PROTEIN AND SMALL MOLECULE ACETYLATION IN PROKARYOTES

by

CHELSEY M. VANDRISSE

(Under the Direction of Jorge C. Escalante-Semerena)

ABSTRACT

Acetylation is a chemical modification conserved in all domains of life. Acetylation has been dominantly defined as a post-translational modification of lysyl residues of proteins carried out by Gcn5 *N*-acetyltransferases (GNATs, PF 00583). Within a single organism, a genome can code for 7-70 acetyltransferases, all with sequence similarity that currently makes it impossible to determine substrate specificity. There has recently been an upsurge of acetyltransferases that target amine groups of diverse small molecule substrates, rather than proteins, generating an additional layer of complexity in characterizing putative acetyltransferases.

In most cases, acetylation of proteins alters the enzyme's activity, providing the cell with a rapid mechanism to adapt to external or internal fluctuations. The protein acetylation paradigm in prokaryotes involves the modification of an active site lysyl residue of acetyl-CoA synthetase (Acs). In many prokaryotic species, Acs is acetylated by a single protein acetyltransferase containing a regulatory domain, that depending on the organism, responds to amino acids, cAMP, $NADP^+$, or unknown molecules. The cell has adapted to require Acs acetylation, for without it, cellular homeostasis is disturbed through improper control of ATP synthesis.

Due to the fact that all organisms carry out acetylation, this thesis focuses on determining acetyltransferase substrate preferences in a variety of prokaryotic organisms. Through this work,

acetyltransferases have been characterized in *Salmonella enterica*, *Streptomyces lividans*, and *Rhodopseudomonas palustris*. In *S. enterica*, characterization of a toxin protein acetyltransferase led to insight into how the cell enters persistence during infection. Additionally, using a *S. enterica* model for phosphinothricin detoxification, acetyltransferases from *Deinococcous radiodurans*, *Geobacillus kaustophilus*, *Bacillus subtilis, Burkholderia xenovorans, Staphylococcus aureus, Acinetobacter baylyi,* and *Escherichia coli* were characterized as small molecule acetyltransferases that detoxified a myriad of methionine derivatives. In *S. lividans*, Acs acetylation was found to be unique in that it required serine acetylation within the active site. Lastly, in *R. palustris*, small molecule acetylation of aminobenzoates was essential to benzoate catabolism and synthesis of reaction center proteins. Although the projects are diverse, each answers unique questions about protein and small molecule acetylation in prokaryotic physiology.

INDEX WORDS: acetylation, bacteria, persistence, benzoate, photosynthesis, *Salmonella*, *Streptomyces, Rhodopseudomonas*

PROTEIN AND SMALL MOLECULE ACETYLATION IN PROKARYOTES

by

CHELSEY M. VANDRISSE

B.S., University of Minnesota, 2008

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

2018

PROTEIN AND SMALL MOLECULE ACETYLATION IN PROKARYOTES

by

CHELSEY M. VANDRISSE

Major Professor:	Jorge C. Escalante-Semerena
Committee:	Diana M. Downs
	Ellen L. Neidle
	William B. Whitman

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2018

DEDICATION

This dissertation is dedicated to myself. Earning a Ph.D. is no joke, and I worked hard.

ACKNOWLEDGEMENTS

I could write another 300 pages thanking and acknowledging those who have made my graduate career at the University of Georgia beyond enjoyable. I would first like to thank my mentor, Jorge C. Escalante-Semerena, who insists on constantly NOT being acknowledged. After reaching the end of my Ph.D., I finally see what he means by this, because a lot of what I've accomplished was due to my own drive and enthusiasm. However, Jorge taught me endless lessons about how to be a scientist and the difference between media and medium or the difference between affect and effect (although he would argue I haven't learned the latter yet, and I would have to agree). Even though my thoughts and ideas were my own, it was still fun to knock on his (always open) door and bounce experiments off of him and receive feedback that allowed me to return to the bench and make my own decisions about my projects. The independence that Jorge allowed me (and required of me) to have really let me develop into the scientist that I wanted to be and not a mold of what he wanted. This mentoring style really worked for me and I hope to take some of these lessons with me when I have my own lab one day. I think my favorite thing about being mentored by Jorge was that I could step into his office anytime to ask any question. I never felt judged or "stupid" and I always received an enthusiastic answer. Jorge doesn't like to be acknowledged for this, because he says it's his job, but I'm going to acknowledge him anyway because I think he did it well. Jorge, if you're reading this, thanks for all you've done for me, for now I have the confidence to be my own scientist. I'm in a position where I no longer need you as a mentor, but could use you as a colleague (when you retire I'd like your HPLCs, FPLCs, and plate readers – maybe your sirtuin grant as well).

