (30 mM) in the presence of *Sl*PatA (p*patA*$^{+}$). Growth experiments were performed at 37°C using a microtiter plate and a microtiter plate reader (Bio-Tek Instruments). Experiments were conducted in triplicate. Error bars represent standard deviations.

DISCUSSION

A new GNAT and a new acyl-CoA synthetase in* S. lividans *reveal a functional lysine acetylation system that may control acetoacetate metabolism in this and other actinomycetes. This work identified two new functions in *S. lividans*. One is a previously uncharacterized GNAT-type protein acetyltransferase with unique domain organization. This GNAT is conserved in some actinomycetes and in the archaeon *Archaeoglobus fulgidus* (Fig. 3.1). The existence of this new GNAT (encoded by EFD66247) was revealed through bioinformatics analysis and DNA sequencing to correct database mistakes in the sequence of the gene. The above-mentioned analysis showed a reversal in the domain organization of the protein (Fig. 3.1) with homology to other protein acetyltransferase (Pat) enzymes previously described. Hence we propose the name *Sl*PatA. The second new function is an acyl-CoA synthetase (encoded by EFD70521) responsible for the activation of acetoacetate to acetoacetyl-CoA, thus we propose the name *Sl*AacS for this protein. The name and functional assignments of these proteins are experimentally supported by *in vivo* and *in vitro* data (Figs. 3.3, 3.4, 3.5, 3.8). Collectively, the data show that *Sl*AacS is used by *S. lividans* to grow on acetoacetate, and that *Sl*PatA catalyzes the acetyl-CoA-dependent acetylation of *Sl*AacS.

The reversal of the domain organization of* Sl*PatA does not affect its acetyltransferase function. In *Sl*PatA, the large *C*-terminal domain (NDP-forming CoA ligase homologue) is found at the *N*-terminus of its homologues in *S. enterica*, *E. coli*, and *R. palustris*. This type of domain inversion has been observed in NDP-forming CoA ligases (51) and phosphoenolpyruvate-dependent sugar:phosphotransferase systems (49). This report expands the phenomenon of domain order reversal to Pat-type GNATs. Further characterization of these multi-domain GNAT enzymes is necessary to understand the how the activity of these enzymes may be regulated.

The role of the large NDP-forming CoA ligase-like domain in *Sl*PatA homologues remains unknown. Recently reported data obtained with *Se*Pat suggest that this large domain may be responsible multimerization (62). Such information is not yet available for *Sl*Pat. The domain inversion in *Sl*PatA may reflect an alternative regulatory mechanism.

The *Sl*PatA *C*-terminal domain contains a proline-rich region with a degenerate G-P-S motif, a signature of the fibrous protein collagen (28). In collagen and bacterial proteins with a G-X-Y motif, these motifs adopt extended fibrillar stretches of amino acids (28, 70). In *Klebsiella pneumoniae* strain FG9 a G-X-Y motif stabilizes homotrimers of the polysaccharide de-branching enzyme pullulanase (10). The role of the proline-rich region of *Sl*PatA in enzyme structure, oligomerization, and activity remains under investigation.

Control of* Sl*AacS activity by* Sl*PatA retains features observed in the control of other acyl-CoA synthetases by lysine acetylation. The mechanism of *Sl*AacS control by acetylation shares several features with previously reported enzymes under the same control. These features are: i) the acetylation site is a conserved lysine residue in the active site (in *Sl*AacS it is Lys617, Table 3.4); ii) activity of the enzyme is abolished upon acetylation; and iii) the modification can be enzymatically removed by a bona fide sirtuin deacetylase (Fig. 3.6). At the moment, what is unclear is which deacetylase in *S. lividans* is responsible for activation of *Sl*AacS^{Ac}. Our data (Fig. 3.9) cannot unambiguously identify the protein deacetylase, if any, that reactivates acetylated *Sl*AacS in *S. lividans*. Further analysis is needed to identify the *Sl*PatA cognate deacetylase.

Pat homologues *Rp*Pat and *Se*Pat (Fig. 3.1) were previously shown to acetylate and inactivate the acetyl-CoA synthetases *Rp*Acs and *Se*Acs, respectively. Notably, *Sl*PatA acetylates *Sl*AacS more efficiently than it does *S. lividans* acetyl-CoA synthetase, *Sl*Acs (Fig. 3.2, 3.5). Mikulik *et al.* demonstrated that Acs from the closely related *S. coelicolor* is acetylated *in vivo* (45), raising the question of which acetyltransferase acetylates *Sl*Acs in *S. lividans*. Work is currently underway to identify the role of *Sl*PatA and other protein acetyltransferases in *Sl*Acs regulation in *S. lividans*.

Physiological roles of acetoacetyl-CoA synthetase acetylation in bacteria, archaea, and eukaryotes. Clearly, *S. lividans* requires AacS activity for growth on acetoacetate (Fig. 3.9), however, it is not so clear whether this enzyme is needed to activate internally generated or externally transported acetoacetate. The chemical instability of acetoacetate (26) makes it likely that the source of substrate for *Sl*AacS is internal.

The *S. lividans* genome encodes five homologues of *E. coli* AtoB, the 3-ketoacyl-CoA thiolase that converts acetoacetyl-CoA and free coenzyme A to two molecules of acetyl-CoA (Fig. 3.4A), which are likely catabolized via the ethylmalonyl-CoA pathway for the degradation of acetate (18), since *S. lividans* lacks the glyoxylate cycle for C_2 compound assimilation (42). Since the conversion of acetoacetate to acetyl-CoA by *Sl*AacS requires input of ATP, we propose that acetylation and inactivation of *Sl*AacS prevents excess ATP and CoA consumption for synthesis of acetyl-CoA. This effect has been demonstrated in *S. enterica* when overproduction of acetyl-CoA synthetase (*Se*Acs) inhibits growth of *S. enterica* on acetate due to ATP depletion and loss of energy charge (9).

Cells of all domains of life synthesize acetoacetyl-CoA synthetase, including archaea, bacteria, nematodes, and mammals (Table 3.4). In the Gram-negative, nitrogen-fixing bacterium *Sinorhizobium meliloti*, the *aacS* homologue *acsA* is required for growth on poly-3-hydroxybutyrate (PHB) cycle intermediates acetoacetate and 3-hydroxybutyrate (8). Some bacteria use the PHB cycle to accumulate and store carbon in a reduced form. PHB can then be used as a carbon an energy source during times of physiological stress (1). Acetylation of AacS may regulate the PHB degradation in *S. meliloti* and other organisms that utilize AacS as a step of the PHB cycle, potentially expanding the role of lysine acetylation in prokaryotic cell physiology. The presence of a homologue of *Sl*Pat in *Archaeoglobus fulgidus* raises intriguing questions regarding the physiological role of lysine acylation in this extremely thermophilic, sulfate-reducing archaeon, especially because we did not find *Sl*Pat homologues in any other archaeal genomes in the databases.

In *Caenorhabditis elegans*, the *Sl*AacS homologue SUR-5 negatively regulates a vulval differentiation pathway (24). In mammals, acetoacetate-CoA synthetase is important for activation of ketone bodies for cholesterol and fatty acid biosynthesis (7, 23). AACS mRNA expression is regulated differentially in adipose tissue of genetically obese and nutritionally obese mice, suggesting that AACS may be important in lipogenesis and ketone body levels during hyper nutritional conditions (71).

AacS homologues contain the conserved active site lysine (Table 3.4). Hence, acetylation of acetoacetyl-CoA synthetase may represent a conserved mechanism of acetoacetyl-CoA synthetase control in these organisms.

***Apparent complexity of protein deacetylation in* S. lividans**. The efficient deacetylation of acetylated *Sl*AacS by a sirtuin deacetylase shows that the posttranslational modification of Lys617 is reversible (Fig. 3.6). The inability of purified *S. lividans* sirtuin-type deacetylase homologues, CobB1 and CobB2, or the zinc-dependent protein deacetylase homolog EFD68590 to deacetylate *Sl*AacS *in vitro* was unexpected because sirtuins from *S. enterica* and *E. coli* deacetylated *Sl*AacS *in vitro* and *in vivo*, respectively (Fig. 3.6A, 3.9B). The identity of the enzyme responsible for *Sl*AacS deacetylation *in vivo* remains unclear. *Frankia* sp. CcI3, a related actinomycete, encodes an *Sl*PatA homologue that clusters with an *Sl*AacS homologue and an EFD68590 (Zn(II)-dependent deacetylase) homologue (Fig. 3.11). This gene clustering suggests that Zn(II)-dependent protein deacetylase EFD68590 may regulate of *Sl*AacS activity.

A triple mutant of the *S. lividans* protein deacetylase homologues grew on acetoacetate (Fig. 3.9). This result was unexpected because of the accumulation of acetylated *Sl*AacS (Fig. 3.7A). We propose that: (i) the level of non-acetylated *Sl*AacS in *S. lividans* supports growth on acetoacetate as a carbon source, and deacetylation would not be necessary for growth under these conditions; (ii) not all the *Sl*AacS in the cell is acetylated, and what remains provides sufficient activity to support growth of *S. lividans* on acetoacetate; or (iii) *S. lividans* may encode an additional class of protein deacetylase capable of reactivating *Sl*AacS.

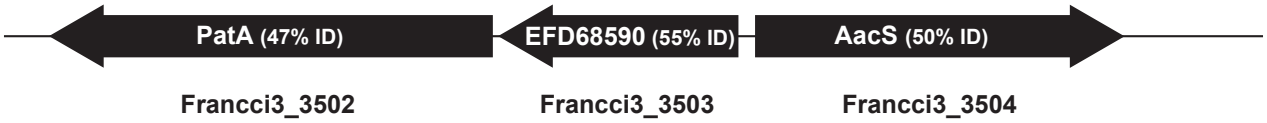


Figure 3.11. *Sl*AacS gene cluster in *Frankia* sp CcI3 suggests a candidate deacetylase for *Sl*AacS. *Frankia* sp. CcI3 encodes a *Sl*PatA, *Sl*AacS, and EFD68590 homologs in a gene cluster. Percent identity (% ID) to the respective *S. lividans* homologs is reported. The figure is drawn to scale.

REFERENCES

1. **Anderson, A. J., and E. A. Dawes.** 1990. Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol. Rev*. **54:**450-472.

2. **Beck, K., and B. Brodsky.** 1998. Supercoiled protein motifs: the collagen triple-helix and the alpha-helical coiled coil. *J. Struct. Biol*. **122:**17-29.

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

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

5. **Bierman, M., R. Logan, K. O'Brien, E. T. Seno, R. N. Rao, and B. E. Schoner.** 1992. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116:**43-49.

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

7. **Buckley, B. M., and D. H. Williamson.** 1973. Acetoacetate and brain lipogenesis: developmental pattern of acetoacetyl-coenzyme A synthetase in the soluble fraction of rat brain. *Biochem. J*. **132:**653-656.

8. **Cai, G. Q., B. T. Driscoll, and T. C. Charles.** 2000. Requirement for the enzymes acetoacetyl coenzyme A synthetase and poly-3-hydroxybutyrate (PHB) synthase for growth of *Sinorhizobium meliloti* on PHB cycle intermediates. *J. Bacteriol.* **182:**2113-2118.

9. **Chan, C. H., J. Garrity, H. A. Crosby, and J. C. Escalante-Semerena.** 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.

10. **Charalambous, B. M., J. N. Keen, and M. J. McPherson.** 1988. Collagen-like sequences stabilize homotrimers of a bacterial hydrolase. *EMBO J*. **7:**2903-2909.

11. **Chater, K. F.** 2006. *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. **361:**761-768.

12. **Choudhary, C., C. Kumar, F. Gnad, M. L. Nielsen, M. Rehman, T. Walther, J. V. Olsen, and M. Mann.** 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325:**834-840.

13. **Courtois, S., C. M. Cappellano, M. Ball, F. X. Francou, P. Normand, G. Helynck, A. Martinez, S. J. Kolvek, J. Hopke, M. S. Osburne, P. R. August, R. Nalin, M. Guerineau, P. Jeannin, P. Simonet, and J. L. Pernodet.** 2003. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol.* **69:**49-55.

14. **Crosby, H. A., E. K. Heiniger, C. S. Harwood, and J. C. Escalante-Semerena.** 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.

15. **Crosby, H. A., D. A. Pelletier, G. B. Hurst, and J. C. Escalante-Semerena.** 2012. System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J. Biol. Chem*. **287:**15590-15601.

16. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97:**6640-6645.

17. **Elion, E. A., P. Marina, and L. Yu.** 2007. Constructing recombinant DNA molecules by PCR, p. Unit 3.17.1-3.17.12. *In* F. M. Ausubel, R. E. R. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, a. Smith, and K. Struhl (ed.), *Current protocols in molecular biology*. Greene Publishing Associates & Wiley Interscience, New York, N.Y.

18. **Erb, T. J., I. A. Berg, V. Brecht, M. Muller, G. Fuchs, and B. E. Alber.** 2007. Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. *Proc. Natl. Acad. Sci. USA* **104:**10631-6.

19. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc. Natl. Acad. Sci. USA* **76:**1648-1652.

20. **Fukui, T., M. Ito, and K. Tomita.** 1982. Purification and characterization of acetoacetyl-CoA synthetase from *Zoogloea ramigera* I-16-M. *Eur. J. Biochem*. **127:**423-428.

21. **Gardner, J. G., F. J. Grundy, T. M. Henkin, and J. C. Escalante-Semerena.** 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD(+) involvement in *Bacillus subtilis*. *J. Bacteriol*. **188:**5460-5468.

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

23. **Geelen, M. J., M. Lopes-Cardozo, and J. Edmond.** 1983. Acetoacetate: a major substrate for the synthesis of cholesterol and fatty acids by isolated rat hepatocytes. *FEBS Lett.* **163:**269-273.

24. **Gu, T., S. Orita, and M. Han.** 1998. *Caenorhabditis elegans* SUR-5, a novel but conserved protein, negatively regulates LET-60 Ras activity during vulval induction. *Mol. Cell. Biol.* **18:**4556-4564.

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

26. **Hay, R. W., and M. A. Bond.** 1967. Kinetics of decarboxylation of acetoacetic acid. *Aust. J. Chem*. **20:**1823-1828.

27. **Horswill, A. R., and J. C. Escalante-Semerena.** 2002. Characterization of the propionyl-CoA synthetase (PrpE) enzyme of *Salmonella enterica*: Residue Lys592 is required for propionyl-AMP synthesis. *Biochemistry* **41:**2379-2387.

28. **Hulmes, D. J.** 1992. The collagen superfamily - diverse structures and assemblies. *Essays Biochem.* **27:**49-67.

29. **Ikeuchi, Y., K. Kitahara, and T. Suzuki.** 2008. The RNA acetyltransferase driven by ATP hydrolysis synthesizes N4-acetylcytidine of tRNA anticodon. *EMBO J.* **27:**2194-2203.

30. **Ito, M., T. Fukui, M. Kamokari, T. Saito, and K. Tomita.** 1984. Purification and characterization of acetoacetyl-CoA synthetase from rat liver. *Biochim. Biophys. Acta* **794:**183-193.

31. **Jenkins, L. S., and W. D. Nunn.** 1987. Genetic and molecular characterization of the genes involved in short-chain fatty acid degradation in *Escherichia coli*: the *ato* system. *J. Bacteriol.* **169:**42-52.

32. **Kawamoto, S., and K. Ochi.** 1998. Comparative ribosomal protein (L11 and L30) sequence analyses of several *Streptomyces* spp. commonly used in genetic studies. *Int. J. Syst. Bacteriol.* **48:**597-600.

33. **Kawamoto, Y., K. Shinozuka, M. Kunitomo, and J. Haginaka.** 1998. Determination of ATP and its metabolites released from rat caudal artery by isocratic ion-pair reversed-phase high-performance liquid chromatography. *Anal. Biochem*. **262:**33-38.

34. **Keller, A., A. I. Nesvizhskii, E. Kolker, and R. Aebersold.** 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74:**5383-5392.

35. **Khan, S. R., J. Gaines, R. M. Roop, 2nd, and S. K. Farrand.** 2008. Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl. Environ. Microbiol.* **74:**5053-5062.

36. **Kieser, T., M. J. Bibb, M. J. Buttner, K. Chater, and D. A. Hopwood.** 2000. Growth and preservation of *Streptomyces*, p. 43-62, *Practical* Streptomyces *Genetics*. John Innes Foundation, Norwich, England.

37. **Kieser, T., M. J. Bibb, M. J. Buttner, K. Chater, and D. A. Hopwood.** 2000. Introduction of DNA into *Streptomyces*, p. 229-252, *Practical* Streptomyces *Genetics*. John Innes Foundation, Norwich, England.

38. **Kieser, T., M. J. Bibb, M. J. Buttner, K. Chater, and D. A. Hopwood.** 2000. Media, buffers, and suppliers, p. 405-420, *Practical* Streptomyces *Genetics*. John Innes Foundation, Norwich, England.

39. **Kim, S. C., R. Sprung, Y. Chen, Y. Xu, H. Ball, J. Pei, T. Cheng, Y. Kho, H. Xiao, L. Xiao, N. V. Grishin, M. White, X. J. Yang, and Y. Zhao.** 2006. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell.* **23:**607-618.

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

41. **Lee, H. J., P. T. Lang, S. M. Fortune, C. M. Sassetti, and T. Alber.** 2012. Cyclic AMP regulation of protein lysine acetylation in *Mycobacterium tuberculosis*. *Nat. Struct. Mol. Biol.*

42. **Lewis, R. A., E. Laing, N. Allenby, G. Bucca, V. Brenner, M. Harrison, A. M. Kierzek, and C. P. Smith.** 2010. Metabolic and evolutionary insights into the closely-related species S*treptomyces coelicolor* and *Streptomyces lividans* deduced from high-resolution comparative genomic hybridization. *BMC Genomics* **11:**682.

43. **Martinez, A., S. J. Kolvek, C. L. Yip, J. Hopke, K. A. Brown, I. A. MacNeil, and M. S. Osburne.** 2004. Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl. Environ. Microbiol.* **70:**2452-2463.

44. **McMahon, M. D., C. Guan, J. Handelsman, and M. G. Thomas.** 2012. Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. *Appl. Environ. Microbiol.* **78:**3622-3629.

45. **Mikulik, K., J. Felsberg, E. Kudrnacova, S. Bezouskova, D. Setinova, E. Stodulkova, J. Zidkova, and V. Zidek.** 2012. CobB1 deacetylase activity in *Streptomyces coelicolor*. *Biochem. Cell Biol.* **90:**179-187.

46. **Nambi, S., N. Basu, and S. S. Visweswariah.** 2010. Cyclic AMP-regulated protein lysine acetylases in mycobacteria. *J. Biol. Chem.* **285:**24313-24323.

47. **Nesvizhskii, A. I., A. Keller, E. Kolker, and R. Aebersold.** 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75:**4646-4658.

48. **Pauli, G., and P. Overath.** 1972. *ato* Operon: a highly inducible system for acetoacetate and butyrate degradation in *Escherichia coli*. *Eur. J. Biochem.* **29:**553-562.

49. **Reizer, J., and M. H. Saier, Jr.** 1997. Modular multidomain phosphoryl transfer proteins of bacteria. *Curr. Opin. Struct. Biol.* **7:**407-415.

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

51. **Sánchez, L. B., M. Y. Galperin, and M. Muller.** 2000. Acetyl-CoA synthetase from the amitochondriate eukaryote Giardia lamblia belongs to the newly recognized superfamily of acyl-CoA synthetases (Nucleoside diphosphate-forming). *J. Biol. Chem.* **275:**5794-5803.

52. **Seow, K. T., G. Meurer, M. Gerlitz, E. Wendt-Pienkowski, C. R. Hutchinson, and J. Davies.** 1997. A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J. Bacteriol.* **179:**7360-7368.

53. **Shaw, K. J., P. N. Rather, R. S. Hare, and G. H. Miller.** 1993. Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. Microbiol. Rev. **57:**138-163.

54. **Shirling, E. B., and D. Gottlieb.** 1966. Methods for characterization of *Streptomyces* species. Int. J. Syst. Bacteriol. **16:**313-340.

55. **Soppa, J.** 2010. Protein acetylation in archaea, bacteria, and eukaryotes. Archaea **2010**.

56. **Spange, S., T. Wagner, T. Heinzel, and O. H. Kramer.** 2009. Acetylation of non-histone proteins modulates cellular signalling at multiple levels. Int. J. Biochem. Cell. Biol. **41:**185-198.

57. **Staack, H., J. F. Binstock, and H. Schulz.** 1978. Purification and properties of a pig heart thiolase with broad chain length specificity and comparison of thiolases from pig heart and *Escherichia coli*. J. Biol. Chem. **253:**1827-1831.

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

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

60. **Takenoya, M., K. Nikolakakis, and M. Sagermann.** 2010. Crystallographic insights into the pore structures and mechanisms of the EutL and EutM shell proteins of the Eut-BMC. J. Bacteriol. **192:**6056-6063.

61. **Thao, S., C. S. Chen, H. Zhu, and J. C. Escalante-Semerena.** 2010. N(epsilon)-Lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. PLoS ONE **5:**e15123.

62. **Thao, S., and J. C. Escalante-Semerena.** 2011. Biochemical and thermodynamic analyses of *Salmonella enterica* Pat, a multidomain, multimeric N(epsilon)-lysine acetyltransferase involved in carbon and energy metabolism. mBio **2:**e00216-11.

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

64. **Tsang, A. W., and J. C. Escalante-Semerena.** 1998. CobB, a new member of the SIR2 family of eucaryotic regulatory proteins, is required to compensate for the lack of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase activity in *cobT* mutants during cobalamin biosynthesis in *Salmonella typhimurium* LT2. J. Biol. Chem. **273:**31788-31794.

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

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

67. **Ward, J. M., G. R. Janssen, T. Kieser, M. J. Bibb, and M. J. Buttner.** 1986. Construction and characterisation of a series of multi-copy promoter-probe plasmid vectors for *Streptomyces* using the aminoglycoside phosphotransferase gene from Tn*5* as indicator. Mol. Gen. Genet. **203:**468-478.

68. **Weinert, B. T., S. A. Wagner, H. Horn, P. Henriksen, W. R. Liu, J. V. Olsen, L. J. Jensen, and C. Choudhary.** 2011. Proteome-wide mapping of the *Drosophila* acetylome demonstrates a high degree of conservation of lysine acetylation. Sci. Signal. **4:**ra48.

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

70. **Xu, Y., D. R. Keene, J. M. Bujnicki, M. Hook, and S. Lukomski.** 2002. Streptococcal Scl1 and Scl2 proteins form collagen-like triple helices. J. Biol. Chem. **277:**27312-27318.

71. **Yamasaki, M., S. Hasegawa, T. Kitani, K. Hidai, and T. Fukui.** 2007. Differential effects of obesity on acetoacetyl-CoA synthetase gene in rat adipose tissues. Eur. J. Lipid Sci. Technol. **109:**617-622.

72. **Yu, B. J., J. A. Kim, J. H. Moon, S. E. Ryu, and J. G. Pan.** 2008. The diversity of lysine-acetylated proteins in *Escherichia coli*. J. Microbiol. Biotechnol. **18:**1529-1536.

73. **Zhang, J., R. Sprung, J. Pei, X. Tan, S. Kim, H. Zhu, C. F. Liu, N. V. Grishin, and Y. Zhao.** 2009. Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. Mol. Cell. Proteomics **8:**215-225.

74. **Zhao, S., W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li, Y. Li, J. Shi, W. An, S. M. Hancock, F. He, L. Qin, J. Chin, P. Yang, X. Chen, Q. Lei, Y. Xiong, and K. L. Guan.** 2010. Regulation of cellular metabolism by protein lysine acetylation. Science **327:**1000-1004.

CHAPTER 4

IDENTIFICATION OF ACETYLATION TARGETS OF THE PROTEIN ACETYLTRANSFERASE PATA IN *STREPTOMYCES LIVIDANS*[1]

[1] Tucker A.C. and J.C. Escalante-Semerena. To be submitted to *Proteomics*.

ABSTRACT

Acetylation of lysine residues affects protein function in all domains of life. Acetyltransferases acetylate specific lysines in substrate proteins. Gcn5-related *N*-acetyltransferases (GNATs) are the single family of acetyltransferases identified in all domains of life. GNATs are the only family of acetyltransferases identified in bacteria. GNATs acetylate a diverse set of substrates. Understanding the role of GNAT-mediated protein acetylation in the metabolism of the actinomycetes is important due to the ability of these organisms to produce diverse natural products. Previously we identified a protein acetyltransferase PatA in *Streptomyces lividans*. Here we correlate protein acetylation profiles of *S. lividans* with production of PatA during growth of this organism in liquid medium. Using 2-D gel electrophoresis, anti-acetyllysine Western blotting, and mass peptide fingerprinting, we have identified new substrates regulated by PatA. Our work identifies new targets of PatA acetylation and expands the role of acetylation in *S. lividans* physiology.

INTRODUCTION

Reversible lysine acetylation is a post-translational modification that can affect enzyme activity, protein-protein interactions, protein-DNA interactions, and protein stability (20, 24, 34). Protein acetylation occurs in all domains of life (33). The study of acetylomes (lysine acetylation profiles of total cellular proteins) in bacteria and eukaryotes has revealed a large number of acetylated proteins in these organisms (19). Few of the acetylation targets are validated and the effects of most of the acetylation events remain unknown. Recent work in the model bacterium *Escherichia coli* suggests that wide-spread, low-level protein acetylation occurs in this organisms due to the presence of acetyl-phosphate, a reactive acetyl donor (41). Our ability to predict physiological impacts of acetylation events is limited in the absence of direct experimental data.

Acetyltransferase enzymes transfer the acetyl group from acetyl-CoA to specific lysine residues on the protein substrates (1, 10, 11, 21). Gcn5-related *N*-acetyltransferases (GNATs) are the only acetyltransferases identified in all domains of life (38). In eukaryotes, GNATs were first identified for

their role in modifying DNA-binding proteins histones (4). Lysine acetylation on histones affects chromatin maintenance by altering DNA-histone interactions (20). Acetylation of the DNA-binding protein Alba in the archaeon *Sulfolobus solfataricus* decreased affinity of Alba for DNA (25, 40). Studies in the bacterium *S. enterica* revealed that reversible lysine acetylation directly regulates metabolism. Acetylation of the catalytic lysine of acetyl-CoA synthetase (Acs) inactivates this enzyme. Acs is a member AMP-forming CoA ligase family. This family of enzymes activates organic acids to CoA thioesters via an acyl-AMP intermediate (15). Adenylylation of the organic acid occurs in the first half reaction and is catalyzed by a conserved lysine. Acetylation of the catalytic lysine abolishes the adenylylation activity of the AMP-forming CoA ligase. This family of enzymes has been identified as a target for lysine acetylation in bacteria and eukaryotes (9, 10, 13, 16, 28, 34, 39).

GNATs exhibit substrate specificity. The α-proteobacterium *Rhodopseudomonas palustris* produces a two-domain protein acetyltransferase (Pat) can distinguish between members of the AMP-forming CoA ligase family (10). Pat contains a large (~700 aa) *N*-terminal domain of unknown function and *C*-terminal GNAT domain (~200 aa). Pat homologs from *E. coli* (Pka) and the closely related γ-proteobacterium *Salmonella enterica* acetylate the AMP-forming CoA ligase Acs (described above). Additionally, *E. coli* Pka and *S. enterica* Pat acetylate ribonuclease R and transcription factor RcsB, respectively. Neither ribonuclease R nor RcsB is structurally related to Acs (23, 24, 37). Together, these findings suggest that Pat homologs can recognize a diverse but specific set of substrates.

The Gram-positive actinomycete *Streptomyces lividans* encodes a Pat homolog, namely PatA, in which the two domains are reversed relative to the Pat homologs in *E. coli*, *S. enterica*, and *R. palustris*. *S. lividans* PatA contains an *N*-terminal GNAT domain linked to a large *C*-terminal domain of unknown function. PatA acetylates acetoacetyl-CoA synthetase (AacS), an AMP-forming CoA ligase in *S. lividans* (39). The Acs homolog in the closely related actinomycete *Streptomyces coelicolor* is acetylated *in vivo* (27), but the acetyltransferase responsible for the acetylation has not been reported. *S. lividans* encodes 72 putative GNATs, suggesting potential for extensive lysine acetylation in this organism.

Understanding metabolic regulation in *Streptomyces* species is of interest due to the ability of these organisms to synthesize diverse secondary metabolites and natural products (7, 8, 26, 31). To characterize the role protein acetylation plays in the physiology of *S. lividans*, we have identified targets of the protein acetyltransferase PatA. We correlate the production of the protein acetyltransferase PatA with distinct protein acetylation profiles during growth in liquid medium. Using 2-dimensional gel electrophoresis, acetyl-lysine western blotting, and mass peptide fingerprinting, we identify 4 AMP-forming CoA ligases that are acetylated in a *S. lividans patA*$^{+}$ strain. We demonstrate that these uncharacterized AMP-forming CoA ligases recognize distinct substrates, implicating PatA regulation in diverse metabolic pathways of *S. lividans.*

EXPERIMENTAL PROCEDURES

Bacterial strains and growth conditions. All strains and plasmids used in this study are listed in Tables 4.1 and 4.2, respectively. *Streptomyces* strains are derivatives of *Streptomyces lividans* TK24. ISP-2 medium (32) or R2YE medium (18) was used to culture *S. lividans* on solid medium. Liquid cultures of *S. lividans* were inoculated with 5 x 10^8 spores germinated as described (17). Briefly, spores were diluted into 1 ml TES buffer (50 mM, pH 8.0) and incubated at 50°C for 10 minutes. One ml of 2X Germination medium was added and cells were grown at 37°C with aeration for 2 hours. Germinated spores were washed with 1 ml TES buffer and resuspended in 1 ml fresh TES buffer to use as an inoculum. *S. lividans* grown in either yeast extract-malt extract (YEME) rich medium or NMMP medium supplemented with β-hydroxybutyrate (10 or 25 mM), glucose (10 mM), *N*-acetylglucosamine (10 or 25 mM), maltose (25 mM), glycerol (25 mM) (18). *S. lividans* liquid cultures were grown in flasks or with marine-grade stainless steel springs to aid in cell dispersion. Strains were cultured for 72 hour at 30°C.

Unless noted otherwise, all *E. coli* strains used were derivatives of *E. coli* MG1655. *E. coli* strains were grown at 37°C in lysogeny broth (LB, Difco) (2). When necessary, antibiotics were used at the following concentrations: ampicillin, 100 μg ml^{-1}; chloramphenicol, 12.5 μg ml^{-1}.

Table 4.1. Strains used in this study.

Strain or Plasmid	Relevant Genotype	Reference[2]
E. coli strains		
JE9314	C41λ(DE3) *pka12*::*kan*$^+$	Lab Collection
S. lividans strains		
TK24	Wild type	Lab Collection
Derivatives of *S. lividans* TK24		
JE16707	Δ*patA*	

[2] Unless otherwise indicated, all strains were constructed during the course of this work.

Table 4.2. Plasmids used in this study.

Plasmids	Relevant Genotype	Reference[3]
pTEV5	N-terminal, rTEV-cleavable His_6-tag overexpression vector, *bla*$^+$	(30)
pEFD68037-1	*S. lividans* MlaL (EFD68037) allele in pTEV5, *bla*$^+$	
pEFD65795-1	*S. lividans* LbuL (EFD65795) allele in pTEV5, *bla*$^+$	
pEFD65796-1	*S. lividans* IbaL (EFD65796) allele in pTEV5, *bla*$^+$	
p*Sl*Paa1	*S. lividans* PaaA (EFD64737) allele in pTEV5, *bla*$^+$	

[3] Unless otherwise indicated, all strains were constructed during the course of this work.

Table 4.3. Primers used in this study.

Name	Sequence[4]
EFD65795 5' NheI pTEV5	GTA**GCTAGC**ATGACCGCACCCGCGCCCCAGC
EFD65795 3' EcoRI pTEV5	ACT**GAATTC**TCAGGGGCGCGTTCCGTACCGC
EFD65796 5' NheI pTEV5	GTA**GCTAGC**ATGACGACGGCCACGGAGCTG
EFD65796 3' EcoRI pTEV5	ACT**GAATTC**TCACCGGAAGTCCTCCTCGCGG
EFD68037 5' NheI pTEV5	GTA**GCTAGC**ATGTCGCCCCGGGAGGACAC
EFD68037 3' EcoRI pTEV5	ACT**GAATTC**TCAGAGCCGGGTGACGTCGAG
EFD64737 NheI pTEV5	GTA**GCTAGC**ATGAGCAGCGAGCCGACGAC
EFD64737 EcoRI pTEV5	ACT**GAATTC**TCACGCGCCCCGCTGGTC

[4] Bold typeface indicates restriction sites

***Preparation of* S. lividans *extracts*.** Twenty or 25 ml of *S. lividans* cultures was diluted into 25 ml distilled water. Cells were harvested by centrifuged 10 min at 4,000 x *g*. Cell pellets were washed once with 25 ml of tris(hydroxymethyl)aminomethane-HCl (Tris-HCl) buffer and excess buffer was removed. Cell pellets were weighed and flash frozen in liquid nitrogen. Cells were stored at -80°C until lysis. *S. lividans* cells were resuspended in 0.5 or 1 ml lysis buffer [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 50 mM), NaCl (100 mM, lysozyme (1 mg ml^{-1}), DNase (50 μg ml^{-1}), SIGMAFAST Protease Inhibitor Cocktail Tablet (EDTA-free, 0.01 tablet ml^{-1}), and phenylmethanesulfonylfluoride (PMSF, 0.5 mM). Cells were placed on ice and lysed by sonication for 2 sec (0.2-s pulse followed by 0.2-s cooling) at level 7 in a model 550 sonic dismembrator (Fisher). Samples were centrifuged at 16,000 x *g* for 30 minutes to remove insoluble material. Soluble proteins were quantified using the Bio-Rad Protein assay according to the manufacturer's protocol (Bio-Rad).

***One-dimensional Western blot analysis*.** For one-dimesional gel analysis Fifty micrograms of cellular protein were resolved using SDS-PAGE (22) and transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore). Rabbit polyclonal *Sl*PatA antiserum (Laboratory Animal Resources, University of Wisconsin, Madison, WI) was used to detect *Sl*PatA (1:10,000 final dilution). Rabbit polyclonal anti-acetyllysine antibodies were used to detect acetylated proteins (Calbiochem, 1:5,000 dilution). Binding of primary antibodies to blots was visualized using alkaline phosphatase-conjugated goat α-rabbit immunoglobulin G (ThermoFisher) and NBT/BCIP 1-Step substrate (Pierce) according to the manufacturers instructions.

***Two-dimensional Western blot analysis*.** Two-dimensional electrophoresis was performed according to the carrier ampholine method of isoelectric focusing (5, 29) by Kendrick Labs, Inc. (Madison, WI). Duplicate gels were run for lysates of *S. lividans* TK24 and *S. lividans* Δ*patA* strains grown on NMMP supplemented with hydroxybutyrate (10 mM) for 60 hours. Isoelectric focusing was carried out in a glass tube of inner diameter 3.3 mm using 2.0% pH 4-8 mix Servalytes (Serva, Heidelberg, Germany; and 2 mM lysine) for 20,000 volt-hrs. One ug of an IEF internal standard, tropomyosin, was added to each sample. This protein migrates as a doublet with lower polypeptide spot

of MW 33,000 and pI 5.2. The enclosed tube gel pH gradient plot for this set of Servalytes was determined with a surface pH electrode.

After equilibration for 10 min in Buffer 'O' (10% glycerol, 50 mM dithiothreitol, 2.3% SDS and 0.0625 M tris, pH 6.8), each tube gel was sealed to the top of a stacking gel that overlaid a 10% acrylamide slab gel (1.00 mm thick). SDS slab gel electrophoresis was carried out for about 5 hrs at 25 mA/gel. The following proteins (Sigma Chemical Co., St. Louis, MO) were used as molecular weight standards: myosin (220,000), phosphorylase A (94,000), catalase (60,000), actin (43,000) carbonic anhydrase (29,000) and lysozyme (14,000). These standards appear as bands at the basic edge of the Coomassie Brilliant Blue R-250-stained gel. The gel was dried between sheets of cellophane with the acid edge to the left.

After slab gel electrophoresis, the gels for blotting were placed in transfer buffer (10 mM CAPS, pH 11.0, 10% MeoH) and transblotted onto a PVDF membrane overnight at 225 mA and approximately 100 volts/ two gels. The following proteins (Sigma Chemical Co., St. Louis, MO) were used as molecular weight standards: myosin (220,000), phosphorylase A (94,000), catalase (60,000), actin (43,000) carbonic anhydrase (29,000) and lysozyme (14,000). These standards appear as bands at the basic edge of the Coomassie Brilliant Blue R-250-stained membrane.

Membranes were destained with 10% methanol prior to blotting. Acetylated lysine residues were detected as described above. Corresponding spots were cut from the duplicate gels, and samples were analyzed by mass spectrometry.

Determination of the acetylated proteins by mass spectrometry. To determine the identity of the acetylated proteins identified by 2-dimensional gel electrophoresis, “In Gel” digestion and mass spectrometric analysis was done at the Proteomics and Mass Spectrometry Facility (University of Georgia). The digestion was performed as outlined on the website: http://mass-spec.uga.edu/protocals/trypsindigest.htm. Mass spectrometry analyses were performed on a Thermo-Fisher LTQ Orbitrap Elite Mass Spectrometer coupled with a Proxeon Easy NanoLC system (Waltham,

MA). Peptides were loaded into a Dionex PepMap 5mm long and 300 µm id precolumn (Sunnyvale,CA) first, then were separated by a self-packed ~12 cm long, 100 µm id column/emitter with 200 Å 5 µM Burker MagicAQ C18 (Auburn, CA), at 500 uL by gradient elution. Briefly, the two-buffer gradient (0.1% formic acid as buffer A and 99.9% acetonitrile with .1% formic acid as buffer B) starts with 5% B, increases to 40% B in 40 minutes, to 60% B in 15 minutes, and to 95% B in 10 minutes. The Top 10 data-dependent acquisition method was used to acquire data. Both MS and MS/MS scans were analyzed using Orbitrap at the resolutions of 120,000 and 30,000, respectively. Proteome Discoverer 1.3 (Thermo-Fisher) software was used with Mascot database search program (Matrix Science, Landon, UK) for protein identifications.

Molecular techniques. DNA manipulations were performed using standard techniques (12). Restriction endonucleases were purchased from Fermentas. DNA was amplified using Pfu Ultra II Fusion DNA polymerase (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 sequencing reactions were resolved at the University of Georgia Genomics facility. Oligonucleotide primer sequences are listed in Table 4.3.

Plasmids used for protein overproduction. The *S. lividans lbuL* (EFD65795), *IbaL* (EFD65796), *mlaL* (EFD68037), *paaL* (EFD64737) genes were amplified from purified *S. lividans* TK24 genomic DNA with the primers listed in Table 3.3. The first codon of EFD65795 (GTG) was changed to the more common ATG start codon. The DNA fragments were digested with NheI and EcoRI and ligated into pTEV5 (30) cut with the same enzymes. The resulting plasmids pEFD68037-1 and pEFD65795-1, pEFD65796-1, and p*Sl*Paa1 direct synthesis of *Sl*MlaL, *Sl*LbuL, *Sl*IbaL, and *Sl*PaaL respectively, with *N*-terminal H_6 tags cleavable by recombinant tobacco etch virus (rTEV) protease prepared as described (3).

***Protein purification*.** Plasmids containing *S. lividans lbuL* (EFD65795), *IbaL* (EFD65796), *mlaL* (EFD68037), or *paaL* (EFD64737) were transformed with pRARE2 (EMD Millipore) into a Δ*pka* derivative of C41λ(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 LB containing ampicillin (100 μg ml^{-1}) and chloramphenicol (12.5 μg ml^{-1}). The cultures were grown shaking at 37°C to A_{600} ~ 0.7 and protein synthesis was induced with IPTG (0.25 mM). Upon induction, the cultures were grown overnight at 30°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 resuspended in 30 ml cold His-Bind buffer (buffer A) [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM)], and imidazole (5 mM)) containing PMSF (1 mM). Cells were placed on ice and lysed by sonication for 2 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*. H_6-tagged proteins were purified from clarified cell extract using a 1 ml settled bed volume of HisPur™ Ni-NTA Resin (Pierce). Unbound proteins were eluted off the column by extensive washing with buffer A. The resin was washed with 10 column volumes of buffer B [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (15 mM)]. H_6-tagged proteins were eluted with 5 CV Buffer C [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (250 mM)]. All fractions containing H_6-tagged proteins were combined. rTEV protease was added to H_6-tagged proteins and the H_6-tagged proteins/rTEV mixtures were incubated at room temperature for 3 h. PMSF was added to the protein mixture and incubated 15 min at room temperature. The protein/rTEV mixtures were dialyzed at 4°C against buffer D (Tris-HCl (50 mM, pH 8), NaCl (500 mM)) twice for 3 hours and again against buffer D containing imidazole (5 mM) for 12 h. After cleavage and dialysis, protein mixtures were passed over 1 ml HisPur™ Ni-NTA Resin (Pierce) using the buffers described above. Cleaved proteins passed through the resin and eluted in the flow-through fractions. Purified proteins were analyzed by SDS-PAGE. Fractions containing target proteins were pooled together. Target proteins were stored in Tris-Cl buffer (50 mM, pH 8.0) containing NaCl (100 mM) and glycerol (20%, v/v). Protein concentrations were determined by measuring absorbance at 280 nm. The molecular weights (MW) and molar extinction (ε) coefficients used to calculate H_6-protein concentrations are as follows: *Sl*MlaL, 60,362 Da, 92,025 mol^{-1}

min^{-1}; *Sl*LbuL, 59,124 Da, 73,395 mol^{-1} min^{-1}; *Sl*IbaL, 61,510 Da, 101,300 mol^{-1} min^{-1}; and *Sl*PaaL, 49,801 Da, 38,640 mol^{-1} min^{-1}.

In vitro *acyl-CoA synthetase assays*. Activities of *Sl*MlaL, *Sl*LbuL, *Sl*IbaL, and *Sl*PaaL were measured using an NADH-consuming assay (9, 14). *R*eactions (100 μl total volume) contained HEPES buffer (50 mM, pH 7.5), TCEP (1 mM), ATP (2.5 mM) CoA (0.5 mM), $MgCl_2$ (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 either acetate, propionate, butyrate, valerate, hexanoate, heptanoate octanoate, nonanoate, decanoate, undecanoate, laurate, tridecanoate, myristate, oxalate, malonate, succinate, glutarate, adipate, pimelate, suberate, azelate, sebacate, undecanedioate, docdecanedioate, fumarate, acetoacetate, β-hydroxybutyrate, butyrate, isovalarate, crotonate, isobutyrate, methylmalonate, benzoate, methylsuccinate, 3-(methylthio)propionate, 3-phenylpropionate, 4-phenylbutyrate, and 2-phenylpropionate (0.2 mM). Reactions were started by the addition of *Sl*MlaL, *Sl*LbuL, *Sl*IbaL, and *Sl*PaaL (10-100 nM). The absorbance at 340 nm was monitored in a 96-well plate using the Spectramax Plus UV-visible spectrophotometer (Molecular Devices). Enzyme activities were calculated as described (14). Specific activity data are presented with standard deviations from triplicate experiments.

RESULTS

PatA is expressed in a growth-phase dependent manner. To identify acetylation targets of PatA, *S. lividans* TK24 wild-type and Δ*patA* strains were grown aerobically on rich medium (YEME) or minimal medium supplemented with glucose, maltose, β-hydroxybutyrate, glycerol, or *N*-acetylglucosamine and the soluble proteins were probed by Western blot for the presence of PatA and the presence of acetylated lysine residues in resolved proteins (Fig. 4.1). As shown in figure 4.2, neither strain exhibited a growth defect on any carbon source tested. If PatA had acetylation targets in *S. lividans*, we expected to observed differences in acetyl-lysine detection between the *S. lividans* wild-type and Δ*patA* strains.

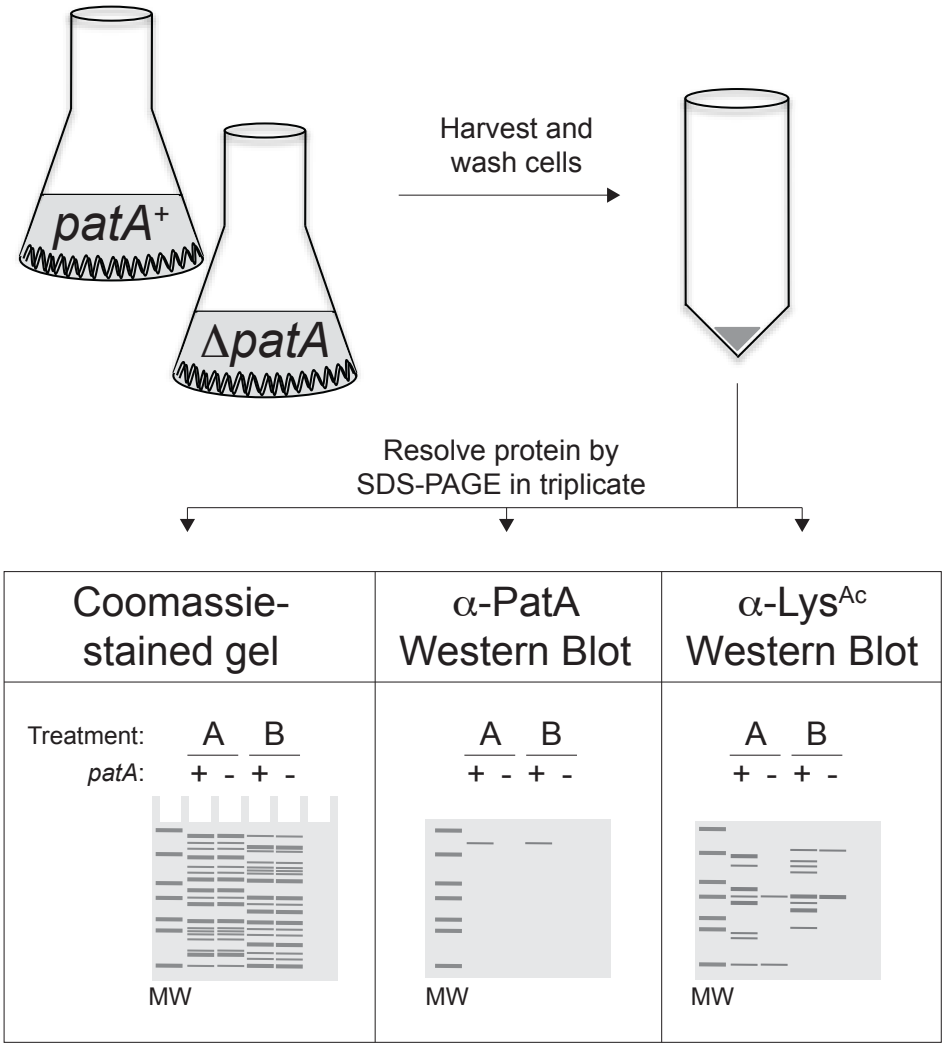


Figure 4.1. Schematic of *S. lividans* cell isolation and Western blot analysis. *S. lividans* was grown in liquid media described in the Experimental Procedures. *S. lividans* was cultured in flasks equipped with marine-grade stainless steel springs to aid in cell dispersal. Cells were harvested by centrifugation and washed before lysis. Cell lysates were clarified by centrifugation and quantified before resolving on triplicate SDS-PAGE gels. One gel was stained with Coomassie Blue to observe proteins and ensure equal loading of lanes. Proteins in the two additional replicate gels were transferred to PVDF membranes. Lysates on one membrane were probed for the presence of PatA using rabbit antiserum. Lysates on the remaining membrane were probed for acetyllysine using anti-acetyllysine antibodies.

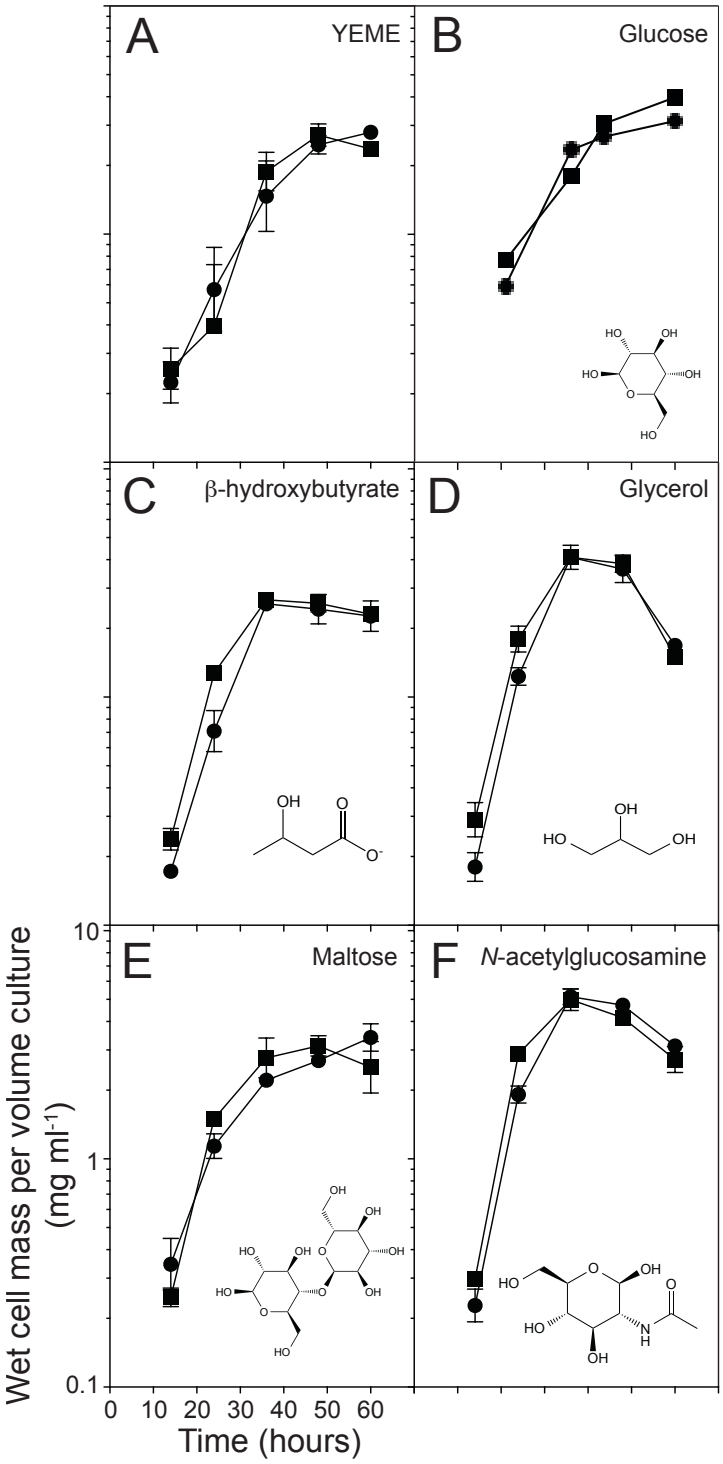


Figure 4.2. Growth of *S. lividans* wild-type and Δ*patA* strains. *S. lividans* wild-type (circles) or Δ*patA* (squares) strains were grown in rich (YEME) medium (A) or minimal liquid media supplemented with glucose (10 mM, B), β-hydroxybutyrate (30 mM, C), glycerol (30 mM, D), maltose (30 mM, E), or *N*-acetylglucosamine (30 mM, F). Cells were harvested at regular intervals by centrifugation. Excess media was removed by pipetting and cell mass was weighed. Growth is reported as wet cell mass per ml of culture. Growth is reported as the mean ± S.D. in (A) and (C-F) or as single value (B).

To correlate the presence of PatA with changes in acetylation profiles, we probed each lysate for PatA and for acetylated lyine residues. As shown in figure 4.3, we can detect PatA in cell lysates when the *S. lividans* wild-type strain is growing in each of the media. *S. lividans* PatA is not detected in the *S. lividans* Δ*patA* strain grown in parallel. Notably, the PatA signal is weaker at the 24-hour time point than at later time points for all conditions. Pat is not detected at the 62-hour time point when the *S. lividans* wild-type strain was grown in media supplemented with β-hydroxybutyrate as the sole carbon source. Similarly, PatA is not detected at the 48- and 62-hour time points when the *S. lividans* wild-type strain was grown on *N*-acetylglucosamine as the sole carbon source. These data suggest that PatA is expressed in a growth phase dependent manner on each carbon source tested.

***PatA controls acetylation of multiple targets in* S. lividans.** When we probed *S. lividans* soluble cellular proteins for acetylated lysine residues, we detect differential banding patterns on each carbon source that correlate with the presence of PatA. It is important to note that we see a high number of proteins acetylated in the absence of PatA in all conditions, suggesting that PatA-independent acetylation occurs in *S. lividans*. As seen in figure 4.3, when *S. lividans* is grown on β-hydroxybutyrate, PatA is detected at 24, 38, and 48 hours. We observed acetylated proteins between 50 and 75 kDa in the *S. lividans* strain that were not detected in the *S. lividans* Δ*patA* strain, suggesting these were targets of PatA. We also detected differences in acetylation patterns when *S. lividans* strains were grown for 62 hours, a time point at which PatA is not detectable by Western blot. This may indicate that undetectable levels of PatA were mediating efficient acetylation. Alternatively, these could be a stable PatA-acetylated protein species with a half-life longer than the PatA protein. We observed similar differences in protein acetylation that correlated with PatA presence when *S. lividans* was grown in minimal medium supplemented with glucose or *N*-acetylglucosamine and when *S. lividans* was grown in rich medium (YEME, Fig. 4.3). Notably, when *S. lividans* was grown on *N*-acetylglucosamine, differences in acetylation patterns between the *S. lividans* wild-type and Δ*patA* strains correlates strongly with the presence of PatA. Significant differences in acetylation are not detected at 48- and 62- hour time points

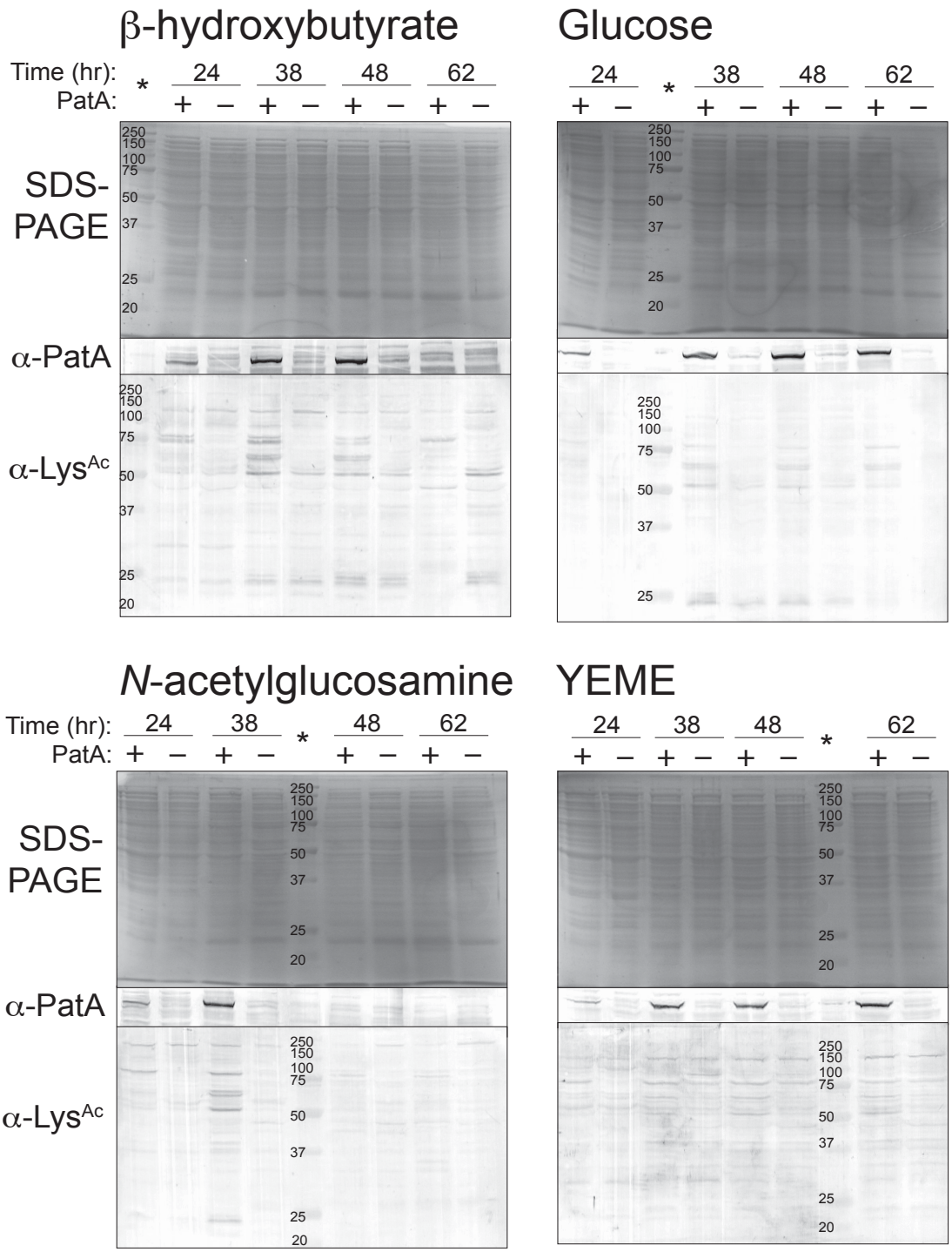


Figure 4.3. Acetylation profiles correlate with PatA in *S. lividans*. *S. lividans* lysates from cells grown on rich medium (YEME) or minimal medium containing Glucose (10 mM), b-hydroxybutyrate (10 mM), or *N*-acetylglucosamine (10 mM) were resolved by SDS-PAGE in triplicate gels (as described in the scheme in figure 4.1). One gel was stained with Coomassie Blue to visualize proteins. The remaining two gels for each condition were transferred to PVDF membranes and probed with antiserum specific for PatA or antibodies specific for acetyllysine. Binding of the primary antibodies was visualized using secondary antibodies conjugated to alkaline phosphatase and NBT/BCIP detection. *, molecular weight standards

when PatA was not detected. Together, these data suggest that PatA controls acetylation of multiple targets in *S. lividans*.

Identification of acetylated proteins in* S. lividans *that are correlated with the production of PatA. We identified the acetylated proteins to determine the substrates of PatA in *S. lividan*s. Soluble proteins from *S. lividans* wild-type and Δ*patA* strains grown on β-hydroxybutyrate were subject to isoelectric focusing followed by SDS-PAGE. Gels were stained to confirm equal loading (Fig. 4.4A,B). *S. lividans* wild-type and Δ*patA* protein extracts were also transferred to PVDF membrane for Western blotting. Acetylated proteins were detected using α-acetyllysine antibodies. We detected 6 distinct acetyllysine spots in lysates of *S. lividans* wild-type strain (Fig. 4.5A,B) that were not detected in lysates of a *S. lividans* Δ*patA* strain (Fig. 4.5C). Acetylated protein spots were excised from the corresponding pI and MW position in the duplicate, coomassie-stained protein gels (Fig. 4.6A,B) for mass peptide fingerprint. Shown in Table 4.4, we identified 4 proteins in 5 of the 6 spots detected. Notably, all are predicted to be members of the AMP-forming CoA ligase family of enzymes. The single acetylated lysine detected on each protein is a predicted catalytic lysine that is conserved among members of the AMP-forming CoA ligase family (15).

Characterization of putative AMP-forming CoA ligases. To determine the substrate specificity of the acetylated AMP-forming CoA ligase enzymes identified by mass spectrometry, we purified each enzyme to homogeneity and measured enzymatic activity in the presence of ATP, CoA and a panel of organic acids listed in the Experimental Procedures (Fig. 4.7). EFD65795 and EFD65796 activate short chain fatty acids. EFD65795 activates several related linear, branched, and unsaturated 4- and 5-carbon fatty acid. Thus, we named this enzyme LbuL for linear, branched, and unsaturated fatty acid CoA ligase. EFD65796 displays a preference for isobutyrate but efficiently activates propionate and butyrate. EFD65796 was named IbuA for isobutyrate CoA ligase. EFD64737 was annotated as a phenylacetic acid CoA ligase (PaaL) and displays a preference for this substrate over related aromatic organic acids.

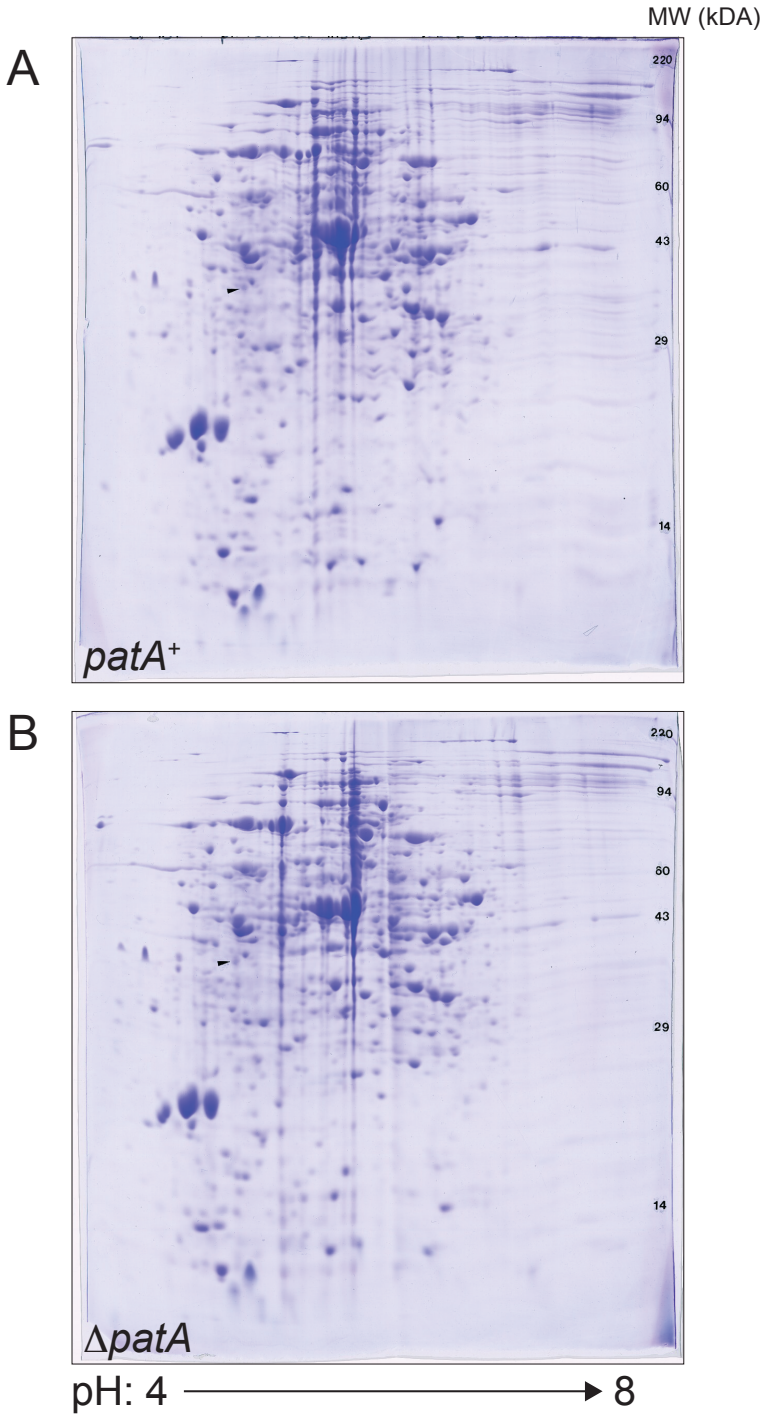


Figure 4.4. Resolution of *S. lividans* wild-type and Δ*patA* strain soluble proteins. Two-dimensional gel electrophoresis was used to resolve the soluble proteins of wild-type (A) and Δ*patA* (B) strains of *S. lividans*. Equal total proteins were loaded for both strainsfor each isoelectric focusing strip (first dimension, pH 4-8). Two technical replicates were run. Shown here is one replicate from each sample stained with Coomassie Blue. The second replicate for each protein sample was transferred to a PVDF membrane for Western blot (see Figure 4.5). MW, molecular weight standards

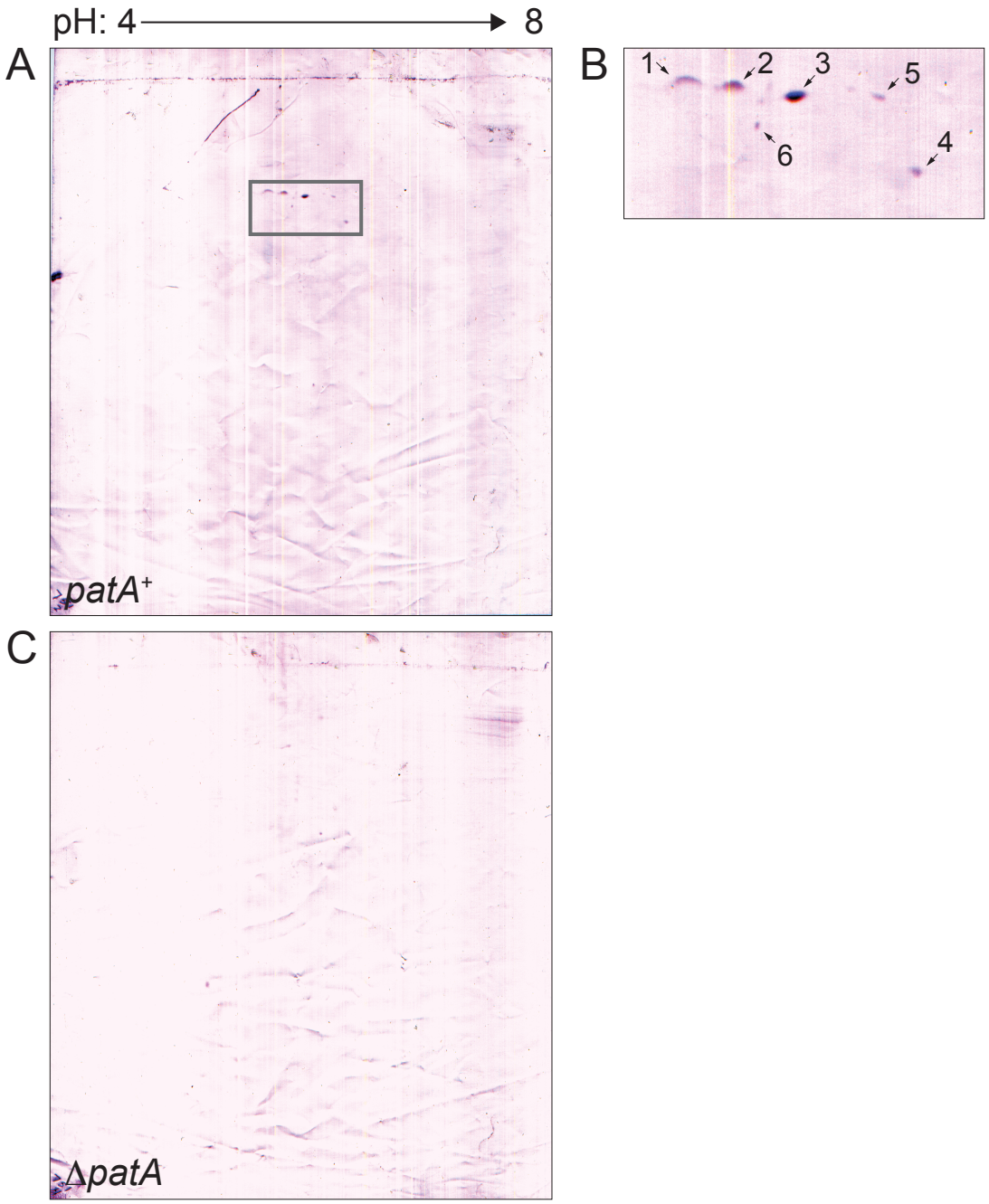


Figure 4.5. Detection of lysine-acetylated proteins in *S. lividans* wild-type and Δ*patA* strain soluble proteins. Identical gels to those described in figure 4.4 were transferred to PVDF membranes for Western blot and probed with anti-acetyllysine antibodies to detect proteins with acetylated lysines. (A) Anti-acetyllysine blot of *S. lividans* wild-type lysates with detected spots surrounded by a black box. (B) The region enclosed in the box in (A) was magnified for clarity and spots were numbered 1-6. (C) Anti-acetyllysine blot of *S. lividans* Δ*patA*.

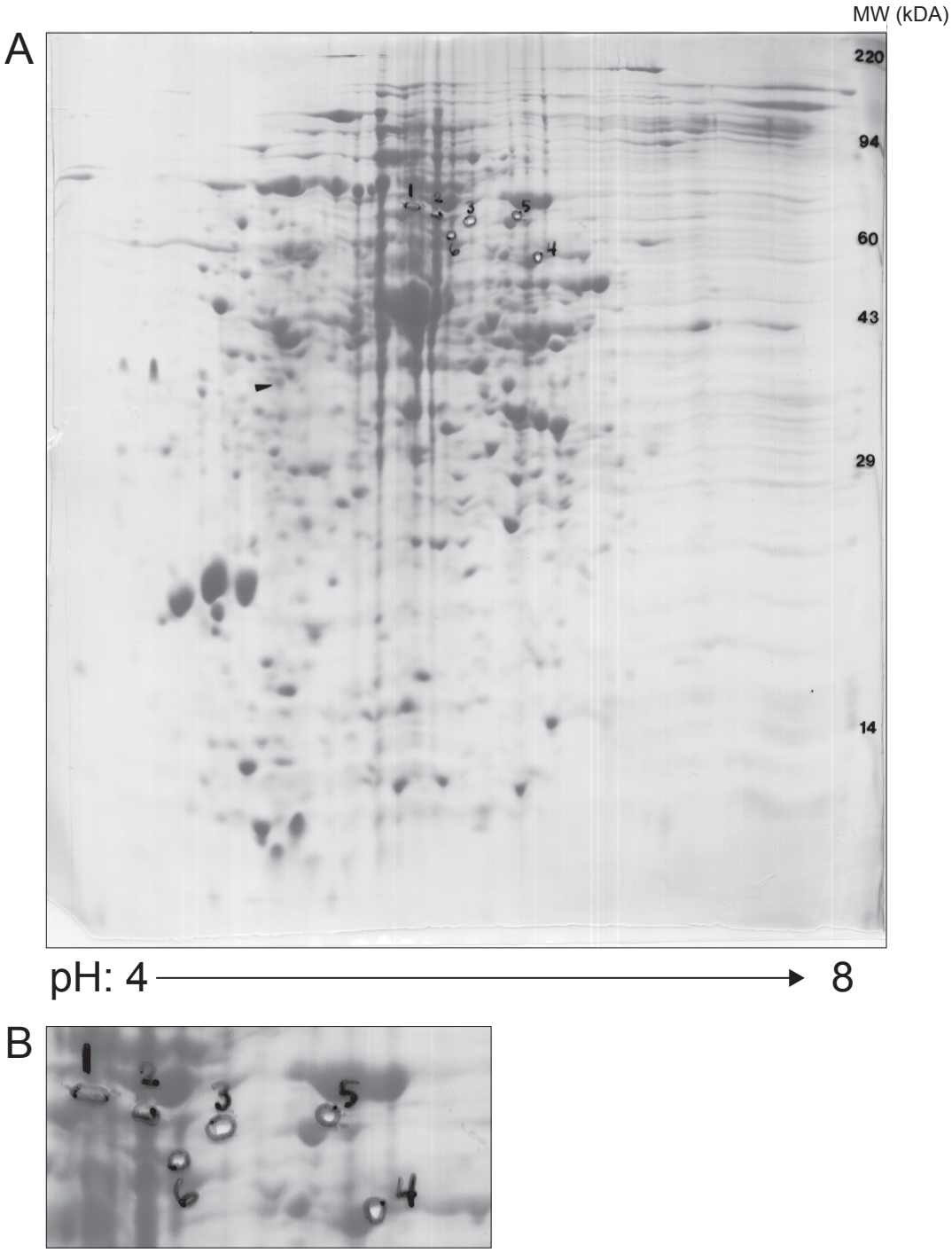


Figure 4.6. Protein spots corresponding to lysine-acetylated proteins were excised after SDS-PAGE. (A) The 6 acetyllysine spots identified by Western blot analysis (Figure 4.5A,B) were excised from the corresponding pI and molecular weight locations on a duplicate gel (Figure 4A). (B) The region corresponding to the location of the spots cut was magnified from clarity. MW, molecular weight standards

Table 4.4. Identification of acetylated proteins.