And to my lab mates: I couldn't have made it through these six years without having you to ask life questions, to ask science questions, to eat lunch in the break room, or just to complain about the world and its injustices. Some may confuse my passion with anger. I know I'm intense pretty much all the time (unless its 6:30 pm, then I'm just "6:30 crazy"), and I thank my lab mates for understanding that side of my personality and giving me room to grow and understand how to funnel my passion into something productive. The phrase, "its 5 o'clock somewhere" will be replaced with "its 4 o'clock somewhere" because that was snack time (tasty snakes time) and it was there that I learned about my lab mates beyond a professional level. I think for a space to be productive, we must connect with each other beyond the surface. So while Jorge may have thought I was wasting time socializing (A LOT), I think that we were actually creating a space where we could have fun and trust each other. In turn, science was fun, work was play, and play was work. In the end, life is just one big party, so why not throw a tequila party in room 216 and call it a day.

Beyond my lab life, I fostered a very rewarding social life in Athens. This city has a lot to offer and I'm grateful to have landed at UGA. Thank you to Jittery Joe's for fueling the writing of this dissertation. Thank you to the South Eastern Women's Wrestling league who allowed me to express my anger and passion through the portrayal of FURRocious Frannie (sorry for breaking my wrist three months prior to graduation, Jorge). My non-science friends (you know who you are), thank you for teaching me something about life outside of the science world and for keeping my social skills intact. Thank you to my family, for loving me unconditionally (sorry not sorry about all the tattoos I obtained in graduate school), and because of your love, I finally feel comfortable to be the most authentic version of myself.

While I'll always remember the results of my projects, I am more likely to remember the people who were around during those projects and the fun we had during that time. Towards the end of my Ph.D., our lab was filled with six women and one male. I am so grateful to have been surrounded by badass women, where we were able to learn how amazing and powerful we were, especially when we stuck together. Early in my Ph.D., Diana Downs pointed out my "flakiness" during my professional interactions. This was mainly due to my lack of confidence and my belief that I could never be as strong, powerful, or smart as a man, so instead I had to be "funny" in order to be liked. Thank you to those that mentored me (Diana, Jorge, Jan, lab mates, and myself) to believe the opposite and allowed me to see my true potential. I'm still down for a good laugh, but now I am confident in my ability to dive deep into science and communicate my intelligence to the scientific community. I will carry on these lessons to people I meet in the future, and mark my words, if you embody misogyny or patriarchal values in my space, I will shut you down.

Thanks for reading,

Chelsey a.k.a. VanD a.k.a #33

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW: PROTEIN ACETYLATION IN BACTERIA

DISSERTATION OVERVIEW

Acetylation is a posttranslational modification conserved in all domains of life that is carried out by *N*-acetyltransferases. While acetylation can occur on N_{α}-amine groups, this thesis will focus on *N*ε-acetylation of lysyl residues and how this changes the cellular physiology of bacteria. Historically, acetylation has been studied in eukaryotes with regards to chromatin maintenance and transcriptional regulation. Recently, bacterial protein acetylation has been given much attention with regards to central metabolism, virulence, transcription, and translation. It seems that processes targeted by bacterial protein acetyltransferases are much more diverse than those studied in eukaryotes, making characterizing putative acetyltransferases challenging. The paradigm for acetylation in bacteria involves the acetylation of acetyl-CoA synthetase, whose activity must be tightly regulated in order to maintain cellular homeostasis. While this paradigm has provided much mechanistic detail for acetylation and deacetylation, this thesis uncovers advancements in the field that are changing the way we understand and interpret protein acetylation mechanisms in bacteria.