Accession number	Annoation	Acetylated site	Spot #
EFD65795	Acyl-CoA synthetase	PMTVSG**K**VRKVELRE	3, 5
EFD65796	Acetyl-CoA synthetase	PKTVSG**K**IRRIELRE	1, 2
EFD68037	Acyl-CoA synthetase	PKTSVG**K**FDKKVLRR	3
EFD64737	Phenylacetate-CoA synthetase	LERSLG**K**IRRVWDQR	4

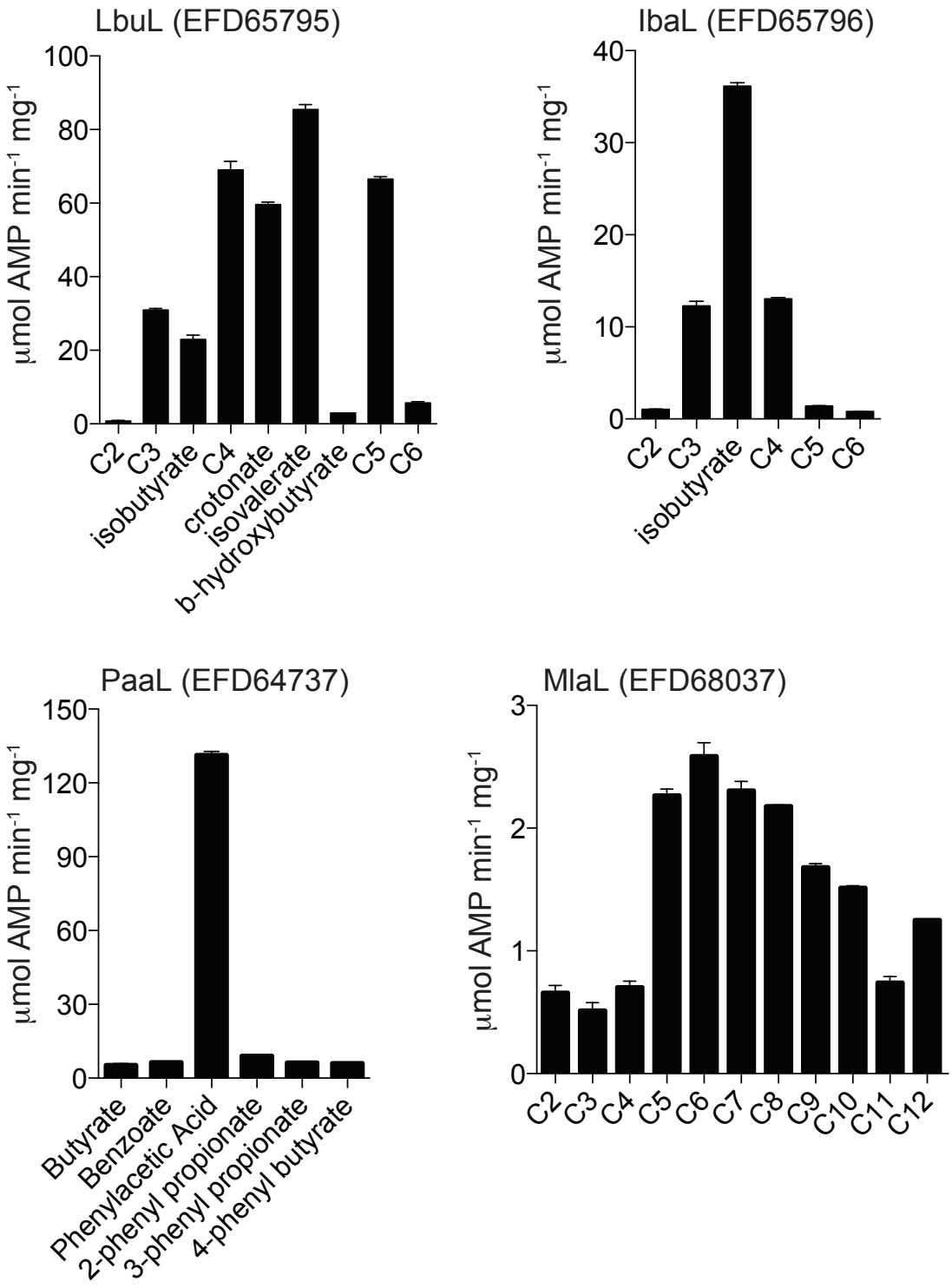


Figure 4.7. Characterization of new AMP-forming CoA ligases. AMP-forming CoA ligases identified as acetylated by mass peptide fingerprinting were tested with a library of organic acid substrates (see Experimental Procedures) to identify the substrate preferences. Activity of each CoA ligase was measured in an NADH comsumption assay. Data represent mean ± S.D. of triplicate experiments.

EFD68037 activates 5- to 12-carbon linear fatty acids and was named MlaL for medium and long chain fatty acid CoA ligase.

DISCUSSION

In this study, we expand the scope of protein lysine acetylation in *S. lividans*. By following the production of PatA protein acetyltransferase during growth of *S. lividans*, we identified 4 proteins whose acetylation was correlated with the presence of PatA *in vivo*. Characterization of the acetylation targets revealed that these proteins were members of the AMP-forming CoA ligase family, and each enzyme recognized a specific subset of organic acids. These finding suggest that PatA regulates activity of enzymes involved in CoA activation in *S. lividans.*

PatA is growth-phase regulated. Knowledge of PatA protein production is limited. Expression of the gene encoding the *E. coli* PatA homolog (Pka) increases 10-fold as *E. coli* transitions from exponential phase into stationary phase when growing on glucose (6). However, expression remains constant when cells are growing on acetate. Here we report that production of PatA is growth phase dependent in *S. lividans* and is maximal when cells are in stationary phase. We observed a significant decrease in PatA levels 12-24 hours after *S. lividans* reached stationary phase when growing in minimal medium supplemented β-hydroxybutyrate or *N*-acetylglucosamine. This suggests that PatA acts during early stationary phase to regulate its targets, including AMP-forming CoA ligases. It is possible that similar decreases in PatA levels may have been observed at time points >62 hr for cells growing in rich medium or in glucose minimal medium.

***PatA regulates AMP-forming CoA ligases in* S. lividans.** Members of this family of enzymes are acetylated on their respective catalytic lysine residues by homologs of PatA in *Salmonella enterica* and *Rhodopseudomonas palustris* (9, 10, 35). AMP-forming CoA ligases are acetylated by GNATs in *Bacillus subtilis* and *Mycobacterium* sp (13, 28). Our study identified 4 AMP-forming CoA ligases that are regulated by *Sl*PatA. Substrates were identified for each of the AMP-forming CoA ligases identified in this study. IbuA and PaaL demonstrated high substrate specificity for a small subset of organic acids

while LbuL and MlaL exhibited broader substrate range. Because MlaL exhibited strikingly lower specific activity with all substrates tested, it is unclear whether we have identified the optimal substrate for this enzyme. Taken together, we have confirmed that each of these proteins is an AMP-forming CoA ligase.

Acetylation of IbuA, PaaL, LbuL, and Mla was detected at a single lysine residue in each protein. Because these were the most abundant acetylation events detected, we hypothesize that *S. lividans* PatA promotes efficient acetylation of AMP-forming CoA ligases *in vivo*. Due to sample preparation and the limit of detection of our Western blotting, we cannot rule out that PatA is mediating efficient acetylation of low-abundance proteins or insoluble proteins that we could not identify with this technique. Proteomic analyses of *S. lividans* at each time point in each growth medium may reveal new acetylation targets that were not detected in this study.

Are the acetylation targets directly modified by PatA? As described above, PatA homologs acetylate AMP-forming CoA ligases in bacteria. Thus, we hypothesize that PatA is directly acetylating the AMP-forming CoA ligases identified in this study. It is possible that PatA indirectly regulates acetylation of these targets through other regulators (e.g. transcriptional regulators, protein acetyltransferases). Direct, *in vitro* evidence of PatA-dependent acetylation would provide support for our hypothesis.

Is activity of the acetylation targets regulated by reversible lysine acetylation? Acetylation of the conserved lysine in AMP-forming CoA ligases results in a loss of activity (34-36). However, it is possible that acetylation of these AMP-forming CoA ligases alters activity or substrate specificity by a new mechanism. *In vitro* acetylation of these targets would clarify the role of acetylation of these AMP-forming CoA ligases. Deacetylation by sirtuin (NAD^+-dependent) deacetylases or histone deacetylases (NAD^+-independent) reactivates members of the AMP-forming CoA ligase family (9, 10, 13, 34). *S. lividans* encodes 2 sirtuins and 1 histone deacetylase (39). Deacetylation of the targets identified in this study would reveal that the modifications identified *in vivo* are reversible and that deacetylation reverses the effects of acetylation.

REFERENCES

1. **Berndsen, C. E., and J. M. Denu.** 2008. Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol.* **18:**682-689.

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

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

4. **Brownell, J. E., and C. D. Allis.** 1995. An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena macronuclei*. *Proc. Natl. Acad. Sci. U S A* **92:**6364-6368.

5. **Burgess-Cassler, A., J. J. Johansen, D. A. Santek, J. R. Ide, and N. C. Kendrick.** 1989. Computerized quantitative analysis of coomassie-blue-stained serum proteins separated by two-dimensional electrophoresis. *Clin. Chem.* **35:**2297-2304.

6. **Castano-Cerezo, S., V. Bernal, J. Blanco-Catala, J. L. Iborra, and M. Canovas.** 2011. cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in *Escherichia coli*. *Mol. Microbiol*. **82:**1110-1128.

7. **Chater, K. F.** 2006. *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. **361:**761-768.

8. **Courtois, S., C. M. Cappellano, M. Ball, F. X. Francou, P. Normand, G. Helynck, A. Martinez, S. J. Kolvek, J. Hopke, M. S. Osburne, P. R. August, R. Nalin, M. Guerineau, P. Jeannin, P. Simonet, and J. L. Pernodet.** 2003. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol.* **69:**49-55.

9. **Crosby, H. A., E. K. Heiniger, C. S. Harwood, and J. C. Escalante-Semerena.** 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.

10. **Crosby, H. A., D. A. Pelletier, G. B. Hurst, and J. C. Escalante-Semerena.** 2012. System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J. Biol. Chem.* **287:**15590-15601.

11. **Crosby, H. A., K. C. Rank, I. Rayment, and J. C. Escalante-Semerena.** 2012. Structure-guided expansion of the substrate range of methylmalonyl-CoA synthetase (MatB) of *Rhodopseudomonas palustris*. *Appl. Environ. Microbiol.* **78:**6619-6629.

12. **Elion, E. A., P. Marina, and L. Yu.** 2007. Constructing recombinant DNA molecules by PCR, p. Unit 3.17.1-3.17.12. *In* F. M. Ausubel, R. E. R. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, a. Smith, and K. Struhl (ed.), *Current protocols in molecular biology*. Greene Publishing Associates & Wiley Interscience, New York, N.Y.

13. **Gardner, J. G., F. J. Grundy, T. M. Henkin, and J. C. Escalante-Semerena.** 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD(+) involvement in *Bacillus subtilis*. *J. Bacteriol.* **188:**5460-5468.

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

15. **Gulick, A. M.** 2009. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* **4:**811-827.

16. **Hallows, W. C., S. Lee, and J. M. Denu.** 2006. Sirtuins deacetylate and activate mammalian acetyl-CoA synthetases. *Proc. Natl. Acad. Sci. U S A* **103:**10230-10235.

17. **Kieser, T., M. J. Bibb, M. J. Buttner, K. Chater, and D. A. Hopwood.** 2000. Growth and preservation of *Streptomyces*, p. 43-62, *Practical* Streptomyces *Genetics*. John Innes Foundation, Norwich, England.

18. **Kieser, T., M. J. Bibb, M. J. Buttner, K. Chater, and D. A. Hopwood.** 2000. Media, buffers, and suppliers, p. 405-420, *Practical* Streptomyces *Genetics*. John Innes Foundation, Norwich, England.

19. **Kim, G. W., and X. J. Yang.** 2011. Comprehensive lysine acetylomes emerging from bacteria to humans. *Trends. Biochem. Sci.* **36:**211-220.

20. **Kouzarides, T.** 2007. Chromatin modifications and their function. *Cell* **128:**693-705.

21. **Kuo, Y. M., and A. J. Andrews.** 2013. Quantitating the specificity and selectivity of Gcn5-mediated acetylation of histone H3. *PLoS One* **8:**e54896.

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

23. **Liang, W., and M. P. Deutscher.** 2010. A novel mechanism for ribonuclease regulation: transfer-messenger RNA (tmRNA) and its associated protein SmpB regulate the stability of RNase R. *J. Biol. Chem.* **285:**29054-8.

24. **Liang, W., A. Malhotra, and M. P. Deutscher.** 2011. Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. *Mol. Cell* **44:**160-166.

25. **Marsh, V. L., S. Y. Peak-Chew, and S. D. Bell.** 2005. Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba. *J. Biol. Chem*. **280:**21122-21228.

26. **McMahon, M. D., C. Guan, J. Handelsman, and M. G. Thomas.** 2012. Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. *Appl. Environ. Microbiol.* **78:**3622-3629.

27. **Mikulik, K., J. Felsberg, E. Kudrnacova, S. Bezouskova, D. Setinova, E. Stodulkova, J. Zidkova, and V. Zidek.** 2012. CobB1 deacetylase activity in *Streptomyces coelicolor*. *Biochem. Cell Biol.* **90:**179-187.

28. **Nambi, S., K. Gupta, M. Bhattacharya, P. Ramakrishnan, V. Ravikumar, N. Siddiqui, A. T. Thomas, and S. S. Visweswariah.** 2013. Cyclic AMP-dependent protein lysine acylation in mycobacteria regulates fatty acid and propionate metabolism. *J. Biol. Chem.* **288**.

29. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250:**4007-4021.

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

31. **Seow, K. T., G. Meurer, M. Gerlitz, E. Wendt-Pienkowski, C. R. Hutchinson, and J. Davies.** 1997. A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J. Bacteriol*. **179:**7360-7368.

32. **Shirling, E. B., and D. Gottlieb.** 1966. Methods for characterization of *Streptomyces* species. *Int. J. Syst. Bacteriol.* **16:**313-340.

33. **Soppa, J.** 2010. Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea* **2010:**pii: 820681.

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

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

36. **Starai, V. J., H. Takahashi, J. D. Boeke, and J. C. Escalante-Semerena.** 2003. Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae*. *Genetics* **163:**545-555.

37. **Thao, S., C. S. Chen, H. Zhu, and J. C. Escalante-Semerena.** 2010. N(epsilon)-Lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. *PLoS ONE* **5:**e15123.

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

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

40. **Wardleworth, B. N., R. J. Russell, S. D. Bell, G. L. Taylor, and M. F. White.** 2002. Structure of Alba: an archaeal chromatin protein modulated by acetylation. *EMBO J.* **21:**4654-4662.

41. **Weinert, B. T., V. Iesmantavicius, S. A. Wagner, C. Scholz, B. Gummesson, P. Beli, T. Nystrom, and C. Choudhary.** 2013. Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Mol. Cell* **51:**265-272.

CHAPTER 5

IDENTIFICATION OF DETERMINANTS WITHIN THE *C*-TERMINAL DOMAIN OF *STREPTOMYCES LIVIDANS* ACETYL-COA SYNTHETASE THAT BLOCK ACETYLATION OF ITS ACTIVE SITE LYSINE *IN VITRO*[1]

[1] Tucker A.C. and J.C. Escalante-Semerena. To be submitted to *Molecular Microbiology*.

ABSTRACT

Reversible lysine acetylation (RLA) is a widespread regulatory mechanism that targets proteins in diverse cellular processes. A strong case has been made for the RLA control exerted by Pat homologues on the broadly distributed AMP-forming acid:CoA ligase (a.k.a. acyl-CoA synthetases) family of metabolic enzymes, with acetyl-CoA synthetase (Acs) being the paradigm in the field. Here we investigate why the Acs homologue in *Streptomyces lividans* is poorly acetylated *in vitro* by the *S. lividans* protein acetyltransferase (Pat) homologue, *Sl*PatA. *Se*Acs-*Sl*Acs chimeras constructed during the course of this work were acetylated by *Sl*PatA *in vitro*, retained most of their activity, and were under RLA control in a heterologous host. We identified *Se*Acs residues *N*- and *C*-terminal to the target lysine that when introduced into *Sl*PatA, rendered the latter under RLA control. These results lend further support to the idea that Pat proteins interact with extensive surfaces of their substrates. Finally, we suggest that acetylation of *Sl*Acs depends on factors or conditions other than those present in our *in vitro* system. We also discuss possible explanations why *Sl*Acs is not controlled by RLA as defined in other bacterial species.

INTRODUCTION

Reversible lysine acetylation (RLA) is a post-translational modification that occurs in all domains of life (28) and affects diverse cellular pathways and functions. Acetyltransferases transfer the acetyl moiety from acetyl-CoA to the ε-amino group of the target lysine. Lysine acetylation can affect enzyme activity (29), protein stability (21), and protein-protein interactions, or DNA binding (17). Yeast Gcn5 protein (yGcn5p)-related *N*-acetyltransferases (aka, GNATs), classified by amino acid sequence and structure (37), are the only class of acetyltransferases found in all domains of life (34). GNATs were first identified for their role in modification of histones (4). Crystal structures and biochemical analyses of the *Tetrahymena* Gcn5, the founding member of the GNAT family, with representative peptides from histones has provided valuable information about the substrate specificity and substrate recognition by GNATs (18, 26).

Members of the GNAT family also acetylate metabolic enzymes. For example, in *Salmonella enterica*, the enzyme acetyl-CoA synthetase (*Se*Acs) is acetylated by the protein acetyltransferase (*Se*Pat), a two-domain acetyltransferase that contains a large domain of unknown function and a *C*-terminal GNAT domain (32). *Se*Acs is a member of the AMP-forming CoA ligase family of enzymes that converts fatty acids to their CoA thioesters via an acyl-AMP intermediate (30). Acetylation of the active site lysine of AMP-forming CoA ligases prevents the adenylylation of the fatty acid. In addition to Pat from *S. enterica*, GNATs are known to acetylate members of the of AMP-forming CoA ligase family (including Acs) in *Rhodopseudomonas palustris* (6, 7), *Bacillus subtilis* (10), and *Mycobacterium smegmatis* (38). The Acs homologue from *Streptomyces coelicolor* is acetylated *in vivo* (22), but the GNAT responsible for acetylation of *S. coelicolor* Acs is unknown.

Knowledge of the interactions of GNATs with their proteins substrates is limited. *R. palustris* encodes a single-domain GNAT (KatA) and homologue of the *Se*Pat GNAT (*Rp*Pat). KatA and *Rp*Pat discriminate between members of the AMP-forming CoA ligase family produced by *R. palustris* (7). In addition to the target lysine, *Rp*Pat recognizes a loop greater than 20 Å from the target lysine, suggesting that Pat enzymes interact with a large surface of the acceptor substrate (8). Introducing this recognition loop into *R. palustris* methylmalonyl-CoA mutase (*Rp*MatB), an AMP-forming CoA ligase that was not a substrate of *Rp*Pat, rendered *Rp*MatB a target of acetylation by *Rp*Pat. Thus, synthetic chimeras of AMP-forming CoA ligases have yielded information about how GNATs recognize protein substrates and have produced AMP-forming CoA ligases that are placed under the regulation of lysine acetylation.

*Rp*Pat and *Se*Pat enzymes acetylate their cognate Acs proteins. Although the GNAT responsible for the acetylation of Acs is *S. coeolicolor* is unknown, the closely related actinomycete *Streptomyces lividans* encodes *Sl*PatA, a homologue of SePat and RpPat enzymes. *Sl*PatA does not efficiently acetylate the *S. lividans* Acs *in vitro* (35), making this the first Acs enzyme that is not acetylated by the Pat homologue from the same organism. *Sl*PatA efficiently acetylates Acs from *S. enterica*. Here we probe the substrate specificity of *Sl*PatA by determining the elements in *S. enterica* Acs (*Se*Acs) that rendered it a better substrate than *S. lividans* Acs (*Sl*Acs). By swapping amino acids from *Se*Acs into the *C*-terminus

of *Sl*Acs, we constructed a *Sl*Acs-*Se*Acs chimera that is efficiently acetylated by *Sl*PatA. The *Sl*Acs-*Se*Acs chimera contains 41 amino acid differences from *Sl*Acs. The *Sl*Acs-*Se*Acs chimera is subject to 40-fold more regulation than *Sl*Acs *in vitro*. In a heterologous model system, we also demonstrate that the *Sl*Acs-*Se*Acs chimera is subject to greater regulation *in vivo* by *Sl*PatA than is *Sl*Acs. In sum, we have identified regions in *Se*Acs that are critical for recognition by *Sl*PatA, and transferring these residues into the poor substrate *Sl*Acs resulted in a *Sl*Acs variant that is now subject to efficient regulation by *Sl*PatA in isolation.

EXPERIMENTAL PROCEDURES

Bacterial strains and growth conditions. All strains and plasmids used in this study are listed in Tables 5.1 and 5.2, respectively. *Escherichia coli* and *Salmonella enterica* strains were grown at 37°C in lysogeny broth (LB, Difco) (2) or no-carbon essential (NCE) minimal medium (1) supplemented with sodium acetate (10 mM), $MgSO_4$ (1 mM), and ampicillin (100 μg ml^{-1}). When necessary, antibiotics were used at the following concentrations: ampicillin, 100 μg ml^{-1}; tetracycline, 10 μg ml^{-1}; chloramphenicol, 12.5 μg ml^{-1}, kanamycin, 50 μg ml^{-1}. L-(+)-arabinose was added at varying concentrations (5 or 200 μM) to induce the expression of *S. enterica acs*, *S. lividans acs*, and *acs* chimeras were cloned into the expression vector pBAD30 (14). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of IPTG (0-500 μM) to induce expression of *S. lividans patA* (EFD66247) clones into the expression vector pSRK-Km.

Molecular techniques. DNA manipulations were performed using standard techniques (9). 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

Table 5.1. Plasmids used in this study.

Plasmid	Primers	Genotype	Source or method
pBAD30		P_{araBAD} expression vector, *bla*$^+$	(14)
p*Sl*Acs6	*Sl*Acs pBAD30 5' *Sl*Acs pBAD30 3'	*S. lividans acs* allele (EFD66247) in pBAD30, *bla*$^+$	Traditional cloning
p*Sl*Acs47	*Sl*Acs-His pBAD30 5' *Sl*Acs-His pBAD30 3'	*S. lividans acs* allele (EFD66247) with *N*-terminal H_6 tag in pBAD30, *bla*$^+$	Traditional cloning
p*Sl*Acs48	*Sl*Acs-His pBAD30 5' *Sl*Acs-His pBAD30 3'	*S. lividans acs –S. enterica acs* chimera allele with *N*-terminal H_6 tag in pBAD30, *bla*$^+$	Traditional cloning
pACS59	*Se*Acs-His pBAD30 5' *Se*Acs-His pBAD30 3'	*S. enterica acs* allele with *N*-terminal H_6 tag in pBAD30, *bla*$^+$	Traditional cloning
pSRK-Km		*lacI*^{q}*-lac* promoter-operator expression vector, *kan*$^+$	(16)
p*Sl*PatA9		*S. lividans patA*$^+$ allele (EFD66247) in pSRK-Km, *bla*$^+$	(35)
pTEV5		*N*-terminal, rTEV-cleavable His_6-tag overexpression vector, *bla*$^+$	(24)
p*Sl*Acs1		*S. lividans acs*$^+$ allele (EFD68454) in pTEV5, *bla*$^+$	(35)
p*Sl*PatA1		*S. lividans patA*$^+$ allele (EFD66247) in pTEV5, *bla*$^+$	(35)
p*Sl*Acs8	*Sl*Acs 5' + NheI *Sl*Acs520*Se*Acs F *Sl*Acs520*Se*Acs R EcoRI + *Se*Acs 3'	A1 chimera: *Sl*Acs 520 *Se*Acs in pTEV5, *bla*$^+$	Overlap-extension PCR, traditional cloning
p*Sl*Acs9	*Sl*Acs 5' + NheI *Sl*Acs550*Se*Acs F *Sl*Acs550*Se*Acs R EcoRI + *Se*Acs 3'	A2 chimera: *Sl*Acs 550 *Se*Acs in pTEV5, *bla*$^+$	Overlap-extension PCR, traditional cloning
p*Sl*Acs12	*Sl*Acs 5' + NheI *Sl*Acs560*Se*Acs F *Sl*Acs560*Se*Acs R EcoRI + *Se*Acs 3'	A3 chimera: *Sl*Acs 560 *Se*Acs in pTEV5, *bla*$^+$	Overlap-extension PCR, traditional cloning
p*Sl*Acs22	*Sl*Acs 5' + NheI *Sl*Acs566*Se*Acs F *Sl*Acs566*Se*Acs R EcoRI + *Se*Acs 3'	A4 chimera: *Sl*Acs 566 *Se*Acs in pTEV5, *bla*$^+$	Overlap-extension PCR, traditional cloning
p*Sl*Acs10	*Sl*Acs 5' + NheI *Sl*Acs582*Se*Acs F *Sl*Acs582*Se*Acs R2 EcoRI + *Se*Acs 3'	A5 chimera: *Sl*Acs 582 *Se*Acs in pTEV5, *bla*$^+$	Overlap-extension PCR, traditional cloning
p*Sl*Acs11	*Sl*Acs 5' + NheI *Sl*Acs617*Se*Acs F *Sl*Acs617*Se*Acs R2 EcoRI + *Se*Acs 3'	A6 chimera: *Sl*Acs 617 *Se*Acs in pTEV5, *bla*$^+$	Overlap-extension PCR, traditional cloning

p*Sl*Acs14	*Sl*Acs 5' + NheI *Se*Acs582*Sl*Acs F *Se*Acs582*Sl*Acs R *Sl*Acs 3' + EcoRI	B1 chimera: *Sl*Acs 550-582 *Se*Acs in pTEV5, bla^+ in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs15	*Sl*Acs 5' + NheI *Se*Acs603*Sl*Acs F *Se*Acs603*Sl*Acs R *Sl*Acs 3' + EcoRI	B2 chimera: *Sl*Acs 550-603 *Se*Acs in pTEV5, bla^+ in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs23	*Sl*Acs 5' + NheI *Se*Acs618*Sl*Acs F *Se*Acs618*Sl*Acs R *Sl*Acs 3' + EcoRI	B3 chimera: *Sl*Acs 550-618 *Se*Acs in pTEV5, bla^+ in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs17	*Sl*Acs 5' + NheI *Se*Acs627*Sl*Acs F *Se*Acs627*Sl*Acs R *Sl*Acs 3' + EcoRI	B4 chimera: *Sl*Acs 550-627 *Se*Acs in pTEV5, bla^+ in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs18	*Sl*Acs 5' + NheI *Se*Acs638*Sl*Acs R	B5 chimera: *Sl*Acs 550-638 *Se*Acs in pTEV5, bla^+ in pTEV5, bla^+	Traditional cloning
p*Sl*Acs19	*Sl*Acs 5' + NheI *Se*Acs643*Sl*Acs R	B6 chimera: *Sl*Acs 550-643 *Se*Acs in pTEV5, bla^+ in pTEV5, bla^+	Traditional cloning
p*Sl*Acs26	*Sl*Acs 5' + NheI *Se*Acs582*Sl*Acs589*Se*Acs F *Se*Acs582*Sl*Acs589*Se*Acs R *Sl*Acs 3' + EcoRI	C1 chimera: *Sl*Acs 550-581 *Se*Acs, 591-627 *Se*Acs in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs27	*Sl*Acs 5' + NheI *Se*Acs590*Sl*Acs597*Se*Acs F *Se*Acs590*Sl*Acs597*Se*Acs R *Sl*Acs 3' + EcoRI	C2 chimera: *Sl*Acs 550-590 *Se*Acs, 598-627 *Se*Acs in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs28	*Sl*Acs 5' + NheI *Se*Acs597*Sl*Acs603*Se*Acs F *Se*Acs597*Sl*Acs603*Se*Acs R *Sl*Acs 3' + EcoRI	C3 chimera: *Sl*Acs 550-597 *Se*Acs, 603-627 *Se*Acs in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs29	*Sl*Acs 5' + NheI *Se*Acs582*Sl*Acs603*Se*Acs F *Se*Acs582*Sl*Acs603*Se*Acs R *Sl*Acs 3' + EcoRI	C4 chimera: *Sl*Acs 550-581 *Se*Acs, 603-627 *Se*Acs in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs44	*Sl*Acs 5' + NheI 7.2 F 7.2 R *Sl*Acs 3' + EcoRI	C5 chimera: *Sl*Acs 615-626 *Se*Acs in pTEV5, bla^+	Overlap-extension PCR, traditional cloning
p*Sl*Acs49	*Sl*Acs K610A F *Sl*Acs K610A R	K610A variant of C3 chimera in pTEV5, bla^+	Site-directed mutagenesis

Table 5.2. Strains used in this study.

Strain	Relevant Genotype and description	Source
S. enterica strains		
JE6583	*metE205 ara-9*	Lab Collection
Derivatives of JE6583		
JE9152	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc)	Lab Collection
JE9894	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc)	Lab Collection
JE13238	*metE205 ara-9* Δ*acs2* Δ*pta127*	Lab Collection
Derivatives of JE9152		
JE18793	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc) / pBAD30 pSRK-Km	
JE18794	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc) / pBAD30 p*Sl*PatA9	
JE18795	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc) / p*Sl*Acs47 pSRK-Km	
JE18796	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc) / p*Sl*Acs47 p*Sl*PatA9	
JE18797	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc) / p*Sl*Acs48 pSRK-Km	
JE18798	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc) / p*Sl*Acs48 p*Sl*PatA9	
JE18799	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc) / pACS59 pSRK-Km	
JE18800	*metE205 ara-9* Δ*acs2* Δ*cobB1330 pat1*::Tn10d(Tc) / pACS59 p*Sl*PatA9	
Derivatives of JE9894		
JE18801	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc) / pBAD30 pSRK-Km	
JE18802	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc) / pBAD30 p*Sl*PatA9	
JE18803	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc) / p*Sl*Acs47 pSRK-Km	
JE18804	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc) / p*Sl*Acs47 p*Sl*PatA9	
JE18805	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc) / p*Sl*Acs48 pSRK-Km	
JE18806	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc) / p*Sl*Acs48 p*Sl*PatA9	
JE18807	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc) / pACS59 pSRK-Km	
JE18808	*metE205 ara-9* Δ*acs2 pat1*::Tn10d(Tc) / pACS59 p*Sl*PatA9	
Derivatives of JE13238		
JE13787	*metE205 ara-9* Δ*acs2* D*pta127* / pBAD30 bla^{+}	
JE14947	*metE205 ara-9* Δ*acs2* D*pta127* / p*Sl*Acs6 bla^{+}	
E. coli strain		
JE9314	C41l(DE3) *pka12*::*kan*$^{+}$	Lab Collection

Unless otherwise noted, all strains were conducted during the course of this study.

PRISM) protocols, and sequencing reactions were resolved at the University of Georgia Genomics Facility. Oligonucleotide primer sequences are listed in Table 5.3.

Plasmids used for protein overproduction. Chimeras of *S. lividans acs* (EFD68454) and *S. enterica acs* genes were amplified from genomic DNA from *S. lividans* TK24 and genomic DNA from *S. enterica* strain TR6583, respectively, with the primers listed in Table 5.3 as described below. The first codon of *Sl*Acs (GTG) was changed to the more common ATG start codon. The DNA fragments were digested with NheI and EcoRI and ligated into pTEV5 (25) cut with the same enzymes. The resulting plasmids p*Sl*Acs8, -9, -10, -11, -12, -14, -15, -17, -18, -19, -22, -23, -26, -27, -28, -29, -44, -49 direct synthesis of *Sl*Acs-*Se*Acs chimeras with *N*-terminal H_6 tags cleavable by recombinant tobacco etch virus (rTEV) protease prepared as described (3).

Construction of chimera plasmids. Fusion plasmids in which the *N*-terminal domain of *Sl*Acs was fused to the *C*-terminal domain of *Se*Acs at residues 520, 550, 560, 566, 582, 617 were generated by overlap-extension PCR (15), followed by traditional cloning into pTEV5. Fusion plasmids in which an internal sequence of *Sl*Acs was replaced by the corresponding sequence *Se*Acs were constructed as follows:

p*Sl*Acs14 (*Sl*Acs 550-582 *Se*Acs) – the nucleotides encoding the first 582 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9, fused to the *C*-terminus of *Sl*Acs, and cloned into pTEV5

p*Sl*Acs15 (*Sl*Acs 550-603 *Se*Acs) – the nucleotides encoding the first 603 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9, fused to the *C*-terminus of *Sl*Acs, and cloned into pTEV5

p*Sl*Acs23 (*Sl*Acs 550-618 *Se*Acs) – the nucleotides encoding the first 618 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9, fused to the *C*-terminus of *Sl*Acs, and cloned into pTEV5

p*Sl*Acs17 (*Sl*Acs 550-627 *Se*Acs) – the nucleotides encoding the first 627 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9, fused to the *C*-terminus of *Sl*Acs, and cloned into pTEV5

p*Sl*Acs18 (*Sl*Acs 550-638 *Se*Acs) – the nucleotides encoding the first 638 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9 and cloned into pTEV5

Table 5.3. Primers used in this study.