INTRODUCTIONAND LITERATURE REVIEW

Posttranslational modifications (PTMs) are a rapid way for the cell to modulate enzyme function or activity. This process allows a cell to quickly respond to internal or external stimuli,

bypassing the need to degrade, transcribe, and translate new proteins. Examples of PTMs include methylation (1), phosphorylation (2, 3), ADP-ribosylation (4), serine/threonine O-acetylation (5), succinylation (6), ubiquintinulation (7), adenylyation (8), S-ntirosylation (9), lipidation (10), glycosylation (11), phosphocholinylation (8), and for the focus of this review, acetylation (12).

Acetylation is a PTM conserved in all domains of life and was first discovered in yeast, where the Gcn5 acetyltransferase protein was found to acetylate lysyl residues of histone tails (12). Acetylation of these lysyl residues abolished their positive charge, blocking the ability of the histone tails to interact with negatively charged phosphate ions of the DNA backbone. The net effect is a relaxation of DNA strands, allowing for transcriptional machinery to access genes for activation (13). Since this discovery in 1960, studies involving acetylation have greatly expanded. A recent review by Hentchel and Escalante (14) provides additional information on the field and this review will focus on a brief background and the current literature dating from 2015 to present.

Bacterial protein acetyltransferases.

Acetylation in prokaryotes is carried out by Gcn5-related *N*-acetyltransferases (GNATs, PF00583), which are named after the yeast Gcn5 histone acetyltransferase. At the expense of acetyl-CoA, acetyltransferases transfer acetyl-moieties to N_α-amine groups of proteins and small molecules or N_ε-amine groups of lysyl residues of proteins (Fig. 1.1). Generally, acetylation of N_α-amine groups (*i.e. N*-termini of proteins) is used as a signal for protein degradation (15) and is irreversible. Acetylation of N_ε-amine groups targets active site lysyl-residues, thereby altering an enzyme's catalytic activity. The important aspect of N_ε-acetylation is its reversibility by protein deacetylases (discussed below), allowing for rapid adjustments to a protein's activity *in vivo*.

An organism's genome can code anywhere from ~1-70 acetyltransferases, each predicted to have different targets. For example, the *Streptomyces* genome codes for 72 acetyltransferases,

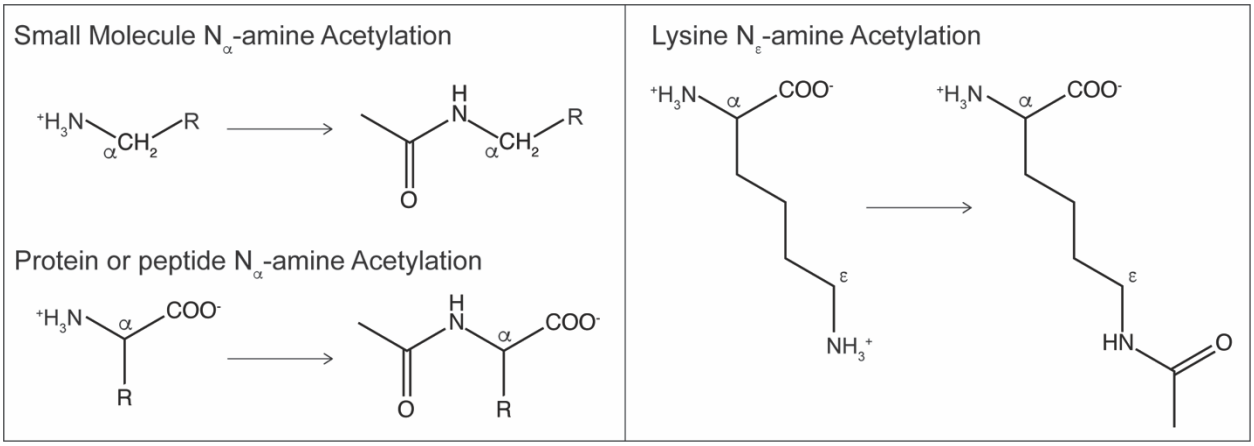


Figure 1.1. Reaction schematic of N_α- and N_ε- acetylation. Acetylation can occur by two primary mechanisms. Acetylation of small molecules or *N*-termini of proteins is characterized by addition of an acetyl group to N_α-amine groups (left panel). In left bottom panel, R represents an amino acid side chain of a protein or peptide. Generally, N_α-amine acetylation is irreversible. Acetylation of lysyl residues of proteins occurs posttranslationally at the N_ε-amine group and can be reversible by deacetylases.