Primer	Sequence
*Sl*Acs pBAD30 5’	ATAGAATTCCCTACAAGGAGAACAACAGCATGAGCAACGAA TCCTTGGCCAAC
*Sl*Acs pBAD30 3’	ATGAAGCTTGCCTGTGGCCGGCTCAGTC
*Sl*Acs-His pBAD30 5’	ATAGAATTCCCTACAAGGAGAACAACAGCATGCACCATCAC CATCACCATGGCAGCGGCATGAGCAACGAATCCTTGGCCAAC
*Sl*Acs-His pBAD30 3’	ATGAAGCTTGCCTGTGGCCGGCTCAGTC
*Se*Acs-His pBAD30 5’	ATAGAATTCCCTACAAGGAGAACAACAGCATGCACCATCAC CATCACCATGGCAGCGGCATGAGCCAAACACATAAACACGC
*Se*Acs-His pBAD30 3’	ATGAAGCTTTTATGACGGCATCGCGATGGCCTG
*Sl*Acs 5’ + NheI	GTAGCTAGCATGAGCAACGAATCCTTGGCCAAC
*Sl*Acs 3’ + EcoRI	ACTGAATTCGCCTGTGGCCGGCTCAGTC
EcoRI + *Se*Acs 3’	ATTGAATTCTTATGACGGCATCGCGATGG
*Sl*Acs520*Se*Acs F	GCCGCGTCGACGACGTGTTAAACGTCTCCGGCCACCG
*Sl*Acs520*Se*Acs R	CGGTGGCCGGAGACGTTTAACACGTCGTCGACGCGGCC
*Sl*Acs550*Se*Acs F	GAGGCGGCCGTGGTCGGCATTCCACACGCTATCAAAGGC
*Sl*Acs550*Se*Acs R	GCCTTTGATAGCGTGTGGAATGCCGACCACGGCCGCCTC
*Sl*Acs560*Se*Acs F	GAGACGACCGGGCAGGCCATCTACGCTTATGTGACGCTCAAC
*Sl*Acs560*Se*Acs R	GTTGAGCGTCACATAAGCGTAGATGGCCTGCCCGGTCGTCTC
*Sl*Acs566*Se*Acs F3	ATTTACGCTTATGTGACGCTCCGCGGCACCACCGCGGAGAGC GAGGACCTGGTCGCCGAG
*Sl*Acs566*Se*Acs R	GCTCTCCGCGGTGGTGCCGCGGAGCGTCACATAAGCGTAAAT
*Sl*Acs582*Se*Acs F	GTCGCCGAGCTGCGCAACTGGGTACGTAAAGAGATTGG
*Sl*Acs582*Se*Acs R2	CCAATCTCTTTACGTACCCAGTTGCGCAGCTCGGCGAC
*Sl*Acs617*Se*Acs F	CATGCGGCGCCTGCTGCGAAAAATCGCGGCGGGCGATACCAGC
*Sl*Acs617*Se*Acs R2	GCTGGTATCGCCCGCCGCGATTTTTCGCAGCAGGCGCCGCATG
*Se*Acs582*Sl*Acs F	GAACTGTACGCGGAGGTGCGCAACCACGTCGGTGCCACGCTC
*Se*Acs582*Sl*Acs R	GAGCGTGGCACCGACGTGGTTGCGCACCTCCGCGTACAGTTC
*Se*Acs603*Sl*Acs F	GTGCTGCACTGGACCGACTCACTGCCGAAGACCCGCTCCGGC
*Se*Acs603*Sl*Acs R	GCCGGAGCGGGTCTTCGGCAGTGAGTCGGTCCAGTGCAGCAC
SeAcs618SlAcs F	ATGCGCCGCATTTTGCGCGACGTCGCGGAGAACCGC
SeAcs618SeAcs R	GCGGTTCTCCGCGACGTCGCGCAAAATGCGGCGCAT
*Se*Acs627*Sl*Acs F	CGATACCAGCAATCTGGGCGATGTCACGACGCTGGCGGACTCG
*Se*Acs627*Sl*Acs R	GTCCGCCAGCGTCGTGACATCGCCCAGATTGCTGGTATCG
SeAcs638SlAcs R	ACTGAATTCTCAGTCCTCGCTGGGGGCGGCCGGGAGCTTCGTC TGGATGAGCTCCACCACGCCAGGATCGGC
SeAcs643SlAcs R	ACTGAATTCTCAGTCCTCGCTGGGGGCGGCCGGGAGCTGCTTC TCTTCGAGCAG
*Se*Acs582*Sl*Acs589*Se*Acs F	AACCACGTCGGTGCCACGCTCGGCCCGCTGGCGACGCCGGAC GTG
*Se*Acs582*Sl*Acs589*Se*Acs R	CGGGCCGAGCGTGGCACCGACGTGGTTGCGCACCTCCGCGTA CAG
*Se*Acs590*Sl*Acs597*Se*Acs F	ATCGCCAAGCCCAAGCGGATCCACTGGACCGACTCACTG
*Se*Acs590*Sl*Acs597*Se*Acs R	GATCCGCTTGGGCTTGGCGATTGGGCCAATCTCTTTACG
*Se*Acs597*Sl*Acs603*Se*Acs F	CTGCTGCCGGTCTCGGAGCTGCCAAAAACCCGTTCC
*Se*Acs597*Sl*Acs603*Se*Acs R	CTCCGAGACCGGCAGCAGCACGTCCGGCGTCGC
*Se*Acs582*Sl*Acs603*Se*Acs F	GGCCCGATCGCCAAGCCCAAGCGGATCCTGCCGGTCTCGGA

	GCTGCCAAAAACCCGTTCC
*Se*Acs582*Sl*Acs603*Se*Acs R	
7.2 F	CCGGCAAGATCATGCGGCGCATTTTGCGCAAAATCGCGG CGGGCGATACCAGCAATCTG
7.2 R	GCGCCGCATGATCTTGCCGGAGCGGGTCTTCGGCAGCTC CGAGACCGGCAG
*Sl*Acs K610A F	CCCGCTCCGGCGCGATCATGCGGCG
*Sl*Acs K610A R	CGCCGCATGATCGCGCCGGAGCGGG

p*Sl*Acs19 (*Sl*Acs 550-643 *Se*Acs) – the nucleotides encoding the first 643 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9 and cloned into pTEV5

p*Sl*Acs26 (*Sl*Acs 550-581 *Se*Acs, 591-627 *Se*Acs) – the nucleotides encoding the first 581 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9 with primers incorporating residues 582-590 from *Sl*Acs, fused to the nucleotides encoding the 64 residues of *Se*Acs fused to *Sl*Acs amplified from p*Sl*Acs17, and cloned into pTEV5

p*Sl*Acs27 (*Sl*Acs 550-590 *Se*Acs, 598-627 *Se*Acs) – the nucleotides encoding the first 590 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9 with primers incorporating residues 591-597 from *Sl*Acs, fused to the nucleotides encoding the 57 residues of *Se*Acs fused to *Sl*Acs amplified from p*Sl*Acs17, and cloned into pTEV5

p*Sl*Acs28 (*Sl*Acs 550-597 *Se*Acs, 603-627 *Se*Acs) – the nucleotides encoding the first 597 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9 with primers incorporating residues 598-602 from *Sl*Acs, fused to the nucleotides encoding the 52 residues of *Se*Acs fused to *Sl*Acs amplified from p*Sl*Acs17, and cloned into pTEV5

p*Sl*Acs29 (*Sl*Acs 550-581 *Se*Acs, 603-627 *Se*Acs) – the nucleotides encoding the first 581 residues of *Sl*Acs fused to *Se*Acs were amplified from p*Sl*Acs9 with primers incorporating residues 582-602 from *Sl*Acs, fused to the nucleotides encoding the 52 residues of *Se*Acs fused to *Sl*Acs amplified from p*Sl*Acs17, and cloned into pTEV5

p*Sl*Acs44 (*Sl*Acs 615-626 *Se*Acs) – the nucleotides encoding the first 614 residues of *Sl*Acs were amplified from p*Sl*Acs1, the nucleotides encoding the final 40 residues of *Se*Acs fused to *Sl*Acs amplified from p*Sl*Acs28, and cloned into pTEV5.

Construction of untagged* Sl*Acs complementation plasmid. The *S. lividans acs* was amplified from p*Sl*Acs1 with the primers that included an optimized ribosome-binding site (Table 5.3). The DNA fragment was cut with EcoRI and HindIII and ligated into pBAD30 (14), cut with the same enzymes. The resulting plasmids p*Sl*Acs6 expresses *Sl*Acs under the control of the P_{araBAD} promoter.

***Construction of* Se*Acs*, Sl*Acs*, and Sl*Acs*-Se*Acs* chimera C3 H_6-tagged complementation vectors.** *S. lividans acs* and the *S. lividans / S. enterica acs* chimera were amplified from p*Sl*Acs1 and p*Sl*Acs28, respectively, using primers that included an optimized ribosome-binding site and an *N*-terminal His_6-tag (Table 5.3). *S. enterica acs* was amplified from genomic DNA isolated from JE6583 using primers that included an optimized ribosome-binding site and an *N*-terminal His_6-tag (Table 5.3).The DNA fragments were cut with EcoRI and HindIII and ligated into pBAD30, cut with the same enzymes. The resulting plasmids p*Sl*Acs47, p*Sl*Acs48, and pACS59 produce *Sl*Acs, *S. lividans / S. enterica acs* chimera C3, and *Se*Acs^{WT}, respectively, with His_6-tags fused *N*-terminal with a Gly-Ser-Gly linker under the control of a the P_{araBAD} promoter.

Protein purification:* Sl*Acs*-Se*Acs* chimera *purification. Plasmids containing *S. lividans acs* - *S. enterica acs* chimeras were transformed with pRARE2 (EMD Millipore) into a Δ*pka* derivative of C41λ(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 LB containing ampicillin (100 mg ml^{-1}) and chloramphenicol (12.5 μg ml^{-1}). The cultures were grown shaking at 25°C to A_{600} ~ 0.7 and protein synthesis was induced with IPTG (0.25 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 resuspended in 30 ml cold His-Bind buffer (buffer A) [*tris*(hydroxymethyl)aminomethane-HCl (Tris-HCl) buffer (50 mM, pH 8), NaCl (500 mM)], and imidazole (5 mM)) containing phenylmethanesulfonylfluoride (PMSF, 1 mM). Cells were placed on ice and lysed by sonication for 2 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*. H_6-*Sl*Acs-*Se*Acs chimera was purified from clarified cell extract using a 1 ml settled bed volume of HisPur™ Ni-NTA Resin (Pierce). Unbound proteins were eluted off the column by washing with buffer A. The resin was washed with 10 column volumes of buffer B [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (15 mM)]. H_6-*Sl*Acs-*Se*Acs chimera was eluted with 5 column volumes of buffer C [Tris-

HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (250 mM)]. All fractions containing H_6-*Sl*Acs-*Se*Acs chimera were combined. rTEV protease was added to H_6-*Sl*Acs-*Se*Acs chimera and the *Sl*Acs-*Se*Acs chimera/rTEV mixture was incubated at room temperature for 3 h. PMSF was added to the protein mixture and incubated 15 min at room temperature. The *Sl*Acs-*Se*Acs chimera/rTEV mixture was dialyzed at 4°C against buffer D (Tris-HCl (50 mM, pH 8), NaCl (500 mM)) twice for 3 h and again against buffer D containing imidazole (5 mM) for 12 h. After cleavage and dialysis, protein mixtures were passed over 1 ml HisPur^{TM} Ni-NTA Resin (Pierce) using the buffers described above. Cleaved *Sl*Acs-*Se*Acs chimera passed through the resin and eluted in the flow-through fractions. Purified *Sl*Acs-*Se*Acs chimera was analyzed by SDS-PAGE. Fractions containing *Sl*Acs-*Se*Acs chimera were pooled together. *Sl*Acs-*Se*Acs chimera was stored in Tris-Cl buffer (50 mM, pH 8.0) containing NaCl (100 mM) and glycerol (20%, v/v). *Sl*Acs concentration was determined by measuring absorbance at 280 nm. The molecular weights and molar extinction coefficients used to calculate H_6-*Sl*Acs-*Se*Acs chimera concentrations are listed in table 5.4.

Protein purification: **Se*Acs purification.*** Plasmid pACS10 was transformed into a Δ*pat* derivative of C41λ(DE3) (JE9314). The resulting strain was grown overnight and sub-cultured 1:100 (v/v) into 2 liters of LB containing ampicillin (100 mg ml^{-1}). The culture was grown shaking at 37°C to A_{600} ~ 0.7 and protein synthesis was induced with IPTG (0.25 mM). Upon induction, the cultures were grown overnight at 30°C. *Se*Acs was purified and stored as described (29). *Sl*Acs^{WT} and *Sl*PatA^{WT} were purified as described (35).

In vitro *CoA ligase assays.* Activity of *Sl*Acs^{WT}, *Se*Acs^{WT}, and *Sl*Acs-*Se*Acs chimera activities were measured using an NADH-consuming assay (6, 11). Reactions (100 µl total volume) contained HEPES buffer (50 mM, pH 7.5), TCEP (1 mM), ATP (2.5 mM) CoA (0.5 mM), $MgCl_2$ (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 Acs (10-100

Table 5.4. Molecular weights and molar extinction coefficients of proteins used in this study.

Protein	MW (Da)	ε (M^{-1} cm^{-1})
*Sl*Acs	71045	135455
*Se*Acs	72153	138770
A1 chimera	71527	150925
A2 chimera	71541	150925
A3 chimera	71500	150800
A4 chimera	71432	147945
A5 chimera	71352	146455
A6 chimera	71115	135455
B1 chimera	71234	139925
B2 chimera	71471	150800
B3 chimera	71471	150925
B4 chimera	71530	150925
B5 chimera	71466	150925
B6 chimera	71751	150925
C1 chimera	71297	145425
C2 chimera	71627	150925
C3 chimera	71429	145425
C4 chimera	71293	139800
C5 chimera	71104	135330
C4 chimera K609A variant	71429	145425
*Sl*PatA	108369	57760

pmol). The absorbance at 340 nm was monitored in a 96-well plate using the Spectramax Plus UV-visible spectrophotometer (Molecular Devices). Enzyme activities were calculated as described (11).

In vitro *protein acetylation assay*. Protein acetylation was observed using radiolabeled Ac-CoA as described (6, 31, 36). Acetylation reactions contained 2(*bis*(2-hydroxyethyl)imino)-2-(hydroxymethyl)-1,3-propanediol (Bis-Tris-HCl) buffer (50 mM, pH 6.0), [1-^{14}C]-Ac-CoA (20 mM), acyl-CoA synthetase or acyl-CoA synthetase *C*-terminal domain (3 μM), glycerol (10%, v/v), and *Sl*PatA^{WT} (1 μM). Reactions (20 μl total volume) were incubated for 60 min at 30°C. Samples (5 μl) were resolved using SDS-PAGE (19) and proteins were visualize by Coomassie Blue R staining (27). Gels were dried and exposed 16 h to a multipurpose phosphor screen (Packard). Labeled proteins were visualized using a Typhoon Trio+ Imager (GE Healthcare) equipped with ImageQuant TL software (GE Healthcare). Acetylation was quantified as digital light units and is reported relative to *Se*Acs^{WT} acetylation.

The effect of acetylation on activity of *Sl*Acs^{WT}, *Se*Acs^{WT}, and *SlAcs-SeAcs* chimera activity was determined as described (6) with modifications. *Sl*Acs^{WT}, *Se*Acs^{WT}, or *SlAcs-SeAcs* (3μM) was incubated with *Sl*PatA^{WT} (1 μM) and 50 μM Ac-CoA for 90 min at 30°C using the buffer system described above. After 90 min, reactions were diluted into HEPES buffer (50 mM, pH 7.5 at 4°C). *Sl*Acs^{WT}, *Se*Acs^{WT}, and *SlAcs-SeAcs* chimera activities were measured as described above.

In vitro *deacetylation assays.* Acetylated *Sl*Acs-*Se*Acs chimera C3 was deacetylated with *S. enterica* CobB$_S$ as described (36). *In vitro* acetylated *Sl*Acs-*Se*Acs chimera C3 (3 mM, radiolabeled) was incubated with *Se*CobB$_S$ (3 μM) in deacetylation buffer containing 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 50 mM, pH 7.0), NAD^+ (1 mM) for 60 min at 37°C (10 μl reaction volume). Reaction mixture samples (5 μl) were resolved by SDS-PAGE, and subjected to phosphor imaging analysis to assess the acetylation state of *Sl*Acs-*Se*Acs chimera C3 after incubation with *Se*CobB$_S$.

RESULTS

S. lividans *acetyl-CoA synthetase (Acs) is functional* in vivo *in* Salmonella enterica. The Acs homologue from *S. lividans* converts acetate to acetyl-CoA *in vitro* (35). Alignment of the *Se*Acs and *Sl*Acs amino acid sequences using BLAST revealed 52% sequence identity and 62% sequence similarity in amino acid sequence. To determine whether or not *Sl*Acs functioned *in vivo*, we used the heterologous host *S. enterica* as a model. *S. enterica* uses two pathways for the conversion of acetate to acetyl-CoA (Fig. 5.1A) (30, 33). One pathway is comprised of *Se*Acs, which catalyzes a two-step conversion of acetate to acetyl-CoA via an acetyl-AMP intermediate. RLA controls *Se*Acs activity. The protein acetyltransferase *Se*Pat acetylates and inactivates of *Se*Acs (discussed further below), and *Se*Acs is deacetylated and reactivated by the sirtuin type deacetylase CobB. In the second pathway, acetate kinase (Ack) and phosphotransacetylase (Pta) catalyze the conversion of acetate to acetyl-CoA via an acetyl-phosphate intermediate. Acs has a higher affinity for acetate than Pta/Ack, and is the preferred pathway when *S. enterica* is growing on low concentrations of acetate (≤10 mM). A *S. enterica* strain lacking the Acs and Ack/Pta pathways fail to grow on acetate (10 mM, Fig. 5.1B, squares). When *Sl*Acs was produced ectopically, growth of an *S. enterica* Δ*acs* Δ*pta* strain was restored (Fig. 5.1B, circles), demonstrating that *Sl*Acs was active and could substitute for *Se*Acs function *in vivo*.

Sl*PatA acetylates the* C-*terminal domain of* Se*Acs, but not* Sl*Acs.* AMP-forming CoA synthetases are two-domain enzymes that activate fatty acids to CoA thioesters in a two-step reaction. In the first half-reaction, an invariant lysine in the small domain (K609 of *Se*Acs) is buried in the active site cleft located between the *N*- and *C*-terminal domains (13). Upon adenylylation of the fatty acid substrate, the *C*-terminal domain undergoes a ~140º domain rotation to allow for the thioesterification of the fatty acyl-AMP intermediate (23). The catalytic lysine of AMP-forming CoA ligases is surface exposed when the enzyme is in the thioester-forming conformation (13), and this likely represents the conformation that is subject to acetylation by Pat.

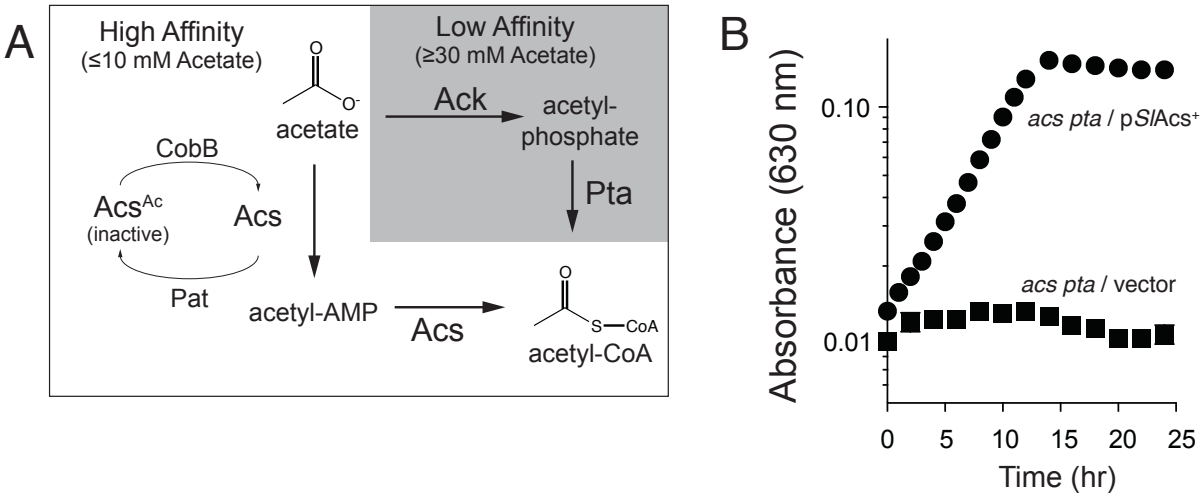


Figure 5.1. *Sl*Acs can substitute for *Se*Acs in *S. enterica* during growth on acetate. (A) *S. enterica* encodes both high affinity and low affinity pathways for acetate utilization. The high affinity pathway is composed of acetyl-CoA synthetase (Acs) and is regulated by protein acetyltransferase (Pat) and sirtuin (CobB). The low affinity pathway is comprised of acetate kinase (Ack) and phosphotransacetylase (Pta). (B) Growth behavior of *S. enterica* 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). Data are plotted as the mean of three data point. All standard deviations are <0.01 absorbance units.

Previously, we demonstrated that *Sl*Acs was a poor substrate for the *S. lividans* Pat homologue (*Sl*PatA) *in vitro* (35). That work identified *Sl*Acs as the first example of an acetyl-CoA synthetase that was not recognized by the cognate Pat protein acetyltransferase *in vitro* (6, 32). However, *Sl*PatA efficiently acetylated and inactivated the acetoacetyl-CoA synthetase *Sl*AacS from *S. lividans*, and the orthologous *Se*Acs enzyme (35), indicating that *Sl*PatA was catalytically active, but somehow unable to acetylate *Sl*Acs *in vitro*.

We considered the possibility that *Sl*Acs favored the adenylyation conformation *in vitro*, which would likely render the target K610 inaccessible to *Sl*PatA due to its location in the *Sl*Acs active site. To differentiate the inaccessibility of *Sl*Acs K610 from the inability of *Sl*PatA to recognize *Sl*Acs, we isolated the *C*-terminal domains of *Se*Acs (a good substrate of *Sl*PatA), and *Sl*Acs. In the absence of the *N*-terminal domain, the Acs active site can no longer form exposing the target lysine, thus rendering it accessible to the acetyltransferase.

Homogeneous *C*-terminal domains of *Sl*Acs and *Se*Acs were incubated in the presence of *Sl*PatA and radiolabeled [1-^{14}C] acetyl-CoA. As shown in figure 5.2A, the *C*-terminal domain of *Se*Acs was acetylated, but the *Sl*Acs *C*-terminal domain was not. These data suggested that the *N*-terminal domain of *Se*Acs was not required for acetylation by *Sl*PatA. Additionally, these results strongly suggested that inaccessibility of *Sl*Acs residue K610 was not the likely the reason why it was poorly acetylated. Based on these data, we hypothesized that regions within the *C*-terminal domain of *Sl*Acs enzyme prevent acetylation of *Sl*PatA. As shown in figure 5.2B, the *C*-terminal domains of *Sl*Acs and *Se*Acs share ~50% sequence identity.

Chimeras of* Sl*Acs and* Se*Acs reveal regions in the* Se*Acs* C*-terminal domain that are critical for acetylation by* Sl*PatA. Based on regions of sequence conservation (Fig. 5.2B), we generated a set of precise fusions between the *Sl*Acs and *Se*Acs that contained varying amounts of the *Se*Acs protein. As shown in figure 5.3A-B, a *Sl*Acs chimera containing the *Sl*Acs *N*-terminal domain fused to the *Se*Acs *C*-terminal (chimera A1) was strongly acetylated by *Sl*PatA, confirming that the *C*-terminal domain of *Sl*Acs was responsible for the poor acetylation of *Sl*Acs^{WT}.

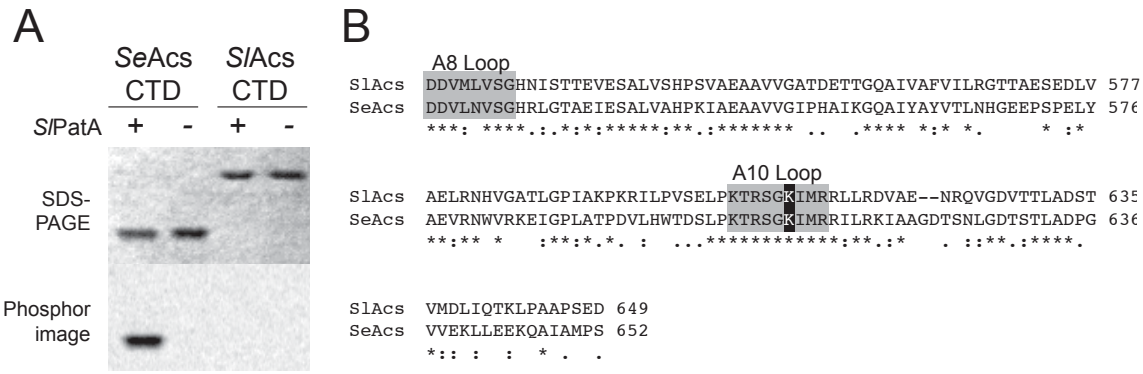


Figure 5.2. *Sl*PatA efficiently acetylates the *C*-terminal domain of *Se*Acs. (A) The *C*-terminal domain of *Sl*Acs^{WT} or *Se*Acs^{WT} was incubated with [1-^{14}C]-acetyl-CoA in the presence or absence of *Sl*PatA^{WT}. Proteins were separated by SDS-PAGE and stained with Coomassie Blue R to visualize proteins. Acetylation was visualized by phosphor imaging. (B) Alignment of the *C*-terminal domain of *Sl*Acs and *Se*Acs. " * " denotes conserved residues; "." denotes similar residues; light gray boxes denote conserved loops of the AMP-forming CoA ligase family (12); black box denotes the catalytic lysine.

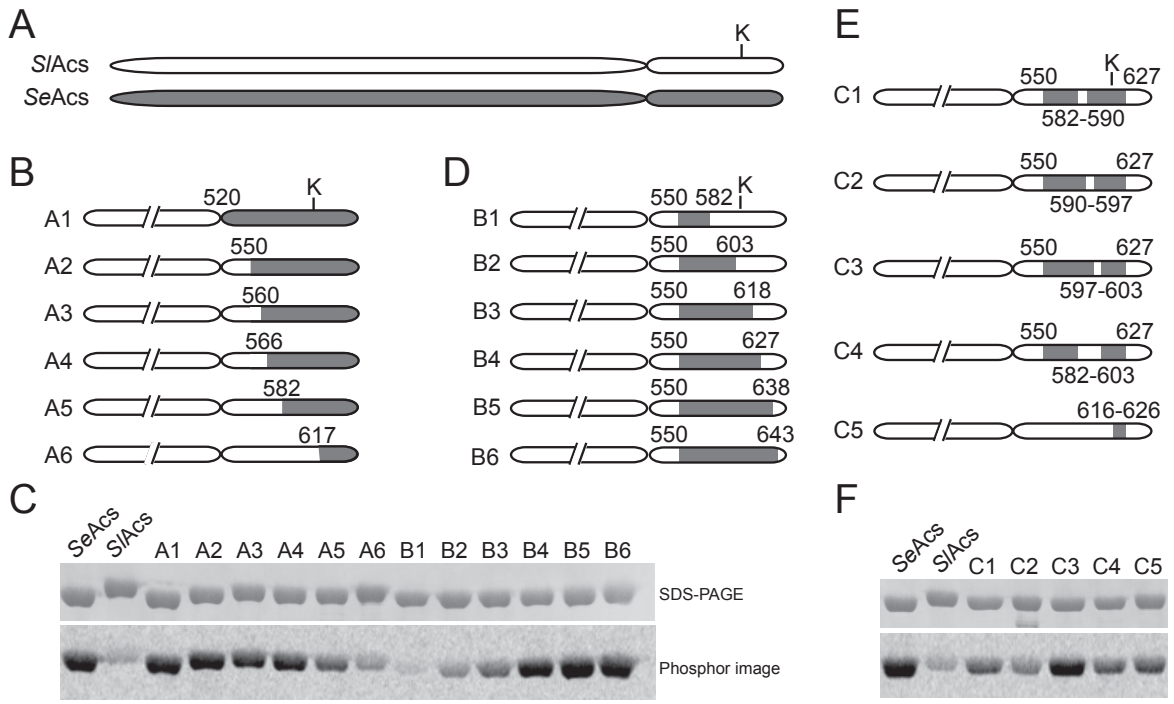


Figure 5.3. Construction and acetylation of *Sl*Acs-*Se*Acs chimeras. (A) A scheme of *Sl*Acs^{WT} (white) and *Se*Acs^{WT} (grey) drawn to scale. Target lysine K610 for *Sl*Acs^{WT} and K609 aligned and depicted by "K". The *N*-terminal domains (520 residues) are shortened with a hatch in all remaining panels. (A) Schematic representation of *Sl*Acs-*Se*Acs chimeras A1-A6 in which the *C*-terminal portion of *Sl*Acs^{WT} (white) was replaced with the corresponding amino acid sequence from *Se*Acs^{WT} (gray). All chimeras are drawn to scale for reference to the target lysine (denoted with "K"). Numbers all denote the fusion points with respect to the *Sl*Acs protein sequence (i.e. either the first residue of *Sl*Acs^{WT} replaced by *Se*Acs^{WT} sequence or the first residue of *Sl*Acs^{WT} after the *Se*Acs^{WT} amino acid sequence). (C) Acetylation of *Sl*Acs-*Se*Acs chimeras A1-A6 and B1-B6 using *Sl*PatA^{WT} and [1-^{14}C]acetyl-CoA. (D) Schematic or *Sl*Acs-*Se*Acs chimers B1-B6 in which internal portion of the *C*-terminal *Sl*Acs domain are replaced with the corresponding sequence from *Se*Acs^{WT}. (E) Schematic of chimeras C1-C5. (F) Acetylation of *Sl*Acs-*Se*Acs chimeras C1-C6.

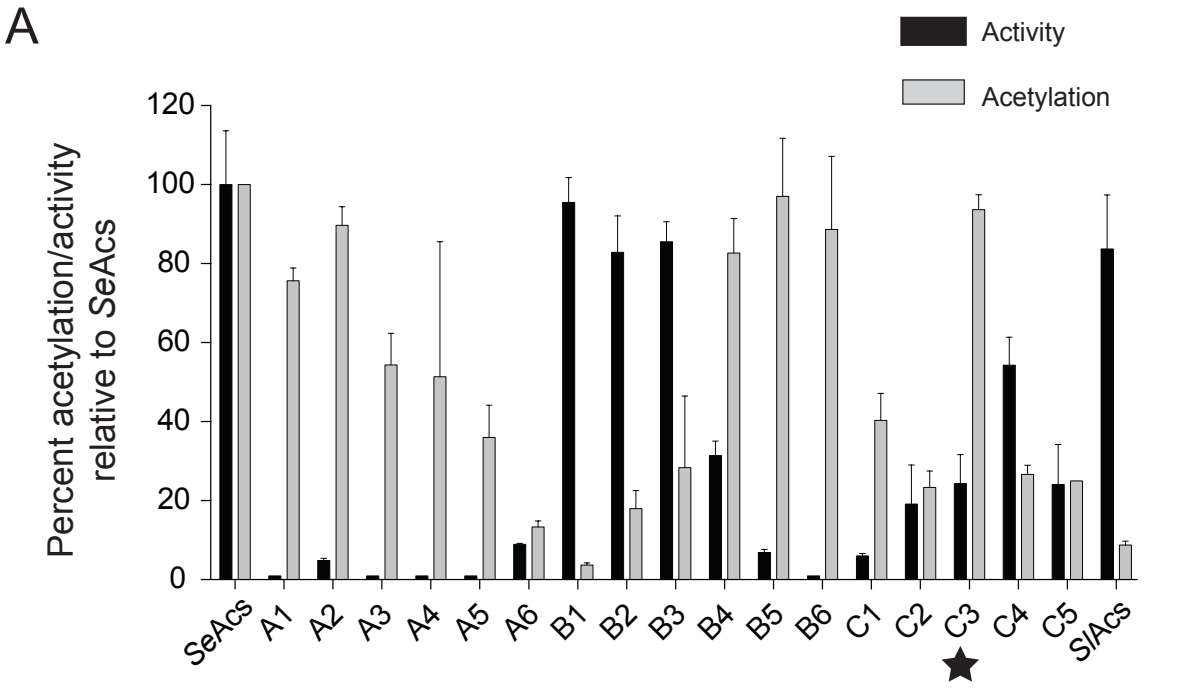
We identified regions of the *Se*Acs *C*-terminal domain important for acetylation by *Sl*PatA by constructing chimeras that contained decreasing amounts of the *Se*Acs *C*-terminal domain relative to chimera A1. To measure the efficiency of acetylation, each chimera was incubated with *Sl*PatA and radiolabeled [1-^{14}C] acetyl-CoA. *Sl*PatA strongly acetylated chimeras containing at least the final 86 amino acids of *Se*Acs (chimeras A1, A2, A3, A4; Fig. 5.3). These chimeras contained at least 43 amino acids *N*-terminal to the acetylation site, a region previously reported to be important for acetylation of homologous AMP-forming CoA ligase enzymes by the *R. palustris* Pat homologue (8).

To narrow down the number of *Se*Acs residues required for acetylation of the *Sl*Acs-*Se*Acs chimeras, we focused on *Sl*Acs-*Se*Acs chimera A2, which had the fewest *Se*Acs-derived residues (Fig. 5.3B), and the highest level of acetylation (Fig. 5.3C).

We generated a second set of chimeras in which various stretches of *Sl*Acs-derived residues were substituted into *Sl*Acs-*Se*Acs chimera A2 (Fig. 5.3D). *Sl*Acs-*Se*Acs chimeras B4, B5, and B6 that contained at least 45 residues of *Se*Acs (including the *Se*Acs^{K609} acetylation site) were strongly acetylated (Fig. 5.3C). Notably, the A10 loop of Acs, which contains the target lysine, is completely conserved between *Se*Acs and *Sl*Acs (Fig. 5.2B). However, 17 amino acids *C*-terminal to the acetylation site of *Se*Acs were required for acetylation by *Sl*PatA. This revealed a previously unrecognized region of the protein important for acetylation. Of this set of *Sl*Acs-*Se*Acs chimeras, chimera B4 was the best substrate of *Sl*PatA and contained the fewest *Se*Acs-derived amino acids.

To determine whether the 77 contiguous *Se*Acs-derived residues of chimera B4 were critical for acetylation, we identified regions of *Sl*Acs and *Se*Acs with low amino acid sequence conservation and introduced those sets of *Sl*Acs residues into chimera B4 (Fig. 5.3E). Of those tested, only *Sl*Acs-*Se*Acs chimera C3 was acetylated with similar efficiency as *Se*Acs (Fig. 5.3F). Acetylation of each chimera was quantified relative to acetylation of *Se*Acs (Fig. 5.4A, gray bars).

Importantly, chimeras containing only the 43 *Se*Acs-derived residues *N*-terminal to K610 (chimera B2) or 17 *Se*Acs-derived residues *C*-terminal to K610 (chimera C5) were <30% acetylated relative to *Se*Acs. Thus, amino acid sequences *N*- and *C*-terminal to the target lysine were important for



B

```
SlAcs   550  GATDETTGQAIVAFVILRGTTAESEDLVAELRNHVGATLGPIAKPKRILPVSE
C3      550  GIPHAIKGQAIYAYVTLNHGEEPSPELYAEVRNWVRKEIGPLATPDVLLPVDS
SeAcs   549  GIPHAIKGQAIYAYVTLNHGEEPSPELYAEVRNWVRKEIGPLATPDVLHWTDS

SlAcs   603  LPKTRSGKIMRRLLRDVA--ENRQVGDVTTLADSTVMD--LIQTKLPAAPS
C3      603  LPKTRSGKIMRRILRKIAAGDTSNLGDVTTLADSTVMD--LIQTKLPAAPS
SeAcs   602  LPKTRSGKIMRRILRKIAAGDTSNLGDTSTLADPGVVEKLLEEKQAIAMPS
```

Figure 5.4. *Sl*Acs-*Se*Acs Chimera C3 is active and efficiently acetylated. (A) Acetyl-CoA synthetase activity of each chimera and *Sl*Acs^{WT} relative to *Se*Acs^{WT} (gray bars). Amount of acetylation in figure 3C and 3F was quantified and normalized to the total acetylation of *Se*Acs (black bars). *Sl*Acs-*Se*Acs chimera C3, the most efficiently acetylated, active chimera with the fewest *Se*Acs^{WT}-derived residues, is noted with a star. Values are reported as the mean ± S.D. of three experiments. (B) Sequence alignment of *Sl*Acs^{WT}, *Se*Acs^{WT}, and chimera C3. Residues in chimera C3 that are derived from the *Se*Acs^{WT} amino acid sequence are highlighted in black.

acetylation by *Sl*PatA, and neither set of amino acids rendered *Sl*Acs a strong acetylation target when introduced independently.

Assessment of the enzymatic activity of the chimeras. Chimeras were tested for their AMP-forming acetyl-CoA ligase forming activity. Although the *Sl*Acs^{WT} and *Se*Acs^{WT} *C*-terminal domains share a high degree of sequence conservation, not all chimeras were active (Fig. 5.4A, black bars). To identify active chimeras that were also targets of acetylation, the acetylation of each chimera was measured relative to *Se*Acs (Fig. 5.4A, gray bars). *Sl*Acs-*Se*Acs chimera C3 (Hereafter referred to as chimera C3) was identified as the single chimera with the fewest *Se*Acs residues that was active and efficiently acetylated by *Sl*PatA. As shown in figure 5.4B, chimera C3 contained 41 amino acid differences from *Sl*Acs.

Chimera C3 activity is modulated by acetylation and deacetylation. As shown in figure 5.5A, the catalytic residue K610 residue is the only residue of chimera C3 that was acetylated. To test whether the activity of chimera C3 was under the control of acetylation, the protein was incubated with *Sl*PatA acetyltransferase in the presence and absence of the acetyl donor, acetyl-CoA. Upon acetylation, chimera C3 activity decreased >98%, similar to the regulation of *Se*Acs activity (Fig. 5.5B). This represented a 40-fold increase in regulation of the chimera C3 over the *Sl*Acs enzyme, which retained ~75% activity upon incubation with *Sl*PatA and Ac-CoA (Fig. 5.5B). Acetylation of *Se*Acs^{WT} is reversed by the sirtuin-type (NAD^+-dependent) deacetylase CobB in *S. enterica,* and deacetylation reactivates the *Se*Acs^{WT} enzyme (29). We tested whether chimera C3 was a deacetylation target by incubating acetylated chimera C3 with *Se*CobB, the co-substrate NAD^+, or both. When both *Se*CobB and NAD^+ were added, we observed a decrease in acetylation of chimera C3^{Ac} (Fig. 5.5C), demonstrating that acetylated chimera C3 was subject a substrate of a sirtuin-type deacetylase.

Sl*Acs*-Se*Acs chimera C3 is acetylated* in vivo *in* S. enterica *by* Sl*PatA*. To determine the efficiency of *Sl*PatA acetylation of Chimera C3 *in vivo*, we used *S. enterica* acetate utilization (Fig. 5.1A) as a heterologous model to demonstrate the effects of *Sl*PatA acetylation on activity of the Acs homologues. In this system, His-tagged chimera C3, *Sl*Acs, and *Se*Acs (H_6-Chimera C3, H_6-*Sl*Acs, H_6-

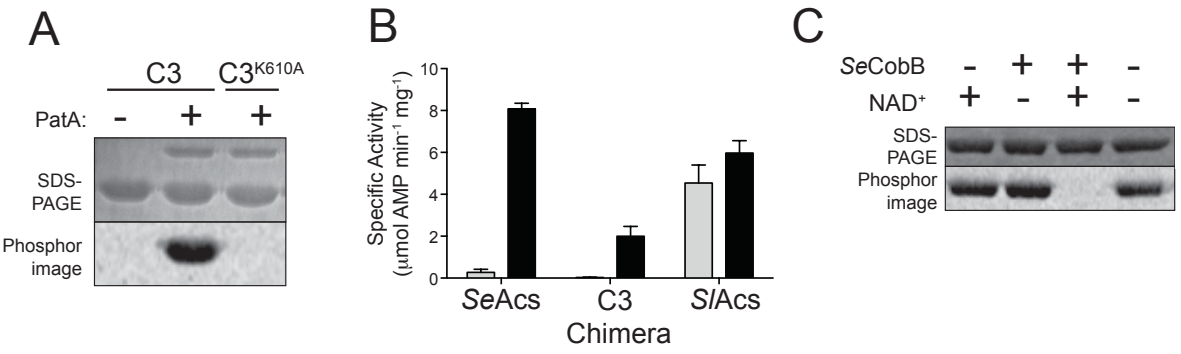


Figure 5.5. *Sl*Acs-*Se*Acs Chimera C3 is regulated by reversible lysine acetylation. (A) *Sl*Acs-*Se*Acs chimera C3 or *Sl*Acs-*Se*Acs chimera $C3^{K610A}$ was incubated with [1-^{14}C]-acetyl-CoA in the presence or absence of SlPatA^{WT}. Proteins were separated by SDS-PAGE and stained with Coomassie Blue R to visualize proteins. Acetylation was visualized by phosphor imaging. (B) *Sl*Acs-*Se*Acs chimera C3, SeAcs^{WT}, or SlAcs^{WT} was incubated with SlPatA^{WT} in the presence (grey bars) or absence (black bars). Samples were removed, diluted, and assayed to measure acetyl-CoA synthetase activity after 90 min incubation with SlPatA^{WT}. Acs activity was measured in an NADH-consumption assay. Values are reported as the mean ± S.D. of three experiments. (C) Sl*Acs*-Se*Acs* chimera C3 previously acetylated by SlPatA^{WT} with [1-^{14}C]-acetyl-CoA was incubated with the addition of SeCobB^{WT} and/or NAD^+. Proteins were resolved by SDS-PAGE and stained with Coomassie Blue R to visualize proteins. Acetylation was visualized by phosphor image.

*Se*Acs, respectively) were produced from plasmids in *S. enterica acs pat cobB* and *S. enterica acs pat cobB*$^+$ strains JE9152 and JE9894, respectively (Table 5.1). Experiments were conducted in *S. enterica pat* strains to prevent acetylation by *Se*Pat. We characterized the effect of *Sl*PatA acetylation on the H_6-Acs homologues by measuring growth of each strain bearing a plasmid with an inducible *Sl*PatA allele or an empty vector. Additionally, we isolated H_6-*Se*Acs^{WT}, H_6-*Sl*Acs^{WT}, and H_6-chimera C3 from cells grown in the presence or absence of *Sl*PatA to quantify the effects of *Sl*PatA acetylation on each Acs protein. As shown in figure 5.6A, production of H_6-chimera C3, H_6-*Sl*Acs, or H_6-*Se*Acs supported growth of *S. enterica acs pat cobB* strain (open symbols). We attributed the lag in the strain producing H_6-chimera C3 to the decreased activity of this chimera (Fig. 5.5B). When production of *Sl*PatA is induced in each strain (25 μM inducer), growth of *S. enterica acs pat cobB* strains producing H_6-*Se*Acs^{WT} or H_6-chimera C3 was significantly reduced, while growth of the *S. enterica acs pat cobB* strain producing H_6-*Sl*Acs^{WT} was unaffected. Importantly, inhibition of an *S. enterica acs cobB* strain producing H_6-*Sl*Acs^{WT} required high levels of *Sl*PatA^{WT} induction (500 μM inducer, Fig. 5.6B). No growth inhibition occurred when *Sl*PatA^{WT} was induced at low levels (≤ 5 μM inducer, Fig. 5.7A). In a *S. enterica acs pat cobB*$^+$ strain, *Se*CobB^{WT} is present to deacetylate and reactivate the Acs enzyme, and we observed no significant growth defects upon *Sl*PatA^{WT} induction of strains expressing H_6-*Sl*Acs^{WT} or H_6-*Se*Acs^{WT} (Fig. 5.7B). However, we did note slight inhibition of growth of a *S. enterica acs pat cobB*$^+$ strain producing H_6-chimera C3. We surmise that this was likely due to the inability of *Se*CobB^{WT} to efficiently deacetylate and reactivate H_6-chimera C3 and restore growth. This idea was supported by the observation that increased induction of *Sl*PatA^{WT} inhibited a *S. enterica acs pat cobB*$^+$ strain producing H_6-chimera C3 (Fig. 5.7C), but not those producing H_6-*Sl*Acs^{WT} nor H_6-*Se*Acs^{WT} (Fig. 5.7D).

Since high levels of *Sl*PatA induction were required to inhibit growth of an *S. enterica acs cobB* strain producing H_6-*Sl*Acs^{WT}, we expected that H_6-*Sl*Acs^{WT} was poorly acetylated by *Sl*PatA^{WT} and thus more active *in vivo*. We also expected higher proportions of acetylated to non-acetylated H_6-*Se*Acs^{WT} and H_6-chimera C3 *in vivo*. To measure the effect of *Sl*PatA^{WT} acetylation on the activity of H_6-*Sl*Acs of, H_6-

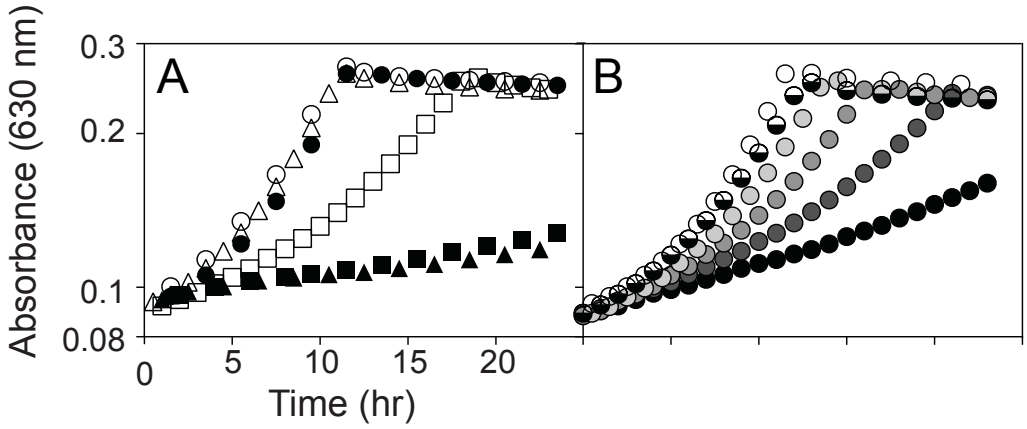


Figure 5.6. *Sl*Acs-*Se*Acs chimera C3 is regulated by *Sl*PatA *in vivo* in *S. enterica*. Growth behavior of *S. enterica* in NCE minimal medium supplemented with acetate (10 mM). (A) Growth of *S. enterica* Δ*acs pat* Δ*cobB* producing H_6-*Se*Acs^{WT} (triangles), H_6-*Sl*Acs^{WT} (circles), or H_6-*Sl*Acs-*Se*Acs chimera C3 (squares) harboring either a plasmid expressing *Sl*PatA^{WT} (filled shapes) or an empty vector (open shapes). All strains were grown in the presence of 25 μM IPTG. (B) Growth of *S. enterica acs pat cobB* (JE9152) producing H_6-*Sl*Acs^{WT} harboring a plasmid producing *Sl*PatA^{WT} induced with IPTG concentrations of 25 μM (open circles), 50 μM (light gray), 100 μM (medium gray), 250 μM (dark gray), or 500 μM (black). For reference, half-filled circles denote the growth of an equivalent strain producing H_6-*Sl*Acs^{WT} harboring an empty vector induced with 500 μM IPTG. Growth experiments were performed at 37°C using a microtiter plate and a microtiter plate reader (Bio-Tek Instruments). Data are plotted as the mean of three data points. All standard deviations are <0.015 absorbance units.

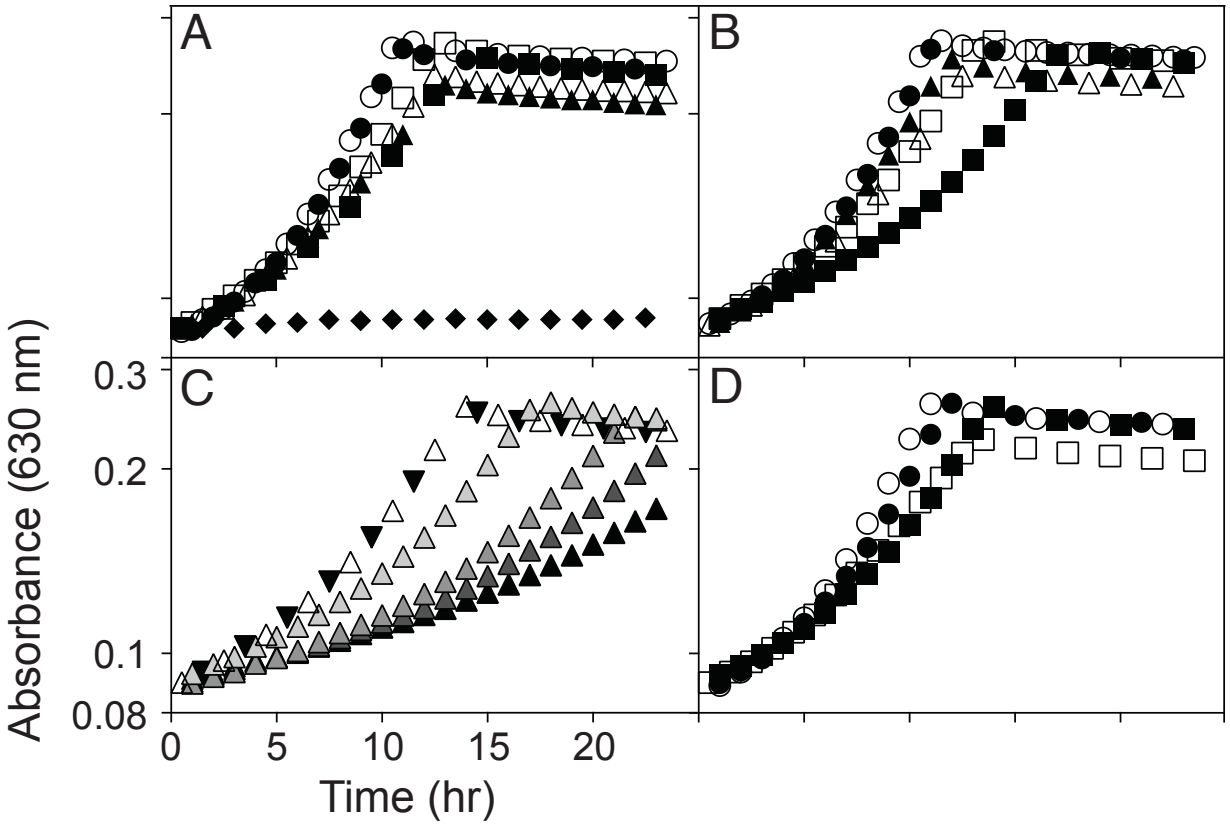


Figure 5.7. Regulation of *Sl*Acs-*Se*Acs chimera C3 by *Sl*PatA^{WT} in *S. enterica*. Growth behavior of *S. enterica* on NCE minimal medium supplemented with acetate (10 mM). (A) Growth of *S. enterica* Δ*acs pat* Δ*cobB* (JE9152) producing H_6-*Se*Acs^{WT} (triangles), H_6-*Sl*Acs^{WT} (circles), H_6-*Sl*Acs-*Se*Acs chimera C3 (squares), or empty vector (diamonds) harboring either a plasmid expressing *Sl*PatA^{WT} (filled shapes) or an empty vector (open shapes). All strains were grown in the presence of 5 μM IPTG. (B) Growth of *S. enterica* Δ*acs pat cobB*$^+$ (JE9894) producing H_6-*Se*Acs^{WT} (triangles), H_6-*Sl*Acs^{WT} (circles), or H_6-*Sl*Acs-*Se*Acs chimera C3 (squares) harboring either a plasmid expressing *Sl*PatA^{WT} (filled shapes) or an empty vector (open shapes). All strains were grown in the presence of IPTG (25 μM). (C) Growth of *S. enterica* Δ*acs pat cobB*$^+$ (JE9894) producing H_6-*Sl*Acs-*Se*Acs chimera C3 harboring a plasmid producing *Sl*PatA^{WT} induced with IPTG concentrations of 10 μM (open triangle), 25 μM (light gray), 500 μM (medium gray), 100 μM (dark gray), or 250 μM (black). For reference, the inverted, filled triangle denote the growth of an equivalent strain producing H_6-*Sl*Acs-*Se*Acs chimera C3 harboring an empty vector (no *Sl*PatA^{WT}) induced with 500 μM IPTG. Additionally, *S. enterica* Δ*acs pat cobB*$^+$ strains (JE9894) producing H_6-*Se*Acs^{WT} (circles) or H_6-*Sl*Acs^{WT} (squares) are shown growing with high induction (250 μM IPTG) of empty vector control (open symbols) or a plasmids expressing *Sl*PatA (D). Growth experiments were performed at 37°C using a microtiter plate and a microtiter plate reader (Bio-Tek Instruments). Data are plotted as the mean of three data points. All standard deviations are <0.015 absorbance units.

*Se*Acs of, and H_6-chimera C3, we grew *S. enterica acs cobB* strains expressing these *acs* alleles while inducing *Sl*PatA^{WT} at low levels (5 μM) to allow for growth and biomass accumulation for all strains (Fig. 5.7A). H_6-*Sl*Acs^{WT}, H_6-*Se*Acs^{WT} and H_6-chimera C3 enzymes were isolated from strains harboring plasmid-borne *Sl*PatA^{WT} or empty vector.

As shown in figure 5.8, activity of the H_6-*Sl*Acs^{WT} enzyme isolated from a strain producing *Sl*PatA^{WT} was not significantly reduced compared to H_6-*Sl*Acs^{WT} isolated from a strain with no *Sl*PatA^{WT}. However, activities of the H_6-*Se*Acs^{WT} and H_6-chimera C3 enzymes were significantly lower when isolated from strains expressing *Sl*PatA^{WT} compared to those with no *Sl*PatA^{WT}. These data suggested that *Sl*PatA^{WT} more efficiently acetylated H_6-*Se*Acs^{WT} and H_6-chimera C3 than it did H_6-*Sl*Acs in a heterologous *in vivo* model.

DISCUSSION

Herein we report the first Acs enzyme that is not a substrate of Pat homologues. This finding is important, since Acs is the paradigm for the analysis RLA in all metabolic systems reported thus far. Our results begin to shed some light onto why the *Sl*Acs is not efficiently acetylated by the *Sl*PatA^{WT} enzyme of *S. lividans*. By constructing chimeras of *Sl*Acs that are acetylated by *Sl*Pat^{WT} and retain biological activity we gained insights into structural, physiological and possibly evolutionary questions raised by this work.

Is Acs activity under RLA control in streptomycetes? At present, the answer to this question is unclear. It is not known whether *Sl*Acs^{WT} is a *bona fide* substrate of *Sl*PatA^{WT} *in vivo* in *S. lividans.* The literature adds to the challenge of determining whether or not in streptomycetes Acs is under RLA control. Work performed by others in *Streptomyces coelicolor* suggested that the Acs enzyme of this actinomycete may be under RLA control, because results of *in vitro* experiments showed that acetylated *Sc*Acs was a substrate of a sirtuin deacetylase present in that bacterium (Mikulik et al., 2012). The same authors also reported the isolation of acetylated *Sc*Acs from

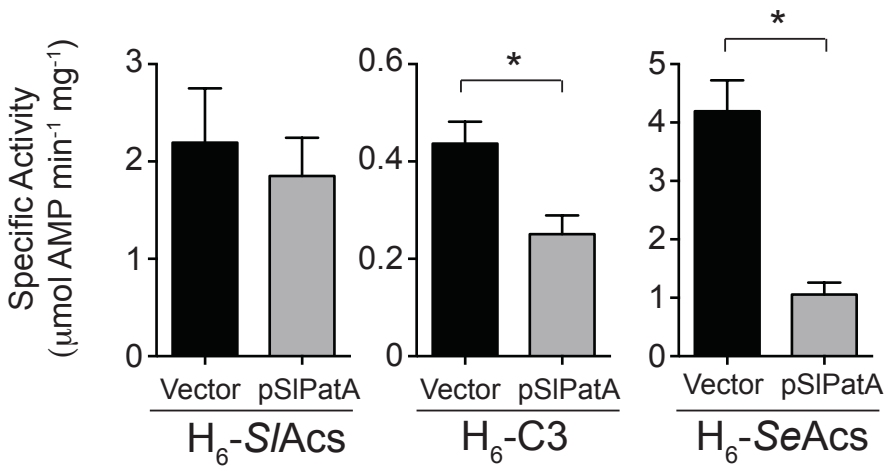


Figure 5.8. Activity of *Sl*Acs-*Se*Acs chimera C3 and *Se*Acs^{WT} is reduced in strains expressing *Sl*PatA^{WT}. H_6-*Sl*Acs-*Se*Acs Chimera C3, H_6-*Sl*Acs, and H_6-*Se*Acs were produced in *S. enterica* Δ*acs pat* Δ*cobB* strain JE9152 harboring either a plasmid producing *Sl*PatA^{WT} or an empty vector. Strains were grown in NCE minimal medium supplemented with acetate (10 mM). Acs activity was measured in an NADH-consumption assay. Values are reported as the mean ± S.D. of three activity measurements. Unpaired t-test was used to compare activity of each enzyme when isolated from either a strain harboring vector only or a plasmid producing *Sl*PatA^{WT}. *, $p<0.0001$

S. coelicolor. Since the *S. coelicolor* genome contains a gene encoding a *Sl*PatA homologue, we was surmised that *Sc*Acs was under RLA control.

Our initial work with the *S. lividans* *Sl*PatA^{WT} and *Sl*Acs^{WT} enzymes paint a complex picture for the regulation of *Sl*Acs^{WT} function in this organism, and by extrapolation, maybe in *S. coelicolor*. Because the specific activity of *Sl*Acs^{WT} is similar to that of *Se*Acs^{WT} *in vitro* (Fig. 5.5B), we hypothesize that *Sl*Acs^{WT} activity is also tightly controlled by *S. lividans*. To account for the inability of *Sl*PatA^{WT} to acetylate *Sl*Acs^{WT}, we propose that *Sl*PatA^{WT} has evolved unique strategies for substrate recognition, or *Sl*PatA^{WT} is not the primary modifier of *Sl*Acs^{WT}. We discuss each possibility further below.

In vitro, Sl*PatA^{WT}* *does not recognize* Sl*Acs^{WT}*. Pat homologues acetylate Acs in *R. palustris* and *S. enterica* (Starai & Escalante-Semerena, 2004c, Crosby et al., 2010). Clearly, acetylation of *Sl*Acs^{WT} by *Sl*PatA^{WT} does not occur efficiently *in vitro* or in a heterologous model system (Figs. 5.4B, 5.5A, 5.6A, 5.8) (Tucker & Escalante-Semerena, 2013). The following possibilities should be taken into consideration when thinking about the potential regulation of *Sl*Acs^{WT} by RLA. First, it is possible that *Sl*Acs^{WT} may have evolved to evade acetylation by *Sl*PatA^{WT}. Secondly, since *S. lividans* encodes ~72 predicted GNAT-type acetyltransferases (Pfam00583) it is possible that one of these GNATs, not *Sl*PatA^{WT}, acetylates *Sl*Acs^{WT}. And thirdly, the reversed domain organization of *Sl*PatA, relative to *Rp*Pat and *Se*Pat, may prevent recognition of *Sl*Acs^{WT} by *Sl*PatA^{WT}.

***Substantial changes in the* C*-terminal domain of* Sl*Acs^{WT}* *lead to its recognition by* Sl*Pat^{WT}*.** Forty-one amino acid changes in the *C*-terminal domain of *Sl*Acs^{WT} were needed to allow *Sl*Pat^{WT} to recognize and acetylate *Sl*Acs (Fig. 5.3). If we assumed that the domain organization of *Sl*Pat^{WT} was not a factor in *Sl*Acs^{WT} recognition, such a large number of substitutions would suggest that the protein underwent evolutionary changes to avoid modification by *Sl*Pat^{WT}. Our results indicate, however, that *Sl*Acs recognition by *Sl*PatA^{WT} is reversible by mutation, and that the resulting *Sl*Acs variant can be reversibly acetylated. Below, we discuss the location of the amino acid changes that turn *Sl*Acs^{WT} into a substrate for *Sl*Pat^{WT} and their importance in recognition by *Sl*PatA^{WT}.

How do changes in the* C*-terminal domain of* Sl*Acs affect its acetylation and activity? Studies of *R. palustris* Pat (*Rp*Pat) substrate specificity indicate that this enzyme recognizes a loop >20 residues *N*-terminal to the target lysine in the substrate protein, suggesting that the *Rp*Pat interacts with a relatively large surface of substrate proteins (Crosby et al., 2012b). Here, we demonstrate that the identities of residues ranging from 8-52 amino acids *N*-terminal to the target lysine of *Se*Acs^{WT}, and 5-17 amino acids *C*-terminal to the target lysine of *Se*Acs^{WT} are critical for recognition of this substrate by *Sl*PatA^{WT} in isolation. Introduction of these sequences into *Sl*Acs, increased the regulation of this enzyme *in vitro* (Fig. 5.5B) and *in vivo* in a heterologous host (Figs. 5.6A, 7). The importance of the combination of *N*- and *C*-terminal residues to the K610 acetylation site is highlighted by the introduction of either the regions yielding only a 2-3 fold increase in acetylation of the substrate (Fig. 5.4A). This indicates that *Sl*PatA^{WT} recognizes several regions of the *Se*Acs *C*-terminal domain including the target lysine, residues *N*-terminal to the target lysine, and residues *C*-terminal to the target lysine. It is possible that these regions of the *Se*Acs *C*-terminal domain are necessary for direct interactions with the *Sl*PatA protein. Alternatively, these regions may be necessary to position the target lysine for entry into the *Sl*PatA active site. The crystal structure of *Sl*PatA^{WT} substrate *Se*Acs^{WT} is known (PDB 1PG3, 1PG4) (Gulick et al., 2003). Comparison of this structure with the structure of *Sl*Acs (structure not known) may distinguish these possibilities. Efforts to obtain the crystal structure of *Sl*Acs are ongoing.

Is* Sl*Acs*^{WT} *regulated by one or more protein acetyltransferases? As mentioned above, *Sl*Acs^{WT} may have evolved to evade acetylation specifically by *Sl*PatA^{WT}, but not by one of the additional 72 predicted GNAT-type acetyltransferases (Pfam00583) encoded by the genome of this bacterium. It is possible that one of these GNATs acetylates *Sl*Acs^{WT} more efficiently than *Sl*PatA^{WT} does. Such a scenario would not be unprecedented. It is known that the genome of *R. palustris* encodes a Pat homologue and a single-domain GNAT protein acetyltransferase that share overlapping protein acetyltransferase substrates, and that both enzymes acetylate these substrates with different affinities (Crosby et al., 2012a). Therefore, the possibility of *Sl*PatA^{WT} not being the sole regulator of *Sl*Acs^{WT} activity in *S. lividans* needs to be further investigated.

***Does the unique domain organization of* Sl*PatA*^{WT} *affect substrate specificity*?** Pat homologues are two-domain proteins composed of a GNAT (acetyltransferase) domain and a large domain of unknown function. In *Sl*PatA^{WT}, the GNAT domain is located at the *N*-terminus of the protein (Tucker & Escalante-Semerena, 2013). In contrast, in *R. palustris* and *S. enterica*, the domain order is reversed (i.e., GNAT domain is at the *C*-terminus of protein). *Sl*PatA^{WT} also has a collagen-like Gly-Pro-Ser motif in the large domain (Tucker & Escalante-Semerena, 2013). *S. enterica* and *R. palustris* Pat homologues efficiently acetylate their cognate Acs enzymes *in vitro* (Starai & Escalante-Semerena, 2004c, Crosby et al., 2010). The alternate domain organization of *Sl*PatA^{WT} may account for the poor acetylation of *Sl*Acs^{WT} compared to *Se*Acs^{WT} and *Sl*AacS^{WT} *in vitro* (Tucker & Escalante-Semerena, 2013). If *Sl*PatA^{WT} has evolved strategies for recognition of protein substrates differently from *Se*Pat and *Rp*Pat, our *in vitro* assay may be missing a factor that promotes efficient *Sl*PatA^{WT} recognition of *Sl*Acs^{WT} such as a small molecule, macromolecule (e.g., protein), or an as-yet-unidentified intracellular condition. If this were the case, the amino acid changes introduced into *Sl*Acs^{WT} to generate the *Sl*Acs-*Se*Acs chimera C3 obviate the need for additional factors or conditions for *Sl*PatA^{WT} recognition.

REFERENCES

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

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

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

4. **Brownell, J. E., and C. D. Allis.** 1995. An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena macronuclei*. *Proc. Natl. Acad. Sci. U S A* **92:**6364-6368.

5. **Chan, C. H., J. Garrity, H. A. Crosby, and J. C. Escalante-Semerena.** 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.

6. **Crosby, H. A., E. K. Heiniger, C. S. Harwood, and J. C. Escalante-Semerena.** 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.

7. **Crosby, H. A., D. A. Pelletier, G. B. Hurst, and J. C. Escalante-Semerena.** 2012. System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J. Biol. Chem.* **287:**15590-15601.

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

9. **Elion, E. A., P. Marina, and L. Yu.** 2007. Constructing recombinant DNA molecules by PCR, p. Unit 3.17.1-3.17.12. *In* F. M. Ausubel, R. E. R. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, a. Smith, and K. Struhl (ed.), *Current protocols in molecular biology*. Greene Publishing Associates & Wiley Interscience, New York, N.Y.

10. **Gardner, J. G., F. J. Grundy, T. M. Henkin, and J. C. Escalante-Semerena.** 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD(+) involvement in *Bacillus subtilis*. *J. Bacteriol*. **188:**5460-5468.

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

12. **Gulick, A. M.** 2009. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol*. **4:**811-827.

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

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

15. **Horton, R. M., S. N. Ho, J. K. Pullen, H. D. Hunt, Z. Cai, and L. R. Pease.** 1993. Gene splicing by overlap extension. *Methods Enzymol*. **217:**270-279.

16. **Khan, S. R., J. Gaines, R. M. Roop, 2nd, and S. K. Farrand.** 2008. Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl. Environ. Microbiol*. **74:**5053-5062.

17. **Kouzarides, T.** 2007. Chromatin modifications and their function. *Cell* **128:**693-705.

18. **Kuo, Y. M., and A. J. Andrews.** 2013. Quantitating the specificity and selectivity of Gcn5-mediated acetylation of histone H3. *PLoS One* **8:**e54896.

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

20. **Lewis, R. A., E. Laing, N. Allenby, G. Bucca, V. Brenner, M. Harrison, A. M. Kierzek, and C. P. Smith.** 2010. Metabolic and evolutionary insights into the closely-related species S*treptomyces coelicolor* and *Streptomyces lividans* deduced from high-resolution comparative genomic hybridization. *BMC Genomics* **11:**682.

21. **Liang, W., A. Malhotra, and M. P. Deutscher.** 2011. Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. *Mol. Cell* **44:**160-166.

22. **Mikulik, K., J. Felsberg, E. Kudrnacova, S. Bezouskova, D. Setinova, E. Stodulkova, J. Zidkova, and V. Zidek.** 2012. CobB1 deacetylase activity in *Streptomyces coelicolor*. *Biochem. Cell Biol.* **90:**179-187.

23. **Reger, A. S., R. Wu, D. Dunaway-Mariano, and A. M. Gulick.** 2008. Structural characterization of a 140° domain movement in the two-step reaction catalyzed by 4-chlorobenzoate:CoA ligase. *Biochemistry* **47:**8016-8025.

24. **Rocco, C. J., K. L. Dennison, V. A. Klenchin, I. Rayment, and J. C. Escalante-Semerena.** 2008. Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli. Plasmid* **59:**231-7.

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

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

27. **Sasse, J.** 1991. Detection of proteins, p. 10.6.1-10.6.8. *In* F. A. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), *Current Protocols in Molecular Biology*, vol. 1. Wiley Interscience, New York.

28. **Soppa, J.** 2010. Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea* **2010:**pii: 820681.

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

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

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

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

33. **Starai, V. J., H. Takahashi, J. D. Boeke, and J. C. Escalante-Semerena.** 2003. Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae*. *Genetics* **163:**545-555.

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

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

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

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

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

CHAPTER 6

INTERACTIONS OF A PROTEIN ACETYLTRANSFERASE DOMAIN WITH AN ACID:COA LIGASE (AMP-FORMING)[1]

[1] Tucker A.C., Rank K.C., Rayment, I. and J.C. Escalante-Semerena. To be submitted to *Molecular Cell*.

ABSTRACT

Reversible lysine acetylation by protein acetyltransferases is a conserved regulatory mechanism that controls diverse cellular pathways. Gcn5-related *N*-acetyl transferases (GNATs), named after their founding member, are found in all domains of life. Crystal structures of Gcn5 bound to peptide substrates revealed substrate-specific contacts with the target lysine and nearby amino acid residues in the acceptor substrate. Here, we map a broader constellation of interactions between a GNAT domain from *S. lividans* and a protein substrate, namely the *C*-terminal domain of *S. enterica* acetyl-CoA synthetase. Negatively charged residues surrounding the GNAT active site and positively charged residues near the target lysine of the acceptor substrate were critical for protein-protein interactions. Using a computer-generated interaction model, we prepared a crosslinked complex of GNAT and the protein substrate that retains catalytic activity, suggesting that the orientation and interactions caused by the restrictive disulfide crosslink represent a physiologically relevant acetylation complex.

INTRODUCTION

Reversible lysine acetylation is an epigenetic regulatory mechanism conserved in all domains of life (44). Proteins targeted by acetylation are involved in diverse cellular processes including chromatin maintenance (27), transcriptional regulation (19), protein stability (33), primary metabolism (13), and cell structure (25). Protein acetyltransferases mediate acetylation of specific lysines within target proteins (2, 13, 30). GNATs (Gcn5-related *N*-acetyltransferases) are present in all domains of life, and were first identified for their role in acetylating DNA-binding histones in eukaryotes (5). The γ-proteobacterium *Salmonella enterica* encodes a multi-domain protein acetyltransferase Pat (*Se*Pat) that contains a GNAT domain fused to the *C*-terminus of a domain of unknown function (Fig. 1A). In this bacterium, *Se*Pat acetylates and inactivates the AMP-forming acetate:CoA ligase enzyme (aka, Ac-CoA synthetase or *Se*Acs), which is involved in short-chain fatty acid utilization (47). *Se*Pat acetylates the active-site residue K609 rendering and enzyme (*Se*Acs^{Ac}) unable to convert acetate to acetyl-AMP (45).