only a few of which are characterized (16, 17). *Salmonella enterica* and *Escherichia coli* are model organisms for acetylation and about half of their 26 acetyltransferases are characterized (18-30). While structures of acetyltransferases have a conserved core domain consisting of β-sheets that bind acetyl-CoA through interactions with the pyrophosphate and pantothenate moieties (31), it is currently impossible to determine acetyltransferase target preference based on sequence or structure alone. Therefore it is difficult to determine if a putative acetyltransferase will target a protein or a small molecule, especially since deletion or overexpression of acetyltransferases do not usually lead to growth phenotypes under standard laboratory conditions.

Regardless of substrate specificity, acetyltransferases can be grouped into different types based on their domain organization. While others may exist, currently there are only four known domain organizations of acetyltransferases (Fig. 1.2) (16, 18, 32, 33). The *Salmonella* and *Escherichia* Pat (for protein acetyltransferase), is the first type I acetyltransferase to be discovered and consists of a large *N*-terminal domain (~700 residues) of unknown function and a GNAT-catalytic *C*-terminal domain (~200 residues). While the physiological function of the large *N*-terminal domain hasn't been determined, studies showed that *Se*Pat binds two molecules of acetyl-CoA, one on the *N*-terminal and *C*-terminal domain (34). It is possible that binding of acetyl-CoA to the large *N*-terminal domain is a regulatory mechanism to cue acetylation, although this has not been studied in depth.

Type II protein acetyltransferases were discovered in *Streptomyces lividans* (PatA) (16) and contain a *N*-terminal GNAT-catalytic domain (~200 residues) with a *C*-terminal regulatory domain (~900 residues) of unknown function. In addition to the reversed domain organization of Type I acetyltransferases, the Type II acetyltransferase of *S. lividans* contains a proline-rich linker that includes a collagen-like GPS motif. While the precise role of the regulatory *C*-terminal domain

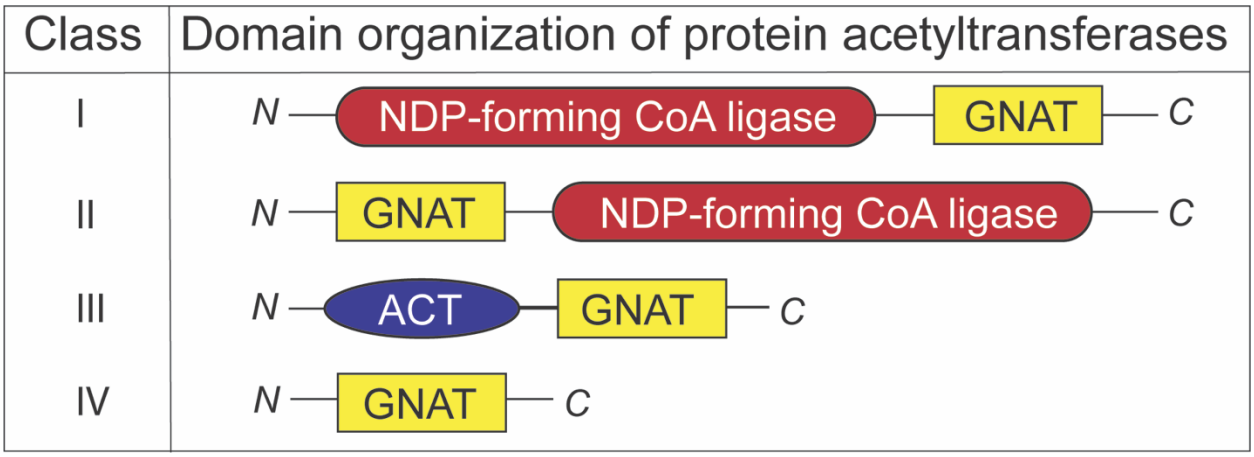


Figure 1.2 Domain organization of bacterial protein acetyltransferases. Acetyltransferases belonging to classes I-IV have been identified within prokaryotic genomes. Type I acetyltransferases contain a large regulatory *N*-terminal domain of unknown function. These domains are homologous to NDP-forming CoA ligases, although they lack a critical active site histidine residue, and do not possess this activity. Within the *C*-terminus of these proteins, a GNAT catalytic domain carries out the acetylation event. Type II acetyltransferases are similar to Type I, but their domain organization is reversed. The catalytic GNAT domain is within the *N*-terminus, with a large regulatory domain on the *C*-terminus. Type III acetyltransferases are also similar to Type I, in that their regulatory domain lies on the *N*-terminus. However, the regulatory domains of Type III acetyltransferases are different in size from that of Type I (~300-400 residues versus 800-900 residues, respectively). An ACT domain (ACT, representing the enzymes initially found with this domain, aspartate kinase, chorismate mutase and TyrA) serves regulatory function. Lastly, Type IV acetyltransferases are unique in that they only contain an catalytic GNAT domain. It is thought that Type I-III acetyltransferases target proteins, while Type IV are able to target both protein and small molecule.

of *Sl*PatA is unknown, removal of this domain renders the catalytic *N*-terminus inactive (35). It appears that other Actinomycetes have Pat enzymes that have similar domain organizations, suggesting acetylation machinery may have evolved within environmental niches. In both type I and type II acetyltransferases, the large regulatory domain has sequence homology to ADP-forming CoA synthetases (not to be confused with AMP-forming CoA synthetases that are regulated by acetylation, discussed below). These regulatory domains do not show catalytic activity as ADP-forming CoA synthetases, as they have a critical histidine residue variation (18) within the predicated catalytic domain. It is not known why these enzymes have evolved with these particular regulatory domains and their function has not been shown.

Type III protein acetyltransferases are similar to type I, in that they contain a *N*-terminal regulatory domain and a *C*-terminal GNAT catalytic domain (Fig. 1.2). They differ, in that their regulatory *N*-terminus is smaller (~300-400 residues) and their specific regulatory functions are known. In *Mycobacterium smegmatis*, the *N*-terminus of PatA binds cyclic-AMP, which increases the activity of this acetyltransferase (36). *Ms*PatA acetylates a universal stress protein (37) and acetyl-CoA synthetase (14). In *Micromonospora* and *Streptomyces*, a type III acetyltransferase (PatB) has an ACT-regulatory domain (ACT, representing the enzymes initially found with this domain, aspartate kinase, chorismate mutase and TyrA) that binds amino-acids to increase its acetyltransferase activity (17, 38). In *Micromonospora*, binding of L-cysteine or L-asparate increased the ability of PatB to acetylate acetyl-CoA synthetase (Acs) of this organism. These ACT-domain acetyltransferases are only found in actinomycetes, raising interesting questions about links in amino acid metabolism and acetylation of these organisms.

Lastly, Type IV acetyltransferases have no regulatory domain and consist only of a catalytic GNAT domain (Fig 1.2.). Acetyltransferases with this domain organization are the most

abundant within a genome. For example, *S. enterica* has a single type I acetyltransferase and 25 type IV acetyltransferases. It is generally found that within any given organism, its genome will code for one to two acetyltransferases that fall into the type I-III category, with the rest of the acetyltransferases being type IV. There are a few known exceptions to this rule, as seen in *Bacillus subtilis*, which has a genome coding for exclusively type IV acetyltransferases. It is generally assumed that type I-III acetyltransferases target protein substrates, with the most commonly known substrate being CoA synthetases, while type IV will target protein or small molecule substrates. To date, no type I-III acetyltransferases have small molecules targets.

Bacterial protein deacetylases.

Reversibility of lysine protein acetylation is carried out by deacetylases. Due to their discovery in yeast where they were shown to de-modify acetylated lysyl residues of histone tails, deacetylase enzymes are categorized as HDACs (for histone deacetylases, PF08295). HDACs can be broken down into four classes, with classes I, II, and IV catalyzing lysine deacetylation without cofactors (39, 40). In these classes of HDACs, Zn(II) is required for deacetylation and binds to a water molecule that then triggers a nucleophilic attack on the carbonyl of the acetyl group. This reaction mechanism releases acetate as a byproduct.