GNATs, including *Se*Pat homologues, acetylate both AMP-forming CoA ligases and structurally unrelated proteins. *R. palustris* encodes a *Se*Pat homologue in addition to a single-domain GNAT (KatA, Fig. 6.1A). *Rp*Pat and *Rp*KatA acetylate specific members of the AMP-forming CoA ligase family of enzymes from *R. palustris* (13). Some *Se*Pat protein substrates are not members of the AMP-forming CoA ligase family. In addition to Acs, *E. coli* and *Se*Pat homologues acetylate ribonuclease R (33) and the transcription factor RcsB (48), respectively. Neither ribonuclease R nor RcsB bears sequence homology to AMP-forming CoA ligases, suggesting that *Se*Pat homologues can acetylate diverse substrates but exhibits substrate specificity as seen with *Rp*Pat acetylation of AMP-forming CoA ligases (14).

Our understanding of the interactions of GNAT proteins with their protein substrates is limited. Three-dimensional crystal structures of the *Tetrahymena* Gcn5 protein in complex with a peptide substrate revealed local interactions between the GNAT and its substrate including the target lysine (8, 40). The *Mycobacterium tuberculosis* protein acetyltransferase PatA (*Mt*PatA) is a two-domain GNAT protein acetyltransferase also known to acetylate members of the AMP-forming CoA ligase family (32). The structure of this enzyme was recently reported. In this report, Lee *et al.* speculated on putative hydrophobic interactions of the enzyme with the lysine side chain of the substrate.

To gain insights into the constellation of interactions between a GNAT and a protein substrate we determined the structure of the GNAT domain of the *S. lividans* PatA (*Sl*PatA) enzyme (Fig. 6.1A), and used the ClusPro 2.0 software program (http://cluspro.bu.edu/login.php) to identify potential interacting surfaces between the GNAT domain of *Sl*PatA and the *C*-terminal domain of *Se*Acs (hereafter *Se*Acs) for which a three-dimensional crystal structure was available (RCSB PDB codes 1PG3, 1PG4) (23). We demonstrate that the GNAT domain of *Sl*PatA acetylated *Se*Acs in *vivo* and *in vitro*, and that reversing charges near a predicted interface prevented interaction and *Se*Acs acetylation. To test the biological relevance of the interaction model we tethered the GNAT to its protein substrate 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 the substrate when the GNAT protein was linked to the protein substrate by direct disulfide bond, demonstrating catalysis in an enzyme-substrate

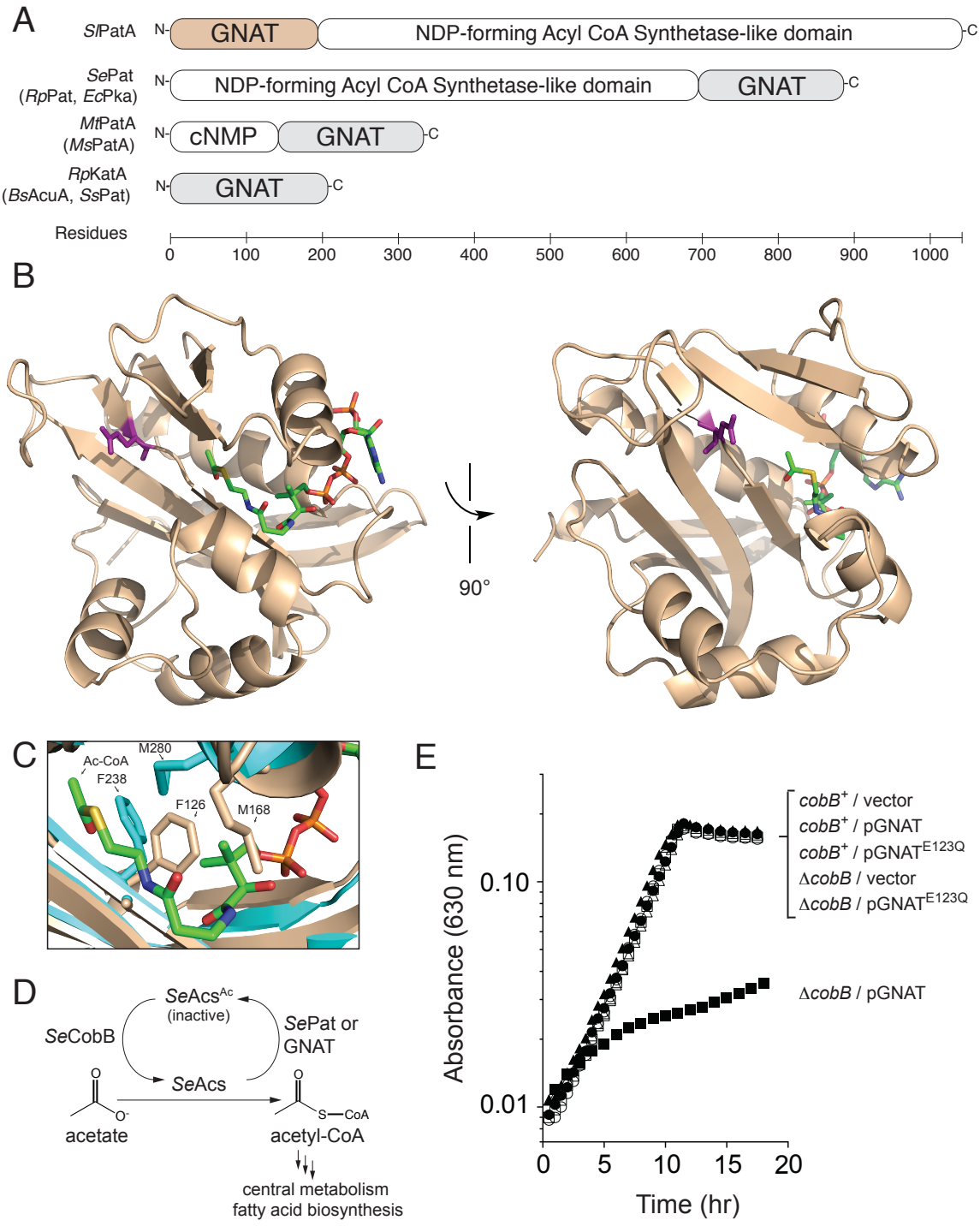


Figure 6.1. *Sl*PatA GNAT domain and *in vivo* function. (A) Domain arrangements of characterized protein acetyltransferases are listed for *Streptomyces lividans* (*Sl*PatA), *Rhodopseudomonas palustris* (*Rp*Pat, *Rp*KatA), *Salmonella enterica* (*Se*PatA), *Escherichia coli* (*Ec*Pka), *Mycobacterium tuberculosis*

(*Mt*PatA), *Bacillus subtilis* (*Bs*AcuA), and *Sulfolobus solfataricus* (*Ss*Pat). GNAT, Gcn5-related *N*-acetyltransferase; cNMP, cyclic mononucleotide monophosphate-binding domain. (B) A ribbon representation of *Sl*PatA GNAT (PBD Accession number XXXX) shown in beige solved to a resolution of 1.4 Å by X-ray crystallography. The predicted catalytic E123 is show in purple represented as sticks. Acetyl-CoA was modeled into the active site based on an alignment with the *Mycobacterium tuberculosis* PatA protein (PDB accession code 4AVB). (C) Alignment of *Sl*PatA GNAT (beige) with *Mycobacterium tuberculosis* PatA (teal) shows the location of *Sl*PatA GNAT residues F126 and M168 that overlap with the modeled Ac-CoA. Homologous residues of *Mt*PatA (F238 and M 280) are shown accommodate the bound Ac-CoA. *Sl*PatA GNAT active site E123 is colored purple and putative active site *Mt*PatA E235 is shown in dark teal. (D) Regulation of *Se*Acs activity by GNAT in *Salmonella enterica*. *Se*Acs converts acetate to acetyl-CoA and can be modified by *Se*Pat or ectopically produced GNAT. The sirtuin *Se*CobB deacetylates and reactivates $Se\text{Acs}^{Ac}$. (E) Growth of *S. enterica* Δ*pat cobB*$^{+}$ and *S. enterica* Δ*pat* Δ*cobB* strains producing GNAT variants 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 <0.008 absorbance units.

complex with potentially limited movement. Together these data support a biologically relevant interaction model between the GNAT and a protein substrate that reveals the importance of electrostatic interactions between the proteins.

EXPERIMENTAL PROCEDURES

Bacterial strains and growth conditions. All strains and plasmids used in this study are listed in the Table 6.1 and Table 6.2, respectively. *Escherichia coli* and *Salmonella enterica* strains were grown at 37°C in lysogeny broth (LB, Difco) (3) or no-carbon essential (NCE) minimal medium (1) supplemented with sodium acetate (10 mM), $MgSO_4$ (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 *S. lividans Sl*PatA GNAT cloned into the expression vector pBAD30 (24).

Molecular techniques. DNA manipulations were performed using standard techniques (15). 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 sequencing reactions were resolved at the University of Georgia Genomics Facility. Oligonucleotide primer sequences are listed in the Table 6.3.

Plasmids used for protein overproduction. *GNAT overproduction plasmids.* The allele encoding the GNAT domain of *S. lividans patA* (M1-L194) was amplified from *S. lividans* TK24 genomic DNA with the primers listed in Table 6.3 for cloning into the overexpression vector pKLD66 (39). The first codon of *Sl*PatA (TTG) was changed to the more common ATG start codon. The GNAT DNA fragment for cloning into pTEV6 was digested with KpnI and SalI and ligated into pTEV6 (39) digested with the same enzymes. The resulting plasmid p*Sl*PatA14 directs synthesis of *Sl*PatA GNAT

Table 6.1. Strains used in this study.

Strain Name	Genotype	Reference[2]
E. coli strains		
JE9314	C41λ(DE3) *pka12*::*kan*$^+$	Lab Collection
JE14224	KDZif1ΔZ	(6)
JE18685	KDZif1ΔZ / pAC-VSVG-Zif-Ap pBRw-GP *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18687	KDZif1ΔZ / pAC-VSVG-Zif-Ap pSlPatA46 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18688	KDZif1ΔZ / pAC-VSVG-Zif-Ap pSlPatA47 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18689	KDZif1ΔZ / pAC-VSVG-Zif-Ap pSlPatA48 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18690	KDZif1ΔZ / pAC-VSVG-Zif-Ap pACS53 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18691	KDZif1ΔZ / pAC-VSVG-Zif-Ap pACS54 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18692	KDZif1ΔZ / pAC-VSVG-Zif-Ap pACS55 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18693	KDZif1ΔZ / pAC-VSVG-Zif-MvaT pBRw-MvaT *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18695	KDZif1ΔZ / p*Sl*PatA43 pBRw-GP *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18696	KDZif1ΔZ / p*Sl*PatA43 pACS53 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18697	KDZif1ΔZ / p*Sl*PatA43 pACS54 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18698	KDZif1ΔZ / p*Sl*PatA43 pACS55 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18699	KDZif1ΔZ / p*Sl*PatA44 pBRw-GP *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18700	KDZif1ΔZ / p*Sl*PatA44 pACS53 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18701	KDZif1ΔZ / p*Sl*PatA44 pACS54 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18702	KDZif1ΔZ / p*Sl*PatA44 pACS55 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18703	KDZif1ΔZ / p*Sl*PatA45 pBRw-GP *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18704	KDZif1ΔZ / p*Sl*PatA45 pACS53 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18705	KDZif1ΔZ / p*Sl*PatA45 pACS54 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18706	KDZif1ΔZ / p*Sl*PatA45 pACS55 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18707	KDZif1ΔZ / pACS50 pBRw-GP *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18708	KDZif1ΔZ / pACS50 p*Sl*PatA46 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18709	KDZif1ΔZ / pACS50 p*Sl*PatA47 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18710	KDZif1ΔZ / pACS50 p*Sl*PatA48 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18711	KDZif1ΔZ / pBRw-GP pACS51 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18712	KDZif1ΔZ / pBRw-GP pACS52 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18713	KDZif1ΔZ / p*Sl*PatA46 pACS51 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18714	KDZif1ΔZ / p*Sl*PatA46 pACS52 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18715	KDZif1ΔZ / p*Sl*PatA47 pACS51 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18716	KDZif1ΔZ / p*Sl*PatA47 pACS52 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18717	KDZif1ΔZ / p*Sl*PatA48 pACS51 *kan*$^+$ *bla*$^+$ *tet*$^+$	
JE18718	KDZif1ΔZ / p*Sl*PatA48 pACS52 *kan*$^+$ *bla*$^+$ *tet*$^+$	
S. enterica strains		
JE16236	*metE205 ara-9* Δ*pat3*	
JE16237	*metE205 ara-9* Δ*pat2* Δ*cobB1375*	

[2] Unless otherwise indicated, all strains and plasmids were constructed during the course of this work

Derivatives of JE16236		
JE16416	*metE205 ara-9* Δ*pat3* / pBAD30 *bla*$^{+}$	
JE17191	*metE205 ara-9* Δ*pat3* / p*Sl*PatA11 *bla*$^{+}$	
JE17922	*metE205 ara-9* Δ*pat3* / p*Sl*PatA21 *bla*$^{+}$	
JE18755	*metE205 ara-9* Δ*pat3* / p*Sl*PatA51 *bla*$^{+}$	
JE18756	*metE205 ara-9* Δ*pat3* / p*Sl*PatA52 *bla*$^{+}$	
Derivatives of JE16237		
JE16419	*metE205 ara-9* Δ*pat2* Δ*cobB1375* / pBAD30 *bla*$^{+}$	
JE17193	*metE205 ara-9* Δ*pat2* Δ*cobB1375* / p*Sl*PatA11 *bla*$^{+}$	
JE17894	*metE205 ara-9* Δ*pat2* Δ*cobB1375* / p*Sl*PatA21 *bla*$^{+}$	
JE18757	*metE205 ara-9* Δ*pat2* Δ*cobB1375* / p*Sl*PatA51 *bla*$^{+}$	
JE18758	*metE205 ara-9* Δ*pat2* Δ*cobB1375* / p*Sl*PatA52 *bla*$^{+}$	

Table 6.2. Plasmids used in this study.

Plasmid	Genotype	Reference[d]
pTEV5	*N*-terminal, rTEV-cleavable His_6-tag overexpression vector, *bla*$^+$	(39)
pBAD30	P_{araBAD} expression vector, *bla*$^+$	
pKLD66	*N*-terminal, rTEV-cleavable His_6-MBP-tag overexpression vector, *bla*$^+$	
pBRωGP	*N*-terminal RNAP ω fusion vector for bacterial 2-hybrid assay, *bla*$^+$	
pACTR-V-Zif-AP	*N*-terminal VSV-G-Zif268 fusion vector for bacterial 2-hybrid assay, *tet*$^+$	
pTEV5 derivatives		
pACS38	*Se*Acs C-terminal domain in pTEV5, *bla*$^+$	
pACS42	*Se*Acs^{R606E} C-terminal domain in pTEV5, *bla*$^+$	
pACS44	*Se*Acs^{R613D} C-terminal domain in pTEV5, *bla*$^+$	
pACS56	*Se*Acs^{A538C} C-terminal domain in pTEV5, *bla*$^+$	
pACS57	*Se*Acs^{H567C} C-terminal domain in pTEV5, *bla*$^+$	
pACS58	*Se*Acs^{D600C} C-terminal domain in pTEV5, *bla*$^+$	
pACS60	SeAcs^{R606E} in pTEV5, *bla*$^+$	
pACS61	SeAcs^{R613D} in pTEV5, *bla*$^+$	
pSlPatA27	GNAT in pTEV5, *bla*$^+$	
pSlPatA33	GNAT^{E121R} in pTEV5, *bla*$^+$	
pSlPatA35	GNAT^{D185R} in pTEV5, *bla*$^+$	
pKLD66 Derivatives		
pSlPatA14	GNAT in pKLD66, *bla*$^+$	
pSlPatA53	GNAT^{S73C} in pKLD66, *bla*$^+$	
pSlPatA54	GNAT^{A110C} in pKLD66, *bla*$^+$	
pSlPatA56	GNAT^{A164C} in pKLD66, *bla*$^+$	
pBAD30 Derivatives		
pSlPatA11	GNAT in pBAD30, *bla*$^+$	
pSlPatA21	GNAT^{E123Q} in pBAD30, *bla*$^+$	
pSlPatA51	GNAT^{E121R} in pBAD30, *bla*$^+$	
pSlPatA52	GNAT^{D185R} in pBAD30, *bla*$^+$	
pBRωGP derivatives		
pSlPatA46	GNAT in pBRωGP, *bla*$^+$	
pSlPatA47	GNAT^{E121R} in pBRωGP, *bla*$^+$	
pSlPatA48	GNAT^{D185R} in pBRωGP, *bla*$^+$	
pACS53	*Se*Acs CTD^{WT} in pBRωGP, *bla*$^+$	
pACS54	*Se*Acs CTD^{R606E} in pBRωGP, *bla*$^+$	
pACS55	*Se*Acs CTD^{R613D} in pBRωGP, *bla*$^+$	

pACTR-V-Zif-AP derivatives		
pSlPatA43	GNAT in pACTR-V-Zif-AP, *tet*$^{+}$	
pSlPatA44	GNAT^{E121R} in pACTR-V-Zif-AP, *tet*$^{+}$	
pSlPatA45	GNAT^{D185R} in pACTR-V-Zif-AP, *tet*$^{+}$	
pACS50	*Se*Acs CTD^{WT} in pACTR-V-Zif-AP, *tet*$^{+}$	
pACS51	*Se*Acs CTD^{R606E} in pACTR-V-Zif-AP, *tet*$^{+}$	
pACS52	*Se*Acs CTD^{R613D} in pACTR-V-Zif-AP, *tet*$^{+}$	

Table 6.3. Primers used in this study.

Primer Name	Primer Sequence[3][4][5]
*Se*Acs CTD pTEV5 F	GTAGCTAGCGACGTGTTAAACGTCTCC
*Se*Acs CTD pTEV5 R	ATTGAATTCTTATGACGGCATCGCGATGGC
*Se*Acs^{R606E} F	CTCACTGCCAAAAACC**GAA**TCCGGCAAAATTATGCGC
*Se*Acs^{R606E} R	GCGCATAATTTTGCCGGA**TTC**GGTTTTTGGCAGTGAG
*Se*Acs^{R613D} F	CCGGCAAAATTATGCGC**GAT**ATTTTGCGCAAAATCGCG
*Se*Acs^{R613D} R	CGCGATTTTGCGCAAAAT**ATC**GCGCATAATTTTGCCGG
*Se*Acs^{A538C} F	GAAATCGAGTCAGCGCTGGTG**TGC**CATCCGAAGATCGCCGAAGC
*Se*Acs^{A538C} R	GCTTCGGCGATCTTCGGATG**GCA**CACCAGCGCTGACTCGATTTC
*Se*Acs^{H567C} F	GTGACGCTCAAC**TGC**GGCGAGGAGCC
*Se*Acs^{H567C} R	GGCTCCTCGCC**GCA**GTTGAGCGTCAC
*Se*Acs^{D600C} F	GTGCTGCACTGGACC**TGC**TCACTGCCAAAAAC
*Se*Acs^{D600C} R	GTTTTTGGCAGTGA**GCA**GGTCCAGTGCAGCAC
*Se*Acs^{R606E} F	CTCACTGCCAAAAACC**GAA**TCCGGCAAAATTATGCGC
*Se*Acs^{R606E} R	GCGCATAATTTTGCCGGA**TTC**GGTTTTTGGCAGTGAG
*Se*Acs^{R613D} F	CCGGCAAAATTATGCGC**GAT**ATTTTGCGCAAAATCGCG
*Se*Acs^{R613D} R	CGCGATTTTGCGCAAAAT**ATC**GCGCATAATTTTGCCGG
GNAT pTEV5 F	GTAGCTAGCATGTCGTACGCGAGCCGTAC
GNAT pTEV5 R	ATTGAATTCTTAGAGATCGAACTCCAG
GNAT pKLD66 F	GTAGGTACCATGTCGTACGCGAGCCGTACTCTGG
GNAT pKLD66 R	ACTGTCGACTCAGAGATCGAACTCCAGGCGGACCACC
GNAT^{S73C} F	CGTACCCTCGCCTG**TGC**GCCAAGGACGTCCAC
GNAT^{S73C} R	GTGGACGTCCTTGGC**GCA**CAGGCGAGGGTACG
GNAT^{A110C} F	GACCGGATCGGG**TGC**GGCGGAACAC
GNAT^{A110C} R	GTGTTCCGCC**GCA**CCCGATCCGGTC
GNAT^{A164C} F	CTACGACCGGATCGGG**TGC**GGCGGAACACCCGC
GNAT^{A164C} F	GCGGGTGTTCCGCC**GCA**CCCGATCCGGTCGTAG
GNAT pBAD30 F	ACCGAATTC*AAGGGAAACCGGT*ATGTCGTACGCGAGCCGTACTC
GNAT pBAD30 R	GTAGGTACCTCAGAGATCGAACTCCAGGCGGACCACC
GNAT^{E123Q} F	GGCCGACGAGGCC**CAG**GTCGCCTTCCTC
GNAT^{E123Q} R	GAGGAAGGCGAC**CTG**GGCCTCGTCGGCC
GNAT^{E121R} F	CACCGCACCGGCCGAC**CGT**GCCGAGGTCGCCTTC
GNAT^{E121R} R	GAAGGCGACCTCGGC**ACG**GTCGGCCGGTGCGGTG
GNAT^{D185R} F	GAAGCGAAGCTTCGAG**CGT**GGGGTGGTCCGCCTG
GNAT^{D185R} R	CAGGCGGACCACCCC**ACG**CTCGAAGCTTCGCTTC
GNAT TwoHyb F	GCTGCGGCCGCCATGTCGTACGCGAGCCGTAC
GNAT TwoHyb R	ATTGGATCCTTAGAGATCGAACTCCAG
*Se*Acs CTD TwoHyb F	CAGGCGGCCGCCGACGTGTTAAACGTCTCC
*Se*Acs CTD TwoHyb R	ATTGGATCCTTATGACGGCATCGCGATGGC

[3] Restriction nuclease sites are underlined
[4] Bold typeface indicates the codon introduced during site-directed-mutagenesis.
[5] Italic typeface indicates a 13 nt region found 5' of *S. enterica pat* that includes the *pat* RBS

with an *N*-terminal H_6-Maltose binding protein tag cleavable by recombinant Tobacco Etch Virus (rTEV) protease prepared as described (4). Plasmids directing synthesis of $GNAT^{S73C}$, $GNAT^{A110C}$, and $GNAT^{A164C}$ variants were generated from the p*Sl*PatA14 plasmid using site-directed mutagenesis with the primers listed in the Table 6.3.

Se*Acs* C-*terminal domain overexpression plasmids.* The allele encoding the *C*-terminal domain of *Se*Acs (D518-S652) was amplified from *S. enterica* LT2 DNA with the primers listed in the Table 6.3 for cloning into the overexpression vector pTEV5. The GNAT DNA fragment for cloning into pTEV5 was digested with NheI and EcoRI and ligated into pTEV5 digested with the same enzymes. The resulting plasmids pACS38 directs synthesis of *Se*Acs CTD with an *N*-terminal H_6 tag cleavable by rTEV protease. Plasmids directing synthesis of *Se*Acs CTD ^{R606E}, *Se*Acs CTD ^{R613D}, *Se*Acs CTD ^{A238C}, *Se*Acs CTD ^{H567C}, and *Se*Acs CTD ^{D600C} were generated from pACS38 using site-directed mutagenesis with the primers listed in Table 6.3.

*Full-length SeAcs overproduction plasmids***.** Plasmids directing the synthesis of *Se*Acs^{R606E} and *Se*Acs^{R613D} were generated from pACS33 using site-directed mutagenesis with the primers listed in the Table 6.3. The resulting plasmids pACS60 and pACS61 direct synthesis of *Se*Acs^{R606E} and *Se*Acs^{R613D}, respectively, each encoding an *N*-terminal H_6 tag cleavable by rTEV protease.

***Overproduction and purification of H_6-tagged proteins*.** Sl*Pat GNAT purification*. Plasmids p*Sl*PatA14, p*Sl*PatA53, p*Sl*PatA54, p*Sl*PatA56 were transformed into the Δ*pka* derivative of *E. coli* C41λ(DE3) (strain JE9314). The resulting strains were grown overnight and sub-cultured 1:100 (v/v) into 2 liters of LB containing ampicillin (100 mg/ml). The cultures were grown shaking at 25°C to A_{600} ~ 0.7 and H_6-MBP-GNAT 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 resuspended in 30 ml cold His-Bind buffer (buffer A) [*tris*(hydroxymethyl)aminomethane-HCl (Tris-HCl) buffer (50 mM, pH 8), NaCl (500 mM)], and imidazole (5 mM)) containing phenylmethanesulfonylfluoride (PMSF, 1 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*. His_6-tagged GNAT 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 GNAT was eluted from the column with Buffer C [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (250 mM)]. All fractions containing H_6-MBP-GNAT were combined. rTEV protease was added to H_6-MBP-GNAT and the GNAT/rTEV mixture was incubated at room temperature for 3 h.. The GNAT/rTEV mixture was dialyzed at 4°C against buffer C (Tris-HCl (50 mM, pH 8), NaCl (500 mM)) twice for 3 hours 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 GNAT was analyzed by SDS-PAGE. Fractions containing GNAT were pooled together. GNAT 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 ethylenediaminetetraacetic acid (EDTA, 0.5 mM) were included in the storage buffer for GNAT cysteine variants. GNAT concentration was determined by measuring absorbance at 280 nm. The molar extinction coefficient used to calculate GNAT concentration was 17,420 M^{-1} cm^{-1}. The purification protocol for GNAT for crystallography 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 10.7 mg/ml before flash freezing into liquid nitrogen.

Selenomethionine-labeled GNAT 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 OD_{600} of ~1. The culture was cooled on ice to 16°C for 10 min, and a defined amino acid mixture containing selenomethionine (50 mg/L) was added to suppress methionine

biosynthesis (CITE). The culture was grown at 37°C for 30 min before the addition of IPTG (1 mM) to induce GNAT expression. The culture was grown overnight at 37°C and the protein was purified and stored for crystallography as described above.

Full-length Se*Acs variant and* Se*Acs CTD*^{WT} *purification.* Plasmids pACS38, pACS44, pACS56, pACS57, pACS58, pACS60, and pACS61 were transformed into a Δ*pka* derivative of *E. coli* C41λ(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 LB containing ampicillin (100 mg/ml). The cultures were grown shaking at 25 °C to A_{600} ~ 0.7 and protein synthesis was induced with IPTG (0.25 mM). Upon induction, the cultures were grown overnight at 25°C. Full-length *Se*Acs proteins and *Se*Acs CTD proteins were purified and stored as described above for GNAT with modifications. During the first purification step, full-length His_6-*Se*Acs proteins were washed with buffer B2 [Tris-HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (20 mM)] before elution with buffer C. His_6-*Se*Acs CTD proteins 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 full-length *Se*Acs proteins and *Se*Acs CTD 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 GNAT cysteine variants. The molar extinction coefficients used to calculate protein concentrations were 72,152 $M^{-1}cm^{-1}$ for full-length *Se*Acs and 15,470 M^{-1} cm^{-1} for *Se*Acs CTD.

*Se*Acs^{WT} was purified and stored as described (45). *Sl*PatA was overproduced and purified as described (51).

Construction of complementation plasmids. The allele encoding the GNAT domain of *S. lividans* PatA (M1-L194) was amplified from *S. lividans* TK24 genomic DNA with the primers listed in the Table 6.3 for cloning into the overexpression vector pBAD30 under the control of the *S. enterica pat* ribosome binding site. The first codon of *Sl*PatA (TTG) was changed to the more common ATG start

codon. The GNAT DNA fragment for cloning into pBAD30 was cut with EcoRI and KpnI and ligated into pBAD30 cut with the same enzymes. The resulting plasmid p*Sl*PatA11 expresses pBAD30 under the control of the arabinose-inducible P_{ara} promoter. The plasmid directing the synthesis of GNAT^{E123Q} was generated by site directed mutagenesis using the primers listed in the Table 6.3.

Plasmids for in vivo two-hybrid system assays. Plasmids pBRωGP and pACTR-V-Zif-AP were used as the bait and prey plasmids for the bacterial two-hybrid assay as described (6). Alleles encoding residues M1-L194 of *Sl*PatA GNAT, residues M1-L194 of *Sl*PatA GNAT^{E121R}, and residues M1-L194 of *Sl*PatA GNAT^{D185R} were amplified from plasmids pSlPatA27, pSlPatA33, and pSlPatA35, respectively, using the primers listed in the Table 6.3. Alleles encoding residues D518-S652 of the *Se*Acs CTD, residues D518-S652 of the *Se*Acs CTD ^{R606E}, and residues D518-S652 of the *Se*Acs CTD ^{R613D} were amplified from pACS38, pACS42, and pACS44, respectively, using the primers listed in the Table 6.3. DNA fragments were digested with NotI and BamHI DNA fragments were ligated into pACTR-V-Zif-AP and pBRωGP, which had been digested with NotI and BamHI. pACTR-V-Zif-AP-based plasmids direct synthesis of an *N*-terminal VGV-G epitope tag, a two-amino acid Gly-Ser linker, and the Zn-finger DNA-binding domain of the murine Zif268 protein (Zif) fused to the cloned allele (GNAT or *Se*Acs CTD). pBRω-GP-based plasmids direct the synthesis of and N-terminal *E. coli* RNA polymeras subunit ω (*rpoZ*) and a three-amino acid Ala-Ala-Ala linker fused to the cloned allele (GNAT or *Se*Acs CTD^{WT}). The resulting plasmids are listed in the Table 6.2.

Two-hybrid system assay and β-galactosidase activity measurement. *E. coli* strain KDZif1ΔZ harboring compatible plasmids were grown overnight in 1 ml NB supplemented with kanamycin (50 μg/ml) and carbenicillin (100 μg/ml). The following day, strains were sub-cultured 1:100 into 200 ml of NB supplemented with kanamycin (50 μg/ml), carbenicillin (100 μg/ml), 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 as described (21, 54), with modifications described (CITE: GRIFFITH 2002 and Zhang 1995), with modifications. Cultures were diluted 20 ml into 80 ml permeabilization solution (Na_2HPO_4 [100 mM], KCl [20 mM], $MgCl_2$ [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 samples were then diluted 25 ml into 150 ml of substrate solution (Na_2HPO_4 [60 mM], NaH_2PO_4 [40 mM], ortho-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 (Molecular Devices). A unit of activity was defined as the amount of enzyme required to catalyze the hydrolysis of 1 nmol of ONPG per min. Specific activity was reported as units per unit of absorbance at 650 nm (A_{650}). Experiments were run twice in biological duplicate. β-galactosidase activity of biological sample was measured in duplicate. Unpaired t-test was carried out on each plasmid combination relative to the double empty vector and both single empty vector controls. Values are reported as the mean ± SD.

AMP-forming CoA ligase assays. *Se*Acs^{WT} activity was measured using an NADH-consuming assay (12, 18). Reactions (100 µl total volume) contained 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (50 mM, pH 7.5), *tris*(2-chloroethyl) phosphate (TCEP, 1 mM), ATP (2.5 mM) CoA (0.5 mM), $MgCl_2$ (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 *Sl*AacS (15 nM). The absorbance at 340 nm was monitored in a 96-well plate using the Spectramax Plus UV-visible spectrophotometer (Molecular Devices). Enzyme activities were calculated as described (18). Values are reported as the mean ± SD.

In vitro *acetylation assays*. Protein acetylation was monitored using radiolabeled Ac-CoA as described (12, 47, 52). *Sl*PatA acetylation reactions contained 2(*bis*(2-hydroxyethyl)imino)-2-(hydroxymethyl)-1,3-propanediol (Bis-Tris-HCl) buffer (50 mM, pH 6.0), [1-^{14}C]-Ac-CoA (50 μM), Acs (3 μM), glycerol (10%, v/v), and *Sl*PatA^{WT} (1 μM). Acetylation reactions catalyzed by *Se*Pat^{WT} were performed in mixtures that contained HEPES buffer (50 mM, pH 7.5), [1-^{14}C]-Ac-CoA (50 μM), Acs (3 μM), and *Se*Pat^{WT} (1 μM). Reactions (100 μl total volume) were incubated for 90 min at 30°C or 37°C for *Sl*PatA^{WT} or *Se*Pat^{WT}, respectively. Samples (10 μl) were removed and quenched at 0, 15, 30, 60, and 90 min. Samples were resolved using SDS-PAGE (31) and proteins were visualize by Coomassie Blue staining (41). Gels were dried and exposed 16 h to a multipurpose phosphor screen (Packard). 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). Values are reported as the mean ± SD.

The effect of acetylation on activity of *Se*Acs^{WT} activity was determined as described (12) with modifications. *Se*Acs (3μM) was incubated with *Sl*PatA (1 μM) or *Se*Pat (1 μM) and 50 μM Ac-CoA for 90 min at 30°C (*Sl*PatA) or 37°C (*Se*PatA) 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. *Se*Acs activity was measured as described above.

GNAT-* S*eAcs CTD*^{WT} *crosslinking and acetylation assays. Cysteine variants of the *Se*Acs CTD^{WT} were exchanged into cross-linking buffer (20 mM HEPES, pH 7.5, 100 mM NaCl, sparged with N_2) using Zeba Spin desalting columns (Pierce). *Se*Acs *C*-terminal domain proteins were immediately diluted into a 10-fold molar excess of aldriothiol-2 (AT-2, 2.05-Å crosslinker, Sigma), *bis*(maleimido)ethane (BMOE, 8.0 Å crosslinker, Thermo), 1,4-*bis*(maleimido)butane (BMB, 10.9 Å crosslinker, Thermo), *bis*(maleimido)hexane (BMH, 16.0 Å crosslinker, Thermo), or 1,4-di-(3'-[2'-pyridyldithio]-propionamido)butane (DPDPB, 19.9 Å crosslinker, bioWorld), and allowed to react for 1 h at room temperature. Activated *Se*Acs CTD variants and GNAT cysteine variants were desalted into

cross-linking buffer as described above. GNAT variants were diluted into a 5-fold molar excess of *Se*Acs CTD and allowed to react for 2 hours 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 crosslinked species.

Crosslinked species were incubated with Bis-Tris (100 mM, pH 6.0) and [1-^{14}C]-Ac-CoA (20 μM; 56.8 mCi mmol^{-1}) at 30ºC for 1 h. Reactions were resolved by non-reducing SDS-PAGE and analyzed as described above.

Crystallization of GNAT. GNAT crystals were grown by small-scale batch crystallization where 4 mL 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 mL of PEG solution (22% (w/v) monomethyl polyethylene glycol 5000 in 3-(N-morpholino)propanesulfonic acid (MOPS) buffer (100 mM, pH 7.0) containing $MgCl_2$ (25 mM), 2,2,2-trifluoroethanol (2% v/v), and glycerol (6%, v/v) (38). Drops were immediately streak-seeded then allowed to grow at RT for 1 week. Crystals grew to maximal dimensions of 300 mm x 100 mm x 100 mm. For freezing, crystals were transferred to PEG solution containing ethyl-CoA (1 mM) and then transferred stepwise to a solution of monomethyl polyethylene glycol 5000 (22%, w/v), glycerol (20%, v/v), MOPS buffer (100 mM, pH 7.0) containing $MgCl_2$ (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.

X-Ray data collection and structural refinement. X-Ray diffraction data for the GNAT crystals were collected at the SBC 19-ID beam line (Advanced Photon Source, Argonne, IL). The data sets were integrated and scaled with the program HKL2000 (36). X-ray data collection statistics are given in Table 6.4. The structure of GNAT 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 GNAT utilizing the programs SHELX, mlphare, DM, and ARP/wARP (7, 11, 42). 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 (7, 35). The native structure of GNAT was refined by iterative cycles of manual model building in Coot and restrained refinement in Refmac 5.6 (16, 43). Data processing and refinement statistics are presented in Table 6.4. The native GNAT L194 structures have been deposited with the PDB under accession code XXXX.