Class III HDACs, also known as sirtuins (PF02416, named after the yeast SIR2 protein), require NAD^+ as a cofactor to remove the acetyl-group from acetylated lysyl residues (41). The acetyl-moiety forms an intermediate with NAD^+, which causes the release of nicotinamide (an inhibitor of sirtuin activity). This leads to deacetylation of the acetyl-lysyl residue and formation of *O*-acetyl-ADP-ribose (*O*AADPR). The fates and physiological roles of *O*-AADPR in prokaryotes isn't fully understood. Prokaryotic genomes tend to code for one to two sirtuin homologues, while some also code for HDACs. For example, *Salmonella enterica* has a single

sirtuin NAD^+-dependent deacetylase (CobB), while *Bacillus subtilis* has one sirtuin deacetylase (SrtN), and one HDAC (AcuC) (42, 43). There are seven isoforms within the class III group of sirtuins (SIRT1-7) and the most commonly studied prokaryotic sirtuins are SIRT5 isoforms (42). However, recent work in *Streptomyces* showed that its genome coded for one SIRT5 deacetylase (CobB), one SIRT4 deacetylase (SrtA), and a HDAC (AcuC) (17). It is predicted these *Streptomyces* enzymes target different acetylated substrates, with CobB targeting acetoacetyl-CoA synthetase (16) and SrtA targeting acetyl-CoA synthetase (17). How these deacetylases have evolved their substrate specificity is of interest.

Important to note, recent work has identified an additional class of sirtuins that do not catalyze a deacetylase reaction, but rather carry out their own PTM. Known as SirTM's, these enzymes in *Stapylococcus aureus* lack the classic histidine active site of deacetylase sirtuins, and instead possess ADP-ribosyltransferase activity. Lipoylation of Lys56 of GcvH (a lipoyl-carrier protein involved in oxidative stress) is required before SirTM can ADP-ribosylate an aspartate residue (D57) of the GcvH-L protein (44). It was predicted that ADP-ribosylation of GcvH-L inhibited its interaction with oxioreductase, allowing the cells to keep its response "off" under non-stress conditions.

The paradigm of reversible lysine acetylation.

The paradigm for protein acetylation in bacteria is the modification of acetyl-CoA synthetase (Acs) (18, 42, 45-48) (Fig. 1.3A). Acs is an AMP-forming CoA synthetase that has multiple conformations (apo, adenylation, and thioesterification). The adenylation conformation allows for binding of acetate and ATP, forming an acyl-AMP intermediate (Fig. 1.3B) (42). The esterification conformation produces acetyl-CoA and releases AMP (Fig. 1.3B). This same mechanism applies to many other bacterial CoA synthetases that utilize a wide range of fatty acids.

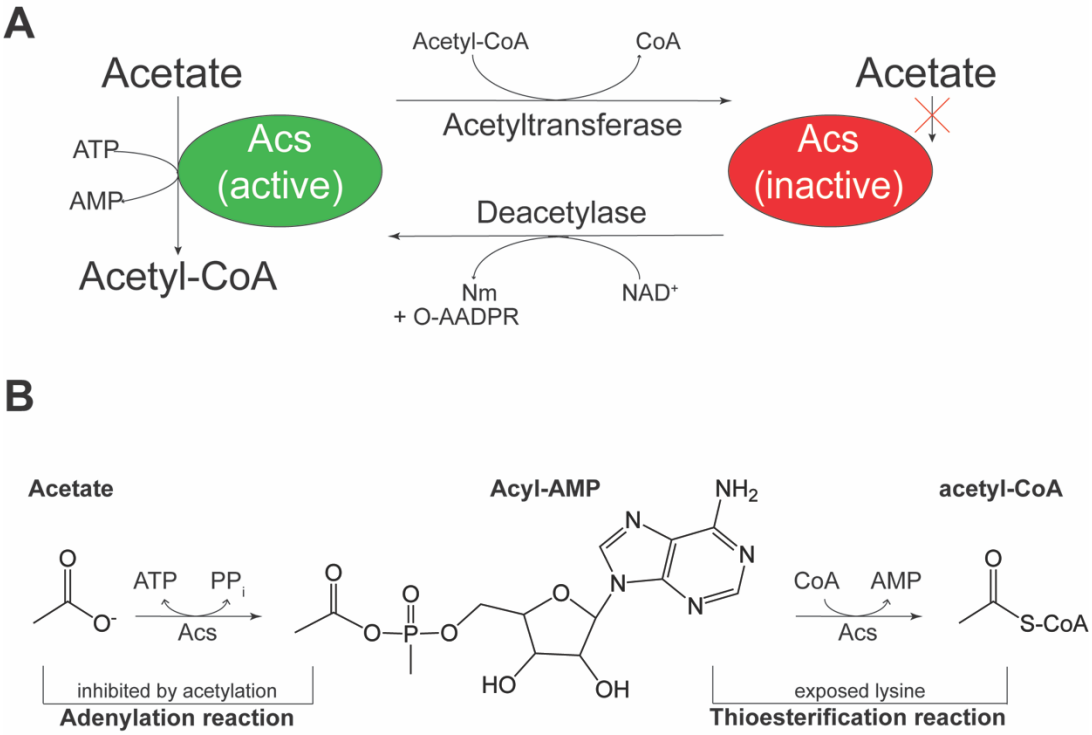


Figure 1.3 Reversible lysine acetylation of the acetyl-CoA synthetase (Acs). A. Pictured is a model of Acs reversible lysine acetylation as studied in *Salmonella enterica*. Acs is a central metabolic enzyme that under low acetate concentrations (10 mM), activates acetate to acetyl-CoA (green protein). A catalytic lysyl residue (Lys609) coordinates the carboxylic acid of acetate into the active site and acetylation of Lys609 halts Acs activity (red protein). The acetyl-group on Lys609 can be removed by a protein deacetylase, returning the enzyme to its active form. In *S. enterica*, the Type I acetyltransferase, Pat, acetylates Acs and the NAD^+-dependent sirtuin deacetylase, CobB, removes this modification. **B.** Activation of acetate to acetyl-CoA occurs in a two-step reaction mechanism. In the adenylation reaction, Lys609 of Acs binds ATP and acetate, forming an acyl-AMP intermediate. In an esterification reaction, AMP is replaced by CoA, releasing free AMP. It is in this conformation that the active site lysyl residue is exposed and likely targeted by acetylation. Acetylation of Lys609 blocks the adenylation reaction and stops the synthesis of acyl-AMP intermediates. Increased levels of AMP in the cell indirectly lead to lower levels of ATP synthesis, and this depletion is the reason this enzyme is tightly regulated by reversible lysine acetylation. *O*-AADPR; *O*-acetyl-ADP-ribose.

High formation of AMP following this esterification can lead to energy disruptions in the cell, through inhibition of ADP synthesis (*i.e.* lower levels of ATP). In *Salmonella enterica*, Lys-609 of Acs coordinates the negatively charged carboxylate of acetate into the active site, where the protein forms the acyl-AMP intermediate (45). Crystal structures show that in the esterification conformation of CoA synthetases, the active site lysyl-residue is surface exposed, making it readily available for acetylation (49). Under conditions where AMP levels are too high, an acetyltransferase can target this residue, thereby inhibiting the initial acyl-AMP forming step. When activity of Acs is needed, a deacetylase can remove the acetyl group on this lysyl residue, returning the protein to its active form (42). This is a prime example of the importance and rapid nature of acetylation in prokaryotes and has been the basis for studying acetylation of other enzymes in the cell.