RESULTS

***Crystal structure of the GNAT domain of the* S. lividans PatA *enzyme at 1.45Å resolution*.** *Sl*PatA is a multidomain protein acetyltransferase with an *N*-terminal GNAT domain and a large *C*-terminal domain of unknown function (Fig 6.1A) (51). To characterize specific interactions of the GNAT domain of *Sl*PatA with protein substrates, we determined the structure of the *Sl*PatA GNAT domain, which contained the catalytic residue E123. For the sake of brevity, hereafter the *Sl*PatA GNAT domain will hereafter be referred to as GNAT. The three-dimensional crystal structure of the GNAT was determined at 1.45Å resolution (Fig. 6.1B, Table 6.4). The resulting model was structurally similar to the active-state GNAT domain of the *M. tuberculosis* PatA (*Mt*PatA, PDB 4AVB, rmsd = 2.2), a two-domain GNAT activated by binding of cyclic-AMP (53). The structure of *Mt*PatA was determined with acetyl-CoA in the active site, and this structure was used to model Ac-CoA into the GNAT structure. GNAT residues Phe126 and Met168 overlapped with the modeled Ac-CoA residue, thus are likely to undergo a shift upon Ac-CoA binding, as observed for the analogous residues of the Ac-CoA bound structure of *Mt*PatA (Phe238 and Met280, Fig. 6.1C).

Sl*PatA GNAT domain is catalytically active in vivo in the absence of the large domain.* We previously demonstrated that *Sl*PatA acetylates the acetyl-CoA synthetase (AMP-forming) from the Gram-negative enterobacterium *Salmonella enterica* (*Se*Acs) (51). *Se*Acs was the first metabolic protein identified as a target of acetylation, and the role of *Se*Acs in *S. enterica* acetate metabolism is well characterized (46). *Se*Acs is responsible for the activation of acetate during growth of *S. enterica* on low concentrations ($\leq$10 mM) of this short-chain fatty acid. *S. enterica* encodes a protein acetyltransferase

Table 6.4. Crystal Structure Data Collection and Refinement Statistics.

Data Collection	GNAT L194
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	40.1, 56.7, 74.0
α, β, γ (°)	90, 90, 90
Wavelength	0.98
Resolution (Å)[a]	25-1.45 (1.48-1.45)
R_{merge}[a]	3.8 (21.2)
$I/\sigma I$[a]	35.4 (10.1)
Completeness (%)[a]	99.8 (98.7)
Redundancy[a]	8.7 (6.6)
Refinement	
Resolution (Å)[a]	25-1.45 (1.48-1.45)
No. reflections[b]	29002 (1540)
R_{work} / R_{free}[c]	0.18/0.22
No. atoms	
Protein	1412
Water	190
Ligand	12
Average *B*-factor (Å^2)	15.4
Ramachandran (%)	
Most favored	98.25
Allowed	1.75
Disallowed	0
R.m.s. deviations	
Bond lengths (Å)	0.025
Bond angles (°)	2.441
PDB accession no.	XXXX

[a]Data in parentheses represent highest resolution shell.

[b]Data in parentheses represent the number of reflections used for the calculation of R_{free}.

[c] $$Rfactor = \sum \left| Fobs\text{-}Fcalc \right| / \sum \left| Fobs \right|$$ (Eq. S1)

Where R_{work} refers to the R_{factor} for the data utilized in the refinement and R_{free} refers to the R_{factor} for 5% of the data that were excluded from the refinement.

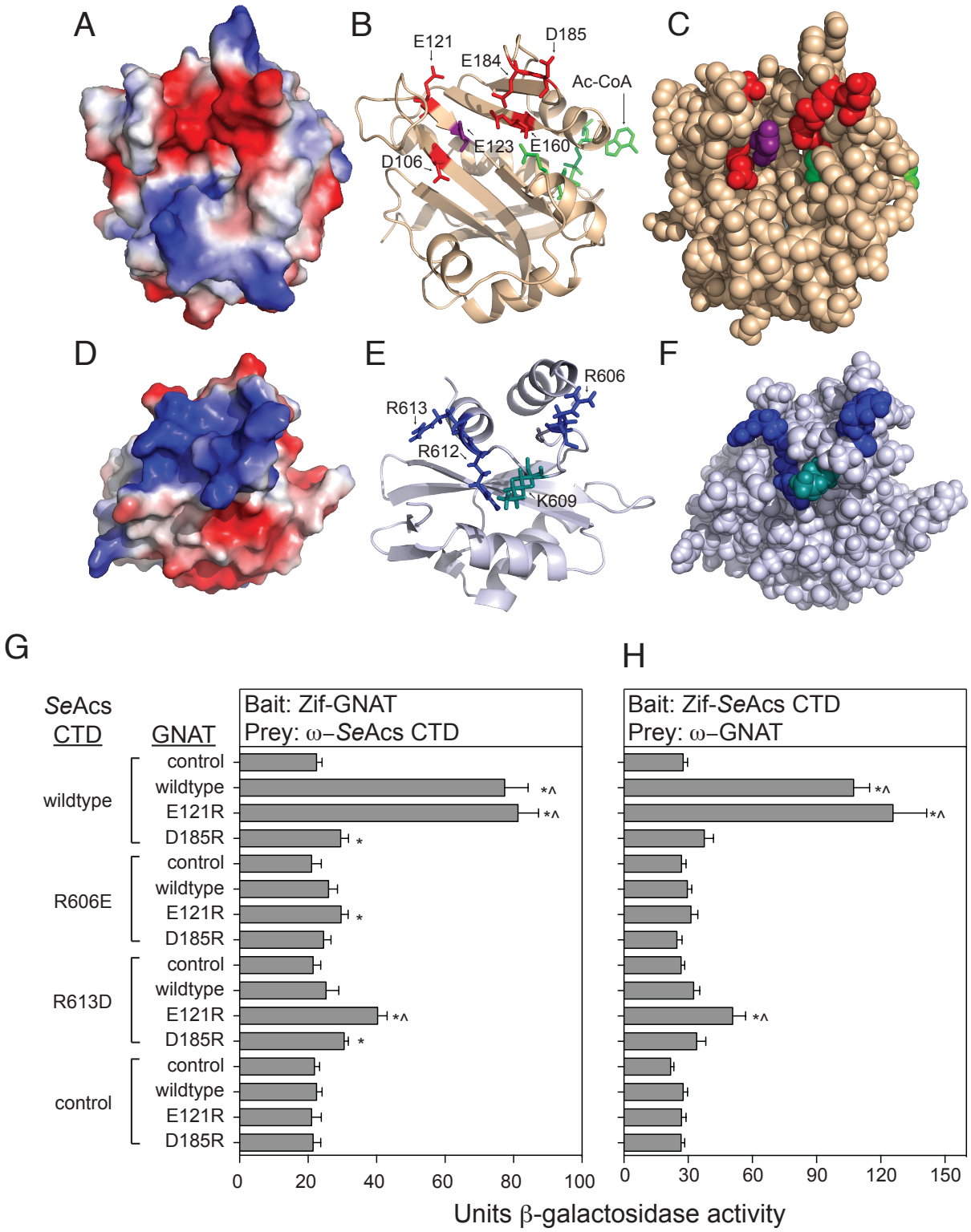


Figure 6.2. Charged surface residues mediate interactions between GNAT and protein substrate. GNAT (A) and *Se*Acs CTD^{WT} (D) as seen in a surface representation with electrostatic potential distribution colored red for negatively charged, white for neutral, and blue for positively charged areas.

Ribbon representations of (B) GNAT modeled with Ac-CoA with the negatively charged residues surrounding the active site E123 (in purple) shown in red and the active site glutamate shown in purple and (E) *Se*Acs CTD^{WT} with the positively charged residues surrounding the target K609 (in teal) in blue. (C,F) Representation of GNAT and *Se*Acs CTD^{WT} as spheres with the same coloring described above. (G) Effect of different Zif-GNAT variants and empty vector on transcription *in vivo* from promoter p*lac*Zif 1-61 with ω-*Se*Acs CTD^{WT} variants or empty vector. (H) Reciprocal bait-prey experiments were conducted with Zif-*Se*Acs CTD^{WT} and ω-GNAT fusions. *E. coli* strain KDZif1ΔZ harboring compatible plasmids directing synthesis of the indicated fusion proteins were grown in the presence of IPTG (200 μM) and assayed for β-galactosidase activity. Values represent mean ± standard deviation. Unpaired t-test was carried out on each plasmid combination relative to the double empty vector and both single empty vector controls. * indicates $p<0.0001$; ^ indicates that the interactions were detected in bait-prey reciprocal experiments.

*Se*Pat homologous to *Sl*PatA in which the domains are reversed (i.e., the *Se*Pat GNAT domain is located *C*-terminal to the large domain, Fig. 6.1A) (47). *Se*Pat acetylates the *Se*Acs catalytic K609 *in vivo* resulting in acetylated and inactivated *Se*Acs (*Se*Acs^{Ac}). The NAD^+-dependent sirtuin-type protein deacetylase *Se*CobB can deacetylate and reactivate *Se*Acs^{Ac} (45). When *S. enterica* is grown in the presence of low acetate concentrations (10 mM), *Se*CobB is required to maintain sufficient *Se*Acs in an active state (Fig. 6.1D). To determine whether or not the *Sl*PatA GNAT domain alone was capable of acetylating *Se*Acs, GNAT was expressed in *S. enterica* strains growing on 10 mM acetate as the sole carbon source. All experiments were conducted in an *S. enterica* Δ*pat* strain to eliminate the effects of *Se*Pat acetylation of *Se*Acs during growth of *S. enterica* on acetate. Expression of GNAT, but not the putative active site variant GNAT^{E123Q}, inhibited growth in a *S. enterica* Δ*pat* Δ*cobB* strain on acetate (Fig. 6.1E). Neither GNAT nor GNAT^{E123Q} inhibited growth of a *S. enterica* Δ*pat cobB*$^+$ strain. Together, these data demonstrated that residue E123 of the GNAT was required for activity, since in its absence (GNAT^{E123Q}) the *S. enterica* Δ*pat* Δ*cobB* strain grew well on acetate. In addition, GNAT-mediated growth inhibition was alleviated in a *Se*CobB-dependent manner (Fig. 6.1E). More importantly, these data indicated that GNAT interacted with and acetylated *Se*Acs *in vivo*.

Amino acid charge reversals at the predicted GNAT-protein substrate interface disrupt interaction. A surface representation of the GNAT revealed its active site residue E123 located within a negatively charged cleft composed of residues D106, E121, E160, E184, and D185 (Fig. 6.2A-C). As mentioned above, a model of the structure *Se*Acs is available (PDB code 1PG4). From this structure one can see that the acetylation site (K609) is located within the *C*-terminal domain of the protein. Notably, electron density was not observed for K609 in the reported *Se*Acs structure. For the purpose of our work, the most common rotamer of lysine was modeled into this structure for reference. The surface of the *Se*Acs *C*-terminal domain (hereafter referred to as *Se*Acs CTD^{WT}) has surface-exposed, positively charged residues R606, R612, and R613 surrounding K609 (Fig. 6.2D-F). Charge residues surrounding the active

site of GNAT and the acetylation site of *Se*Acs suggested that electrostatic interactions mediated GNAT-*Se*Acs binding during the acetylation reaction.

We used a bacterial two-hybrid assay to quantify the interactions of the GNAT and *Se*Acs CTD^{WT} *in vivo* when the charged residues on the protein surfaces were substituted with residues of the opposite charge. For all bacterial two-hybrid interaction experiments, GNAT variants and *Se*Acs CTD^{WT} variants were each tested as bait and prey (Fig. 6.2G-H). Interaction patterns were similar in each experiment, but only interactions found to be statistically significant in both bait-prey reciprocal experiments are discussed below.

As shown in figure 6.2G-H, variant GNAT^{D185R} showed significantly reduced interactions with *Se*Acs CTD^{WT}, whereas variant GNAT^{E121R} interacted with *Se*Acs CTD^{WT} despite the charge reversal, suggesting that the charge of this residue was not critical for the interaction between the proteins. Variants of the *Se*Acs CTD^{WT} were also constructed to test the role of the positively charged residues surrounding the aceylation target K609. *Se*Acs CTD^{R606E} and *Se*Acs CTD^{R613D} variants exhibited significantly reduced interactions with the GNAT (Fig. 6.2G-H).

Interactions were tested between each GNAT variant and each *Se*Acs CTD^{WT} variant to determine whether or not the double charge reversals restored interactions between the domains. As shown in figure 6.2G-H, weak interactions between GNAT^{E121R} and *Se*Acs CTD^{R613D} variants were detected, demonstrating that the positive charge introduced into the GNAT^{E121R} variant weakly, but detectably, compensated for the *Se*Acs CTD^{R613D} charge reversal that abolished interaction with the wild-type GNAT. The role of this interaction is further discussed below.

***Residue D185 of the GNAT mediates interaction with the protein substrate* in vivo.** We assessed the effects of the amino acid charge reversals on GNAT activity *in vivo* during growth on 10 mM acetate. Due to the strong interaction of the GNAT^{E121R} variant with *Se*Acs CTD^{WT} in the bacterial two-hybrid assay, we predicted the GNAT^{E121R} variant would interact with full-length *Se*Acs^{WT} *in vivo*. As shown in figure 6.3A, the GNAT^{E121R} variant significantly inhibited growth of a *S. enterica* Δ*pat* Δ*cobB* strain but not the growth of a *S. enterica* Δ*pat cobB*$^{+}$ strain, confirming that the GNAT^{E121R} variant

acetylated and inactivated *Se*Acs^{WT} *in vivo*. The GNAT^{D185R} variant only slightly inhibited growth of the *S. enterica* Δ*pat* Δ*cobB* strain, supporting our findings that the GNAT^{D185R} variant exhibited significantly weaker interactions with the *Se*Acs CTD^{WT} in the bacterial-two-hybrid assay (Fig. 6.2G-H). When present at higher levels, however, the GNAT^{D185R} variant inhibited the growth of the *S. enterica* Δ*pat* Δ*cobB* strain but not the growth of the *S. enterica* Δ*pat cobB*$^{+}$ strain (Fig. 6.3B). These results suggested that higher levels of GNAT^{D185R} variant compensated for the weakened interaction with full-length *Se*Acs^{WT} and inhibited growth of *S. enterica* on 10 mM acetate.

Se*Acs^{WT}* *residues R606 and R613 are required for efficient recognition by protein acetyltransferases.* The *Se*Acs CTD^{R606E} and *Se*Acs CTD^{R613D} variants did not interact with the GNAT as indicated by the bacterial two-hybrid interaction analysis. To determine whether or not *Se*Acs CTD^{WT} variants were misfolded, full-length *Se*Acs^{R606E} and *Se*Acs^{R613D} variants were isolated and their specific activities were compared to that of the *Se*Acs^{WT} enzyme. We obtained the following results: *Se*Acs^{WT}, 8.3 ± 0.5 μmol AMP min^{-1} mg^{-1}; *Se*Acs^{R606E}, 5.1 ± 0.2 μmol AMP min^{-1} mg^{-1}; and *Se*Acs^{R613D}, 2.8 ± 0.2 μmol AMP min^{-1} mg^{-1}. We interpreted these results to suggest that the variants were not misfolded when produced *in vivo,* and that the results of the bacterial two-hybrid assays were likely reflective of the presence or absence of interactions.

To determine whether or not the loss of interaction of *Se*Acs^{R606E} and *Se*Acs^{R613D} variants with *Sl*PatA^{WT} blocked acetylation, we measured the rate of acetylation of the *Se*Acs^{R606E} and *Se*Acs^{R613D} variants using full-length *Sl*PatA^{WT} enzyme and [1-^{14}C]-acetyl-CoA. As shown in figures 6.4A and 6.4B, *Se*Acs^{WT} enzyme was acetylated by *Sl*PatA^{WT} after 90 min, while the *Se*Acs^{R606E} and *Se*Acs^{R613D} variants were ~30% acetylated in the same time period. To determine whether or not these residues were broadly required for acetylation by Pat homologues, we measured the rate of acetylation of each variant by the *Se*Pat enzyme*,* the cognate acetyltransferase for *Se*Acs in *S. enterica*. *Se*Pat fully acetylated *Se*Acs^{WT} within 30 min, while *Se*Acs^{R606E} and *Se*Acs^{R613D} variants were acetylated ~15% and ~30%, respectively, after 90 min of incubation with *Se*Pat (Fig. 6.4A-B). The signal obtained in the phosphor image at time 0

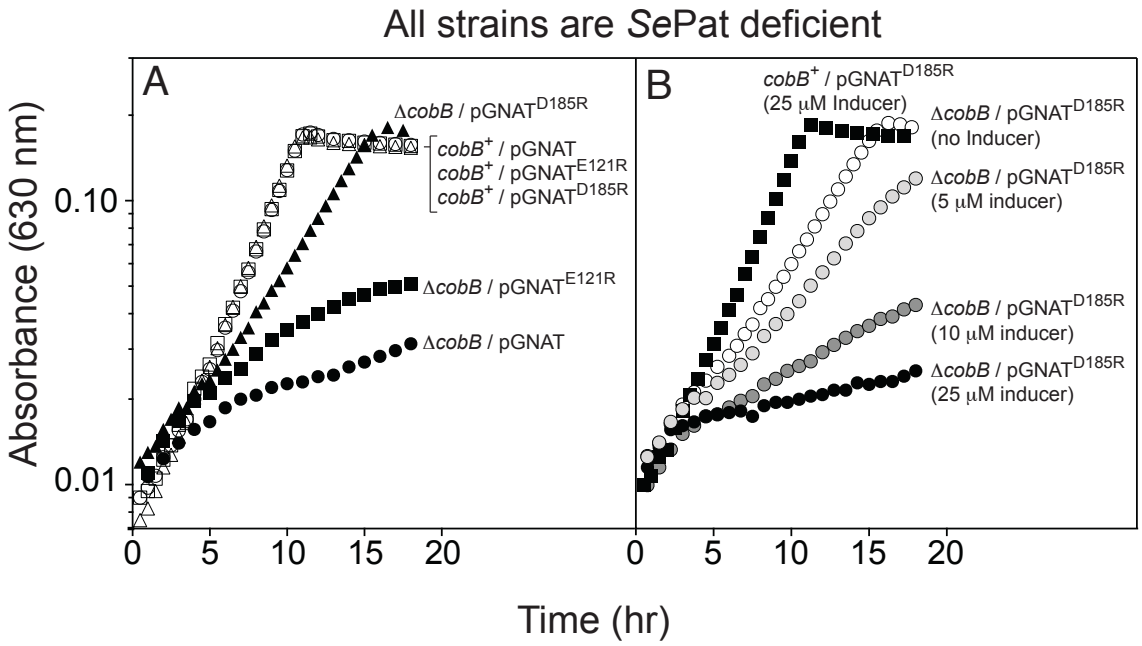


Figure 6.3. GNAT^{D185R} variant has decreased activity *in vivo*. (A) *S. enterica* Δ*pat cobB*$^{+}$ (open shapes) and *S. enterica* Δ*pat* Δ*cobB* (filled shapes) strains producing GNAT (circles), GNAT^{E121R} (squares), and GNAT^{D185R} (triangles) in the absence of inducer during growth on NCE minimal medium supplemented with acetate (10 mM). (B) Growth of *S. enterica* Δ*pat cobB*$^{+}$ (squares) and *S. enterica* Δ*pat* Δ*cobB* (circles) strains producing GNAT^{D185R} 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.

for the reaction mixture containing *Se*Pat^{WT} + *Se*Acs^{WT} was due to the rapid enzymatic acetylation of *Se*Acs^{WT} by *Se*Pat^{WT}. Control reactions lacking *Se*Pat^{WT} indicated that radiolabel was not associated with *Se*Acs^{WT}, even after 90 min of incubation with [1-^{14}C]-acetyl-CoA (Fig. 6.5). In contrast, neither *Sl*PatA^{WT} nor *Se*Pat^{WT} efficiently acetylated the *Se*Acs^{R606E} and *Se*Acs^{R613D} variants. As shown in figure 6.4C, *Se*Acs^{WT} was almost completely inactivated by either *Sl*PatA^{WT} or *Se*Pat^{WT} (>96% loss of activity), while *Se*Acs^{R606E} and *Se*Acs^{R613D} variants displayed significant activity after incubation with *Sl*PatA^{WT} (>40% activity retained) or *Se*Pat^{WT} (>80% activity retained).

Interaction model for GNAT and protein substrate. Members of the GNAT acetyltransferase family have been crystallized with peptide substrates to visualize local interactions with peptide residues surrounding the target lysine. Limited information is available about the constellation of interactions that occur between the GNAT and a full-length protein acetylation target. GNAT and full-length *Se*Acs^{WT} interactions were modeled using the ClusPro 2.0 server to visualize the possible interactions during the protein acetylation reaction (9, 10, 28, 29). Because charged residues were demonstrated to be important for the GNAT-*Se*Acs^{WT} interactions, the interaction models prioritizing electrostatic interactions were analyzed further. Crystal structures of the yeast Gcn5 protein (yGcn5p) bound to a peptide substrate demonstrated that the carboxyl group of the putative catalytic E123 was approximately 8 Å away from the epsilon-amino group of the target lysine (40). Thus, the 10 best-fit models were evaluated for conformations that placed the *Se*Acs^{WT} K609 residue within 15 Å from the GNAT E123 active site residue to account for the absence of the electron density of K609 in the *Se*Acs CTD^{WT} structure (described above). The best-fit interaction model is shown in Figure 6.6A.

GNAT-substrate crosslinking reveals a selective orientation for enzyme-substrate interaction. To test the ClusPro interaction model, GNAT and *Se*Acs CTD^{WT} were crosslinked in 9 orientations with crosslinkers of distinct lengths. To generate crosslinked complexes, residues near the prediction interaction surfaces of GNAT and *Se*Acs CTD^{WT} were identified, and a single-cysteine variant was constructed for each chosen residue (Figure 6.6A-B, GNAT residues S73, A110, A164; *Se*Acs CTD^{WT} residues A538, H567, D600). Purified single-Cys variants were crosslinked using sulfhydryl-specific

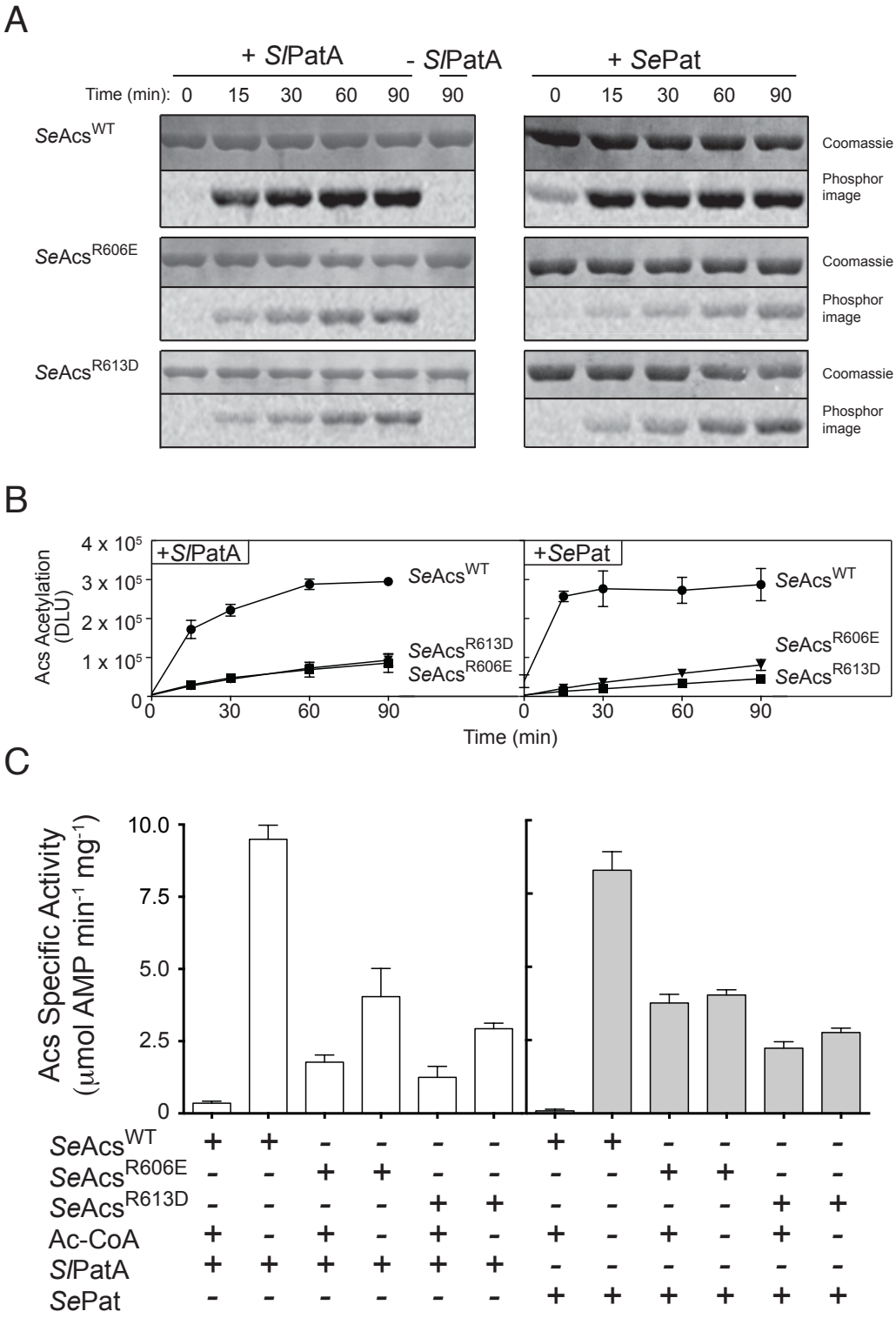


Figure 6.4. *Se*Acs^{R606E} and *Se*Acs^{R613D} are resistant to acetylation by protein acetyltransferases. (A) *Se*Acs^{WT}, *Se*Acs^{R606E}, and *Se*Acs^{R613D} were incubated with *Sl*PatA or *Se*Pat in a 3:1 molar ratio

(*Se*Acs:*Sl*PatA or *Se*Pat) the presence of [1-^{14}C]-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. Acetylation was visualized by phosphor imaging. Reaction controls lacking *Se*Pat are included in Figure 6.5. Gels and phosphor images were cropped to the *Se*Acs bands. Phosphor signal associated with each band was quantified as described in the experimental procedures (B). (C) *Se*Acs, *Se*Acs^{R606E}, and *Se*Acs^{R613D} were incubated with *Sl*PatA (white bars) or *Se*Pat (gray bars) at the ratio described above in the presence or absence of acetyl-CoA. After 90 minutes, the *Se*Acs activity was measured in an NADH-consumption assay. All data points are mean ± standard deviation.

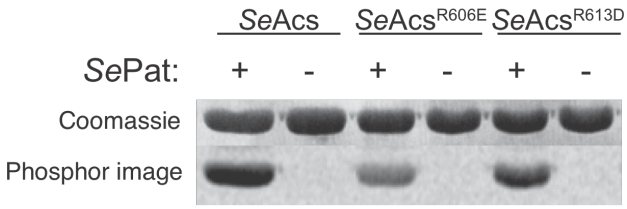


Figure 6.5. *Se*Acs^{R606E} and *Se*Acs^{R613D} are acetylated by *Se*Pat^{WT}. (A) *Se*Acs, *Se*Acs^{R606E}, and *Se*Acs^{R613D} were with incubated with *Se*Pat in a 3:1 molar ratio (*Se*Acs:*Se*Pat) the presence of [1-^{14}C]-Ac-CoA. Samples were quenched after 90 minutes and separated by SDS-PAGE and stained with Coomassie Blue to visualize proteins. Gel and phosphor image were cropped to the *Se*Acs^{WT} bands.

crosslinkers with reported lengths of 2.05 Å to 19.9 Å (Fig. 6.6C). Crosslinking in all 9 configurations was successful with at least 4 different length crosslinkers (Fig. 6.6E).

The ability of the GNAT to acetylate *Se*Acs^{WT} residue K609 in each crosslinked species was used as a measure of the biological relevance of the orientation caused by the crosslinker length constraints. In the absence of a crosslink, GNAT acetylation of *Se*Acs CTD^{WT} was efficient only when GNAT is in molar excess of the *Se*Acs CTD^{WT} (Fig. 6.6D). As seen in figure 6.6E, GNAT acetylates *Se*Acs CTD^{WT} in all of the crosslink orientations when th spacer length is ≥8 Å except in the GNAT^{A110C}-*Se*Acs CTD^{A538C} crosslinked complex. Acetylation was relatively weaker when a 20-Å linker (DPDPB) joined GNAT and *Se*Acs CTD^{WT} relative to linkers ranging in size from 8-16.1Å indicating that increased spacer length did not necessarily promote interaction and acetylation. Notably, only the GNAT^{S73C}-*Se*Acs CTD^{H567C} crosslinked complex resulted in detectable acetylation with a direct disulfide bond between the cysteine residues of each protein. As a point of reference, Ser73 of GNAT and His567 of *Se*Acs CTD^{WT} were positioned near one another in the interactions model (Fig. 6.6A), and this GNAT-*Se*Acs CTD^{WT} complex was predicted to mimic the ClusPro interaction model most closely.

DISCUSSION

Characterization of GNAT interactions with a protein substrate. Here we characterized interactions between a GNAT protein acetytransferase domain and a protein substrate, advancing our understanding of how GNATs and their acceptor substrates productively interact to modify the target lysine. *In vitro* and *in vivo* data support the conclusion that a negatively charged residue in the GNAT active site cleft and two positively charged residues near the substrate lysine in the substrate protein are critical for acetylation.

Computer-assisted electrostatic interaction model-building allowed us to construct a direct disulfide crosslink between the GNAT and protein substrate that retained high levels of catalytic activity. The data support an interaction model that locks the GNAT and its substrate in a conformation that likely

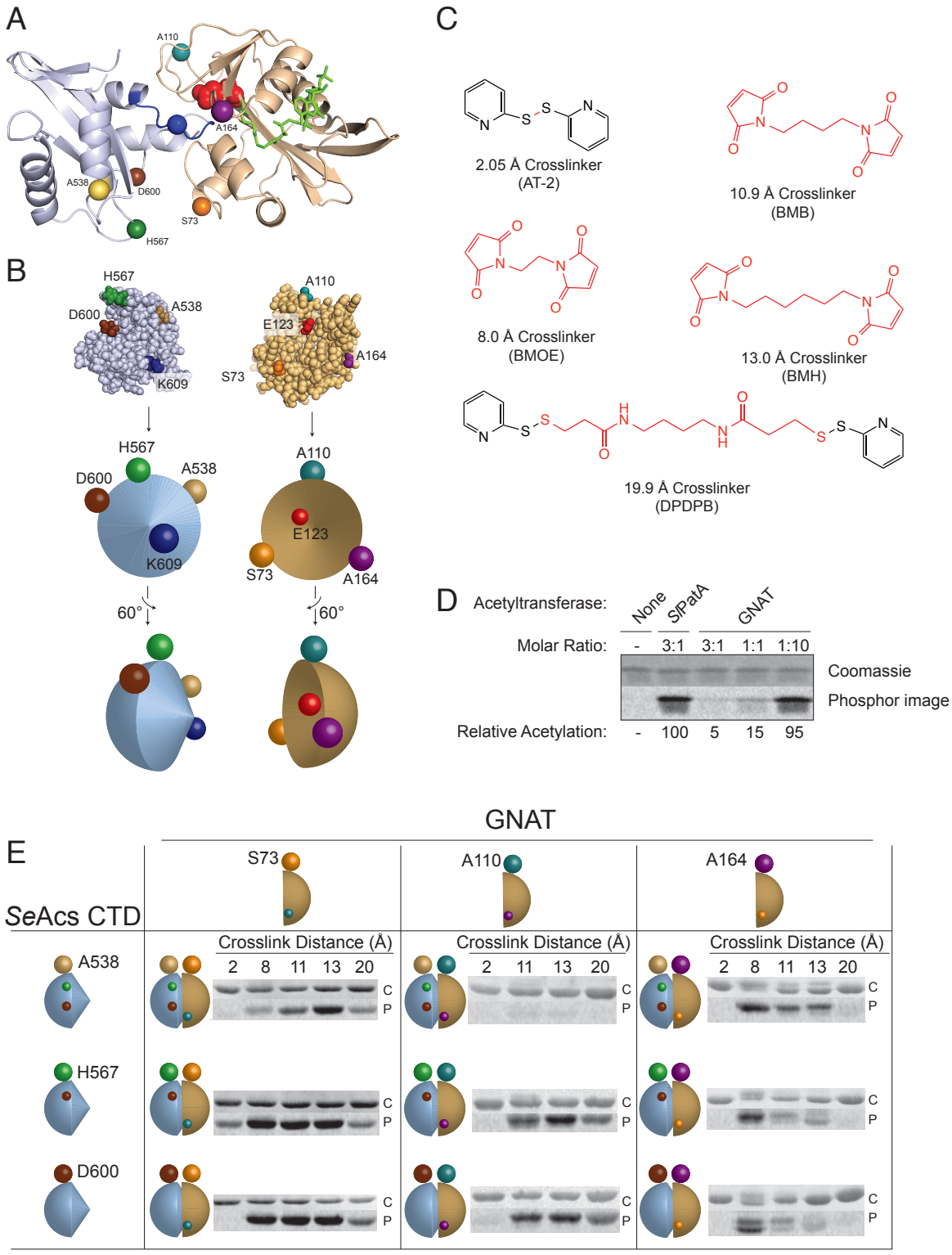


Figure 6.6. Direct crosslinking of GNAT^{S73C} and *Se*Acs CTD^{H567C} with a disulfide bond results in a reactive complex. (A) Modeling of the GNAT-*Se*Acs^{WT} interface. GNAT active site E123 (red) and *Se*Acs^{WT} target K609 (blue) are shown as spheres. The *Se*Acs^{WT} A10 loop is shown in dark blue. Residues

chosen for the crosslinking analysis are highlighted. (B) *Se*Acs CTD^{WT} (light blue) and GNAT (beige) shown as spheres with GNAT^{E123}, *Se*Acs CTD^{K609}, and crosslinking residues colored. Each structure was simplified into a schematic to demonstrate the relative location of the crosslinking locations on the predicted interaction surfaces. (C) Chemical structures of 2.05 Å crosslink (AT-2), 8.0 Å crosslink (BMOE), 10.9 Å crosslink (BMB), 13.0 Å crosslink (BMH), and 19.9 Å crosslink (DPDPB). (D) *Se*Acs CTD was incubated with full-length *Sl*PatA^{WT} at a 3:1 molar ratio (*Se*Acs CTD^{WT}:*Sl*PatA^{WT}) or with GNAT at 3:1, 1:1, and 1:10 molar ratio (*Se*Acs CTD^{WT}:GNAT) in the presence of [1-^{14}C]-Ac-CoA. Samples were quenched after 60 minu, separated by SDS-PAGE and stained with Coomassie Blue to visualize proteins. Acetylation was visualized by phosphor imaging. Gel and phosphorimage were cropped to the *Se*Acs CTD^{WT} band. Acetylation was quantified relative to *Sl*PatA^{WT} + *Se*Acs CTD^{WT} signal and is reported as the mean of three independent experiments. S.D. for all values was ≤18% of the mean value. (E) Crosslinked *Se*Acs CTD^{WT} and GNAT complexes were normalized to 2 ug per reaction based on densitometry analysis of Coomassie-stained gels. Complexes were incubated in the presence of [1-^{14}C]-Ac-CoA for 60 min and visualized as described above except using non-denaturing SDS-PAGE. Gels and phosphor images were cropped to the *Se*Acs CTD^{WT} and GNAT complex bands.

resembles a native interaction orientation between the two proteins *in vivo*. Below, we discuss our data in support of the interaction model presented in figure 6.6A.

The putative GNAT active site is located within a negatively charged cleft. As seen in figure 6.2A-C, the putative catalytic E123 of GNAT is found within a cleft containing negatively charged residues, one of which (D185) was shown to be important for interaction with the protein substrate (Fig. 6.2G). GNAT residue D185 is approximately 4.4 Å away from *Se*Acs residue R612 in the ClusPro interaction model of GNAT and *Se*Acs (Fig. 6.7A). We conclude that the observed poor interactions between the GNAT^{D185R} variant and *Se*Acs^{WT} are explained by strong charge repulsions.

In the GNAT structure residue D185 occurs in a negatively charged loop along with residue Glu184, which is conserved in the structure of the GNAT domain of protein acetyltransferases like *Mt*PatA (E294, D295), a GNAT known to regulate AMP-forming CoA ligases (Fig. 6.8A). Analogous negatively charged loops are predicted in structure models of the *Se*Pat GNAT domain (E172, E173) and the *Rp*KatA GNAT (E156) generated using the Phyre 2.0 software program (Fig. 6.8B-C) (26). In the crystal structure of the *Sulfolobus solfataricus* protein acetyltransferase Pat (*Ss*Pat), the residues forming this loop (E136, D137, D138) are excluded from the active site cleft (Fig. 6.8D). *Ss*Pat^{WT} was crystallized after limited proteolysis and the absence of amino acids (specifically aa42-52 that link helix α2 to strand β2) may have distorted the active site cleft. *Ss*Pat^{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 *Ss*Pat^{WT} active site cleft may represent distinct substrate specificity. Together, the GNAT structures and structural model reveal a conserved negatively charged loop that may mediate interactions of GNATs with protein substrates.

Positive charges surrounding target lysine of the protein substrate mediate interactions with the GNAT. Residue K609 of *Se*Acs^{WT} is located within the structurally conserved A10 loop found in AMP-forming CoA ligases (22). In the crystal structure of *Se*Acs^{WT}, the A10 loop contains positively charged residue equivalent to residue R612. Residues R606 and R613 flank the A10 loop (Fig. 6.2E) and are conserved in representative Acs homologues from the three domains of life (46). The relevance of

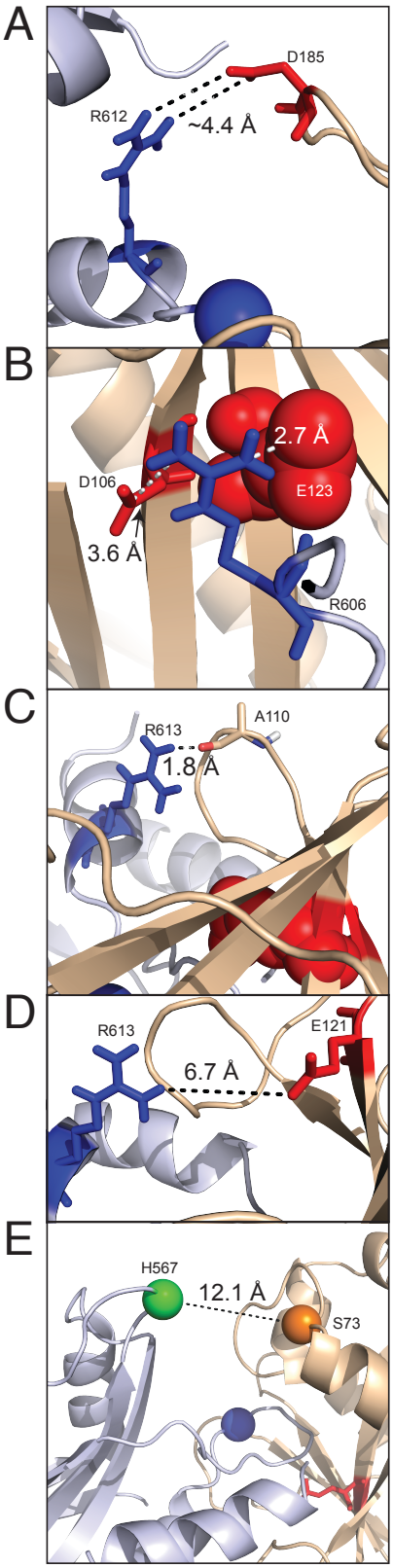


Figure 6.7. Interaction model of GNAT and protein substrate supported by interaction data. (A) D185 side chain of GNAT (beige) is predicted to approach R612 side chain of *Se*Acs CTD^{WT} (light blue).

For conservation of this negatively charged loop, see Figure 6.8. (B) The *Se*Acs CTD R606 guanidinium moiety is predicted to interact with the side chains of GNAT residues D106 and E123, the putative active site residue. (C) *Se*Acs CTD residue R613 is predicted to interact with the backbone carbonyl group of GNAT A110. (D) The closest potential intermolecular interaction partner for E121 is *Se*Acs CTD R613. (E) *Se*Acs CTD H567 and GNAT residue S73 are positioned near one another in the ClusPro model. In all panels, the backbone atoms of the *Se*Acs K609 acetylation site residue are represented as a blue sphere and the GNAT putative active site glutamate is represented by red spheres.

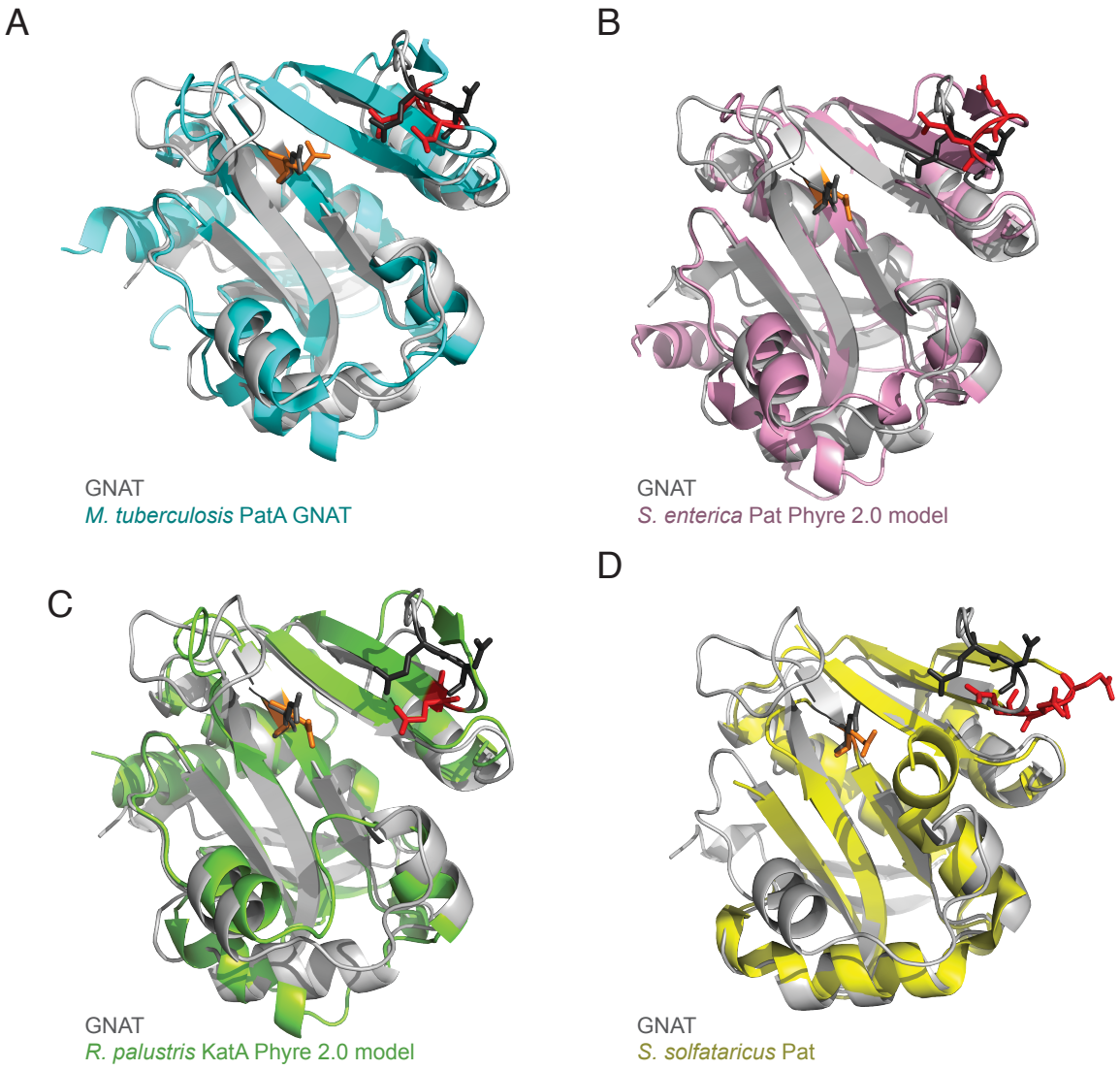


Figure 6.8. Conservation of a negatively charged loop near the GNAT active site. *Sl*PatA GNAT is shown aligned with (A) the crystal structure of *M. tuberculosis* PatA (PDB accession code 4AVB), (B) a structure model of *S. enterica* Pat GNAT, (C) a structure model of *R. palustris* KatA, and (D) the structure of *S. solfataricus* Pat. The negatively charged loop residues of *Sl*PatA GNAT and aligned structures are shown in black and red, respectively. For reference, the putative active site glutamate of *Sl*PatA GNAT and the aligned structures are represented as sticks in light gray or orange, respectively. Structure models were generated using the Phyre 2.0 server. Alignments were generated in PyMol (Delano, 2002).

residues R606 and R613 explains the clearly weakened interaction of *Se*Acs^{WT} with GNAT and substantially reduced acetylation of K609 (Figs. 6.2G-H, 6.4A-C).

In the ClusPro interaction model, R606 is predicted to be near both D106 (3.6 Å) and E123 (2.7 Å) (Fig. 6.7B). We propose that these electrostatic interactions orient residue E123 for catalysis. In the wild-type enzyme, the guanidinium group nitrogen of R613 is placed approximately 1.8 Å away from the carbonyl group of alanine in the polypeptide chain. Consistent with this model was the result obtained after reversing the charge at that position in the *Se*Acs^{R613D} variant, which eliminated this contact (Fig. 6.7C), and the acetylation of *Se*Acs CTD residue K609.

Interactions were detected between *Se*Acs CTD^{R613D} and GNAT^{E121R} in a charge swap experiment. In the ClusPro interaction model (Fig. 6.7D), residues R613 of *Se*Acs CTD^{WT} and E121 of the GNAT are separated by 6.7 Å, suggesting that electrostatic interactions between these residues may not be occurring directly. It is possible that the positive charge introduced by GNAT^{E121R} interacts with other residues in the negatively charged GNAT active site cleft (Fig. 6.2A) to accommodate the negative charge introduced in the substrate *Se*Acs CTD^{R613D}.

In vitro, full-length *Se*Acs^{R606E} and *Se*Acs^{R613D} have decreased, but substantial enzymatic activity (>60% and >30% of wild-type *Se*Acs activity, respectively) suggesting that the positive charge side chains at these locations are important for interactions with the GNAT and for optimal enzymatic activity.

***The large domain of* Sl*PatA is important for substrate recognition by the GNAT domain*.** Single-domain GNAT protein acetyltransferases efficiently acetylate protein substrates (13, 17, 34). Why do Pat homologues have a large accessory domain in addition to the GNAT domain (Fig. 6.1A)? We have learned from work with the *Sl*PatA homologue in *S. enterica* and histone acetyltransferase (HAT1) in eukaryotes, that the large domains of protein acetyltransferases are necessary for catalytic activity (37, 49). Thao *et al.* reported that the large domain found in the *S. enterica* Pat (*Se*Pat) homologue was required for oligomerization and activity (50). The large domain may regulate binding of the Ac-CoA molecule or acceptor substrate to the GNAT, and this would account for the inefficient acetylation of *Se*Acs CTD^{WT} seen with the GNAT domain alone (Fig. 6.6D). Our data suggest that one role of the large

domain in Pat homologues is to somehow enhance the interactions of the GNAT domain with its protein substrate, since crosslinking of the GNAT domain to its substrate leads to efficient substrate acetylation. The large domain may enhance interactions by directly contacting the acceptor substrate or inducing structural changes in the GNAT domain upon Ac-CoA binding to the large domain.

GNAT-protein substrate complexes reveal productive catalysis in specific orientations. Results from crosslinking experiments revealed important, productive orientations of the GNAT and *Se*Acs CTD^{WT} that resulted in substrate acetylation. Of the 36 possible complex orientations, only the GNAT^{S73C}-*Se*Acs CTD^{H567C} complex was catalytically active when a disulfide bond (2.05 Å) directly linked the cysteine residues introduced into the GNAT and *Se*Acs CTD^{WT} proteins. Such a direct, restrictive crosslink limits the possible number of GNAT and *Se*Acs^{WT} interactions, but catalysis occurred with high efficiency (Fig. 5E). Thus, we hypothesize that the relative domain orientations resulting in the GNAT^{S73C}-*Se*Acs CTD ^{H567C} complex represents a strong candidate for the biologically relevant orientation of GNAT with the *Se*Acs CTD^{WT} during interaction and acetylation. Consistent with this hypothesis, the GNAT^{S73C}-*Se*Acs CTD^{H567C} complex most closely mimics the interaction model generated using the ClusPro software (figures 6.6A, 6.7E).

Eight of the nine orientations resulted in catalytically productive interactions when domains were joined with a linker ≥8 Å suggesting that linkers providing potential distance and flexibility between the domains generally resulted in detectable interactionsmol. Acetylation was undetectable in the GNAT^{A110C}-*Se*Acs CTD^{A538C} crosslinked complexes suggesting that this crosslinked orientation disfavors catalysis when tethered at any distance. The 20-Å linker (DPDPB) resulted in relatively less acetyation than the 8-16.1 Å linkers in most cases. This may be due to the distance introduced by the crosslinker disfavoring protein interaction (20) or to the difference in functional groups in the sulfhydryl-reactive crosslinker (DPDPB) compared to the maleimide crosslinkers (BMOE, BMB, and BMH).

REFERENCES

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

2. **Berndsen, C. E., and J. M. Denu.** 2008. Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol*. **18:**682-689.

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

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

5. **Brownell, J. E., J. Zhou, T. Ranalli, R. Kobayashi, D. G. Edmondson, S. Y. Roth, and C. D. Allis.** 1996. *Tetrahymena* histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84:**843-851.

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

7. **CCP4.** 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50:**760-763.

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

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

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

11. **Cowtan, K., and P. Main.** 1998. Miscellaneous algorithms for density modification. *Acta Crystallographica* **D54:**487-493.

12. **Crosby, H. A., E. K. Heiniger, C. S. Harwood, and J. C. Escalante-Semerena.** 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.

13. **Crosby, H. A., D. A. Pelletier, G. B. Hurst, and J. C. Escalante-Semerena.** 2012. System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J. Biol. Chem*. **287:**15590-15601.

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

15. **Elion, E. A., P. Marina, and L. Yu.** 2007. Constructing recombinant DNA molecules by PCR, p. Unit 3.17.1-3.17.12. *In* F. M. Ausubel, R. E. R. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, a. Smith, and K. Struhl (ed.), *Current protocols in molecular biology*. Greene Publishing Associates & Wiley Interscience, New York, N.Y.

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

17. **Gardner, J. G., F. J. Grundy, T. M. Henkin, and J. C. Escalante-Semerena.** 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD(+) involvement in *Bacillus subtilis*. *J. Bacteriol*. **188:**5460-5468.

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

19. **Glozak, M. A., N. Sengupta, X. Zhang, and E. Seto.** 2005. Acetylation and deacetylation of non-histone proteins. *Gene* **363:**15-23.

20. **Green, N. S., E. Reisler, and K. N. Houk.** 2001. Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. *Protein Sci.* **10:**1293-1304.

21. **Griffith, K. L., and R. E. Wolf, 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.

22. **Gulick, A. M.** 2009. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol*. **4:**811-827.

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

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

25. **Hubbert, C., A. Guardiola, R. Shao, Y. Kawaguchi, A. Ito, A. Nixon, M. Yoshida, X. F. Wang, and T. P. Yao.** 2002. HDAC6 is a microtubule-associated deacetylase. *Nature* **417:**455-458.

26. **Kelley, L. A., and M. J. Sternberg.** 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4:**363-371.

27. **Kouzarides, T.** 2007. Chromatin modifications and their function. *Cell* **128:**693-705.

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

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

30. **Kuo, Y. M., and A. J. Andrews.** 2013. Quantitating the specificity and selectivity of Gcn5-mediated acetylation of histone H3. *PLoS One* **8:**e54896.

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

32. **Lee, H. J., P. T. Lang, S. M. Fortune, C. M. Sassetti, and T. Alber.** 2012. Cyclic AMP regulation of protein lysine acetylation in *Mycobacterium tuberculosis*. *Nat. Struct. Mol. Biol.*

33. **Liang, W., A. Malhotra, and M. P. Deutscher.** 2011. Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. *Mol. Cell* **44:**160-166.

34. **Marsh, V. L., S. Y. Peak-Chew, and S. D. Bell.** 2005. Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba. *J. Biol. Chem*. **280:**21122-21228.

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

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

37. **Polesskaya, A., and A. Harel-Bellan.** 2001. Acetylation of MyoD by p300 requires more than its histone acetyltransferase domain. *J. Biol. Chem.* **276:**44502-44503.

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

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

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

41. **Sasse, J.** 1991. Detection of proteins, p. 10.6.1-10.6.8. *In* F. A. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), *Current Protocols in Molecular Biology,* vol. 1. Wiley Interscience, New York.

42. **Sheldrick, G. M.** 2008. A short history of SHELX. *Acta Crystallogr.* A. **64:**112-122.

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

44. **Soppa, J.** 2010. Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea* **2010:**pii: 820681.

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

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

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

48. **Thao, S., C. S. Chen, H. Zhu, and J. C. Escalante-Semerena.** 2010. N(epsilon)-Lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. *PLoS ONE* **5:**e15123.

49. **Thao, S., and J. C. Escalante-Semerena.** 2012. A positive selection approach identifies residues important for folding of Salmonella enterica Pat, an N(epsilon)-lysine acetyltransferase that regulates central metabolism enzymes. *Res. Microbiol.* **163:**427-435.

50. **Thao, S., and J. C. Escalante-Semerena.** 2011. Biochemical and thermodynamic analyses of *Salmonella enterica* Pat, a multidomain, multimeric N(epsilon)-lysine acetyltransferase involved in carbon and energy metabolism. *mBio* **2:**e00216-11.

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

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

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

54. **Zhang, X., and H. Bremer.** 1995. Control of the *Escherichia coli* rrnB P1 promoter strength by ppGpp. *J. Biol. Chem.* **270:**11181-11189.

CHAPTER 7

CONCLUSIONS AND FUTURE DIRECTIONS

SUMMARY AND CONCLUSIONS

Overview. Roles of reversible lysine acetylation (RLA) systems have been described in *S. enterica* (24, 25, 27, 28), *E. coli* (1, 15, 16), *B. subtilis* (9-11), *R. palustris (6-8),* and mycobacteria (14, 20, 21, 31). This work expanded the RLA network of proteins by identifying a new isoform of the sirtuin deacetylase in *S. enterica* that is conserved in γ-Proteobacteria. Members of the RLA system in the actinomycete *S. lividans* were also identified. *Sl*PatA is a protein acetyltransferase with unique domain organization. *Sl*PatA regulates members of the AMP-forming CoA ligase (ACL) family of enzymes in *S. lividans*. However, not all *S. lividans* ACL enzymes are efficiently acetylated by *Sl*PatA. The results in Chapter 6 define *Sl*PatA interactions with protein substrates at the atomic level. Together this work expands the RLA system in bacteria and defines constrains of RLA modifying enzymes and their substrates.

Multiple deacetylase isoforms are conserved in γ-proteobacteria. The effect RLA on cellular metabolism is most extensively studied in *S. enterica* (2, 12, 13, 24-26, 28-30). Chapter 2 describes an additional regulator in this organism. *S. enterica* synthesizes two isoforms of the CobB deacetylase from in-frame start codons in one open reading frame. The previously unidentified longer CobB isoform ($CobB_L$) encodes an arginine-rich, 37-amino acids extension. Both isoforms are active *in vivo* and *in vitro*. The long isoform is conserved in γ-Proteobacteria. The plant pathogen *Erwinia amylovora* only encodes the long form, suggesting that this isoform is sufficient for deacetylation in this organism. The physiological role of the $CobB_L$ isoform is undetermined.

Sl*PatA regulates ACL enzymes in* S. lividans. Lysine acetylation regulates members of the ACL family of enzymes in diverse bacteria. Work in Chapters 3 and 4 demonstrated that *S. lividans* encodes a

protein acetyltransferase, *Sl*PatA, with unique domain organization that regulates five ACL enzymes *in vivo*, including acetoacetyl-CoA synthetase (AacS). AacS is found in all domains of life, and acetylation of AacS may be a conserved regulatory mechanism. *Sl*PatA was also found to controls acetylation of at least four other ACL enzymes in *S. lividans*. Acetylation of ACL enzymes in bacteria is thought to control ATP consumption in the ACL reaction (2) or to maintain CoA homeostasis by sensing increases in Ac-CoA levels (7).

Engineering* Sl*Acs to be an* Sl*PatA substrate reveals* Sl*PatA specificity. *Sl*PatA discriminated between ACL enzymes *in vitro*. *Sl*PatA does not efficiently acetylate *S. lividans* acetyl-CoA synthetase (*Sl*Acs) *in vitro*. However, *S. enterica* Acs (*Se*Acs) is a strong substrate for *Sl*PatA. In Chapter 5, amino acid sequences from *Se*Acs were swapped into *Sl*Acs to generate a *Sl*Acs variant that is regulated by *Sl*PatA acetylation. Regions both *N*-terminal and *C*-terminal to the target lysine in *Sl*Acs and *Se*Acs were important for *Sl*PatA recognition. This work, along with previous work in *R. palustris* (8), demonstrated the effectiveness of using protein engineering to study RLA substrate specificity and to construct proteins that are subject to native RLA systems.

Physiologically relevant GNAT-substrate acetylation complex. GNATs interact with multiple residues in the protein substrate surrounding the target lysine (22). Chapter 6 describes a broader constellation of interactions that occur between the GNAT and the protein substrate. The structure of the *Sl*PatA GNAT revealed a negatively-charge active site cleft. Efficient interaction between *Sl*PatA GNAT and the protein substrate requires at least one of these negatively charged residues of the *Sl*PatA GNAT and two positively charged residues near the target lysine in the protein substrate. A computer-generated docking model of the *Sl*PatA GNAT with a protein substrate was generated to visualize potential protein-protein interactions. We validated this model by crosslinking *Sl*PatA GNAT and the protein substrate in this conformation and demonstrating catalytic interaction. To my knowledge, this is the first description of a physiologically relevant protein acetylation complex.

FUTURE DIRECTIONS

Physiological role of the positively-charged* N-*terminus of* S. enterica *sirtuin. *S. enterica* sirtuin CobB does not require the *N*-terminal extension found on the $CobB_L$ isoform for deacetylation of *S. enterica* Acs *in vitro* or *in vivo*. Because of the positively charged nature of the $CobB_L$ *N*-terminal extension, I hypothesize that it binds a negatively charged molecule in the cell. Immunoprecipitation of CobB from *S. enterica* did not yield any protein binding partners, but crosslinking may be required for detection of interaction partners. Alternatively, $CobB_L$ may be interacting with other cellular molecules including phospholipids, nucleic acids, or small molecules.

***Regulation of* Sl*PatA in* S. lividans.** Little is known about how acetyltransferases respond to intracellular levels of the substrate Ac-CoA. The *S. enterica* Pat protein acetylatransferase (*Se*Pat) binds Ac-CoA cooperatively, and is likely sensitive to the Ac-CoA levels of the cell (29). Homology of *Sl*PatA to *Se*Pat suggests the same feature may be true of *Sl*PatA. All known targets of *Sl*PatA acetylation are ACL enzymes that produce acyl-CoA species. As seen is suggested in *R. palustris* (7), the ACL enzymes regulated by acetylation produce acyl-CoA species that are likely broken down to produce Ac-CoA. *Sl*PatA may be sensing the downstream acetyl-CoA product of these enzymes and regulating their activity to maintain Ac-CoA homeostasis. *Se*Pat acetylation of ACL enzymes also prevents ACL-mediated depletion of the cellular ATP pool (2). Thus, *Sl*PatA may detect Ac-CoA as a read-out of ACL activity.

*Sl*PatA encodes a proline rich region that includes a degenerate collagen repeat G-P-S that is cleaved by a collagenase *in vitro*. Collagen like repeats form fibrillar structures (32) and mediate multimerization (3) in other bacterial proteins. *Sl*PatA does not multimerize *in vitro* (A. C. Tucker and J. C. Escalante-Semerena, unpublished work). Multimerization of *Sl*PatA may require features absent in a pure system. The collagen-like motif may be used to interact with *S. lividans* proteins in the cell that could localize, enhance, or inhibit *Sl*PatA activity.

Effect of acetylation on Sl*PatA *substrates. One major question in the field of protein acetylation is how the bacteria link the relative proportion of acetylated substrates to intracellular signals. In chapter 3, we demonstrate that H_6-tagged acetoacetyl-CoA synthetase (H_6-AacS) is acetylated by *Sl*PatA in wild-

type *S. lividans* when growing on acetoacetate as a carbon source. Thus, the RLA system of *S. lividans* including *Sl*PatA and the *S. lividans* deacetylases maintained a sufficient active population of ectopically-produced H_6-AacS to support growth on acetoacetate without depleting cellular ATP pools (described above). Characterizing the timing and efficiency of *Sl*PatA substrate acetylation may reveal the signals and hierarchy of metabolic control for enzymes in the RLA.

Role of* S. lividans *deacetylases. The role of the *S. lividans* deacetylases is more complex than has been described for *S. enterica*. A deacetylase-deficient strain of *S. enterica* grows poorly on low levels of acetate due to the acetylation and inactivation of Acs by Pat (25). Growth of a *S. lividans* strain deficient of all known deacetylases is indistinguishable from wild-type when cells are growing on acetoacetate. This is unexpected because AacS (described above) has cannot be deacetylated and reactivated in a deacetylase-deficient *S. lividans* strain. Presumable *Sl*PatA is present to acetylate AacS because we detected acetylation of H_6-AacS under identical conditions in a wild-type *S. lividans* strain (described above). *S. lividans* may encode a member of a previously-uncharacterized deacetylase family. Alternatively, *Sl*PatA activity may be low when *S. lividans* grows on acetoacetate, and this level of activity does not inactivate the entire AacS population.

S. lividans encodes two sirtuin deacetylases and one Zn^{2+}-dependent deacetylase, and our attempts to purify active enzymes were unsuccessful. The closely-related actinomycete *Streptomyces coelicolor* encodes homologs to each *S. lividans* deacetylase, and one of the sirtuin homologs, CobB1, has documented activity *in vitro* (18). It is unclear why *S. lividans* encodes multiple deacetylases. *R. palustris* and *B. subtilis* each encode a sirtuin and a Zn^{2+}-dependent deacetylase, and both deacetylate ACL family enzymes (6, 10, 11). Presumably sirtuin activity would be regulated by NAD^+ availability in the cell while activity of the Zn^{2+}-dependent deacetylase would be under other mechanisms of control (e.g. transcriptional or translational). These deacetylases may regulate protein deacetylation in different physiological conditions or may act in concert to regulate different substrates, though these two possibilities are not mutually exclusive.

What role does protein acetylation play in* Streptomyces *physiology? *Streptomyces* species produce diverse natural products (4, 5, 17, 23). It is possible that RLA affects physiology of this organism by altering carbon flow in primary and secondary metabolism. We have identified 5 enzymes of the ACL family that are regulated by *Sl*PatA. Understanding the function of acetylation of these and other targets this would enhance current approaches to engineer *Streptomyces* species to produce natural products (19).

Does* Sl*PatA regulate Acs in* S. lividans*? Work in Chapter 5 describes the construction of a *Sl*Acs variant that is regulated by *Sl*PatA outside of the natural host. It is unclear why *Sl*PatA does not recognize SlAcs since *Sl*PatA efficiently acetylates the closely related homolog *S. enterica* Acs. I pursued the *Sl*Acs structure to visualize structural differences between this enzyme and *Se*Acs, for which a structure is known (PDB: 1PG4). However, *Sl*Acs did not yield crystals despite attempts with multiple *Sl*Acs variants, *Sl*Acs substrate/product combinations, and *Sl*Acs chemical derivatizations.

If *Sl*PatA acetylates *Sl*Acs in *S. lividans*, we should be able to detect the modification on *Sl*Acs isolated from an *S. lividans* wildtype strain but not *Sl*Acs isolated from a *S. lividans* Δ*patA* strain. Initial attempts to isolate *Sl*Acs from a *S. lividans* Δ*patA* strain were unsuccessful. If *Sl*PatA does regulated *Sl*Acs *in vivo*, *Sl*PatA may require an accessory factor to recognize *Sl*Acs. This would render *Sl*PatA regulation of Acs more complex than is seen in other bacterial systems. Alternatively, another *S. lividans* GNAT may be responsible for acetylation of *Sl*Acs.

Visualizing the protein acetylation complex. This work established and supported the first model of interactions between a GNAT and a protein substrate. *S. lividans* PatA and *R. palustris* Pat (7) can distinguish between closely related ACL substrates, suggesting that sequence and structural changes in ACL enzymes block these interactions. *E. coli* Pka acetyltransferase recognizes structurally unrelated substrates including ACL enzymes and to ribonuclease R (1, 15, 16). Does the GNAT domain of *E. coli* Pka recognize these structurally unrelated substrates in similar manners? Is the *E. coli* Pka active site permissive to specific structural differences? Further structural work to characterize the architectures of GNAT active sites, their protein substrate, and closely related non-substrate proteins will define the structural constraints of RLA systems.

REFERENCES

1. **Castaño-Cerezo, S., V. Bernal, J. Blanco-Catala, J. L. Iborra, and M. Canovas.** 2011. cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in *Escherichia coli*. *Mol. Microbiol*. **82:**1110-1128.

2. **Chan, C. H., J. Garrity, H. A. Crosby, and J. C. Escalante-Semerena.** 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.

3. **Charalambous, B. M., J. N. Keen, and M. J. McPherson.** 1988. Collagen-like sequences stabilize homotrimers of a bacterial hydrolase. *EMBO J*. **7:**2903-2909.

4. **Chater, K. F.** 2006. *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361:**761-768.

5. **Courtois, S., C. M. Cappellano, M. Ball, F. X. Francou, P. Normand, G. Helynck, A. Martinez, S. J. Kolvek, J. Hopke, M. S. Osburne, P. R. August, R. Nalin, M. Guerineau, P. Jeannin, P. Simonet, and J. L. Pernodet.** 2003. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol.* **69:**49-55.

6. **Crosby, H. A., E. K. Heiniger, C. S. Harwood, and J. C. Escalante-Semerena.** 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.

7. **Crosby, H. A., D. A. Pelletier, G. B. Hurst, and J. C. Escalante-Semerena.** 2012. System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J. Biol. Chem.* **287:**15590-15601.

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

9. **Gardner, J. G., and J. C. Escalante-Semerena.** 2008. Biochemical and mutational analyses of AcuA, the acetyltransferase enzyme that controls the activity of the acetyl coenzyme a synthetase (AcsA) in *Bacillus subtilis*. *J. Bacteriol.* **190:**5132-5136.

10. **Gardner, J. G., and J. C. Escalante-Semerena.** 2009. In *Bacillus subtilis*, the sirtuin protein deacetylase encoded by the *srtN* gene (formerly *yhdZ*), and functions encoded by the *acuABC* genes control the activity of acetyl-CoA synthetase. *J. Bacteriol.* **191:**1749-1755.

11. **Gardner, J. G., F. J. Grundy, T. M. Henkin, and J. C. Escalante-Semerena.** 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD(+) involvement in *Bacillus subtilis*. *J. Bacteriol.* **188:**5460-5468.

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

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

14. **Lee, H. J., P. T. Lang, S. M. Fortune, C. M. Sassetti, and T. Alber.** 2012. Cyclic AMP regulation of protein lysine acetylation in *Mycobacterium tuberculosis*. *Nat. Struct. Mol. Biol.*

15. **Liang, W., and M. P. Deutscher.** 2012. Post-translational modification of RNase R is regulated by stress-dependent reduction in the acetylating enzyme Pka (YfiQ). *RNA* **18:**37-41.

16. **Liang, W., A. Malhotra, and M. P. Deutscher.** 2011. Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. *Mol. Cell* **44:**160-166.

17. **McMahon, M. D., C. Guan, J. Handelsman, and M. G. Thomas.** 2012. Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. *Appl. Environ. Microbiol.* **78:**3622-3629.

18. **Mikulik, K., J. Felsberg, E. Kudrnacova, S. Bezouskova, D. Setinova, E. Stodulkova, J. Zidkova, and V. Zidek.** 2012. CobB1 deacetylase activity in *Streptomyces coelicolor*. *Biochem. Cell Biol.* **90:**179-187.

19. **Mitchell, W.** Natural products from synthetic biology. *Curr. Opin. Chem. Biol.* **15:**505-515.

20. **Nambi, S., S. Badireddy, S. S. Visweswariah, and G. S. Anand.** 2012. Cyclic AMP-induced conformational changes in mycobacterial protein acetyltransferases. *J. Biol. Chem*. **287:**18115-18129.

21. **Nambi, S., K. Gupta, M. Bhattacharya, P. Ramakrishnan, V. Ravikumar, N. Siddiqui, A. T. Thomas, and S. S. Visweswariah.** 2013. Cyclic AMP-dependent protein lysine acylation in mycobacteria regulates fatty acid and propionate metabolism. *J. Biol. Chem*. **288**.

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

23. **Seow, K. T., G. Meurer, M. Gerlitz, E. Wendt-Pienkowski, C. R. Hutchinson, and J. Davies.** 1997. A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J. Bacteriol.* **179:**7360-7368.

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

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

26. **Starai, V. J., J. G. Gardner, and J. C. Escalante-Semerena.** 2005. Residue Leu-641 of acetyl-CoA synthetase is critical for the acetylation of residue Lys-609 by the Protein acetyltransferase enzyme of *Salmonella enterica*. *J. Biol. Chem.* **280:**26200-26205.

27. **Starai, V. J., H. Takahashi, J. D. Boeke, and J. C. Escalante-Semerena.** 2004. A link between transcription and intermediary metabolism: a role for Sir2 in the control of acetyl-coenzyme A synthetase. *Curr. Opin. Microbiol*. **7:**115-119.

28. **Starai, V. J., H. Takahashi, J. D. Boeke, and J. C. Escalante-Semerena.** 2003. Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae*. *Genetics* **163:**545-555.

29. **Thao, S., and J. C. Escalante-Semerena.** 2011. Biochemical and thermodynamic analyses of *Salmonella enterica* Pat, a multidomain, multimeric N(epsilon)-lysine acetyltransferase involved in carbon and energy metabolism. *mBio* **2:**e00216-11.

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

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

32. **Xu, Y., D. R. Keene, J. M. Bujnicki, M. Hook, and S. Lukomski.** 2002. Streptococcal Scl1 and Scl2 proteins form collagen-like triple helices. *J. Biol. Chem.* **277:**27312-27318.