Sources and fates of acetyl-CoA

Acetyl donors for acetylation can come from two known sources, acetyl-CoA (Ac-CoA) and acetyl-phosphate (Ac-P). Metabolically, Ac-P is generated under high acetate conditions (50 mM) where acetate kinase (AckA) activates acetate to acetyl-CoA. Acetyl-CoA is then converted to Ac-P by the phosphotransacetylase (Pta) (50, 51). The high energy Ac-P molecule can be used as an acetyl donor for chemical acetylation of lysyl residues of proteins (discussed below in protein acetylation: enzyme *vs* chemical). Alternatively, under low acetate conditions (10 mM), acetyl-CoA synthetase (Acs) activates acetate to acetyl-CoA. Independently of acetate metabolism, acetyl-CoA can originate from metabolism of other short chain fatty acids or via central metabolism of glucose molecules. Proteolysis and oxidation of amino acids can also lead to increased levels of acetyl-CoA (Fig. 1.4).

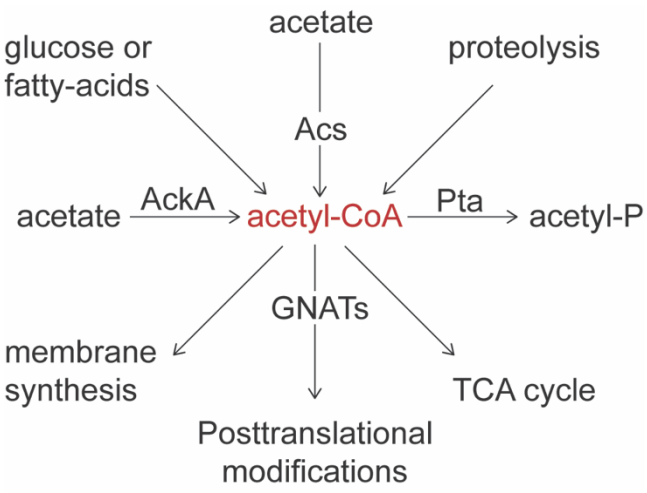


Figure 1.4. Sources and fates of acetyl-CoA. Under low acetate conditions (10 mM), acetate can be converted to acetyl-CoA via Acs (as described in Figure 3). Under high acetate conditions (50 mM), acetate kinase (AckA), activates acetate to acetyl-CoA, while the phosphotransacetylase (Pta) converts acetyl-CoA to acetyl-phosphate (acetyl-P). Metabolism of glucose and other short chain fatty acids also lead to the production of acetyl-CoA. Proteolysis leads to an increased accumulation of free amino acids, which are eventually deaminated and oxidized to acetyl-CoA. All of these sources of acetyl-CoA can funnel this high energy molecule into synthesis of fatty acids for membrane production, lead into the TCA cycle for production of reducing equivalents during respiration, or can be utilized as source for acetylation reactions.

Once acetyl-CoA is produced, it has a variety of fates within the cell. Acetyl-CoA can be funneled into the TCA-cycle for production of reducing equivalents during respiration, for the production of fatty acids for membrane synthesis, and for use in posttranslational acetylation reactions (Fig. 1.4). It is believed that acetyl-CoA levels trigger acetyltransferases to modify metabolic enzymes, such as CoA synthetases, although how this cascade is initiated is not well understood.

PROTEIN ACETYLATION: ENYZMATIC *VS.* CHEMICAL

Currently, it should be considered that enzymatic protein acetylation is of high physiological relevance. Acetylation machinery has evolved and been encoded by the genome, with each acetyltransferase predicted to have a unique and specific target. While it is not known what triggers activity of an acetyltransferase, we know that acetylation is necessary to maintain cellular homeostasis (as described above), cue transcription of essential genes, and alter cellular metabolism based on environmental stressors or niches. This process has been shown to be of importance because deletion of acetyltransferases under specific conditions changes the metabolism or physiology of a cell. The enzymatic nature of an acetyltransferase lies within its ability to extract protons from lysyl residues, which is carried out by an active site glutamate (Fig. 1.5) (31). At cellular pH, lysyl-residues are protonated and positively charged due to the the pK_a (10.53) of this side chain. The negatively charged catalytic glutamate residue acts as a base that facilitates a water-mediated proton extraction from the lysyl residue. The deprotonated lysyl residue then undergoes a nucleophilic attack on the carbonyl-group of acetyl-CoA, generating acetyl-lysyl residues.

While the catalysis is specific and the core domain containing this glutamate is highly conserved, the residues important for recognition of acetyltransferase substrates is poorly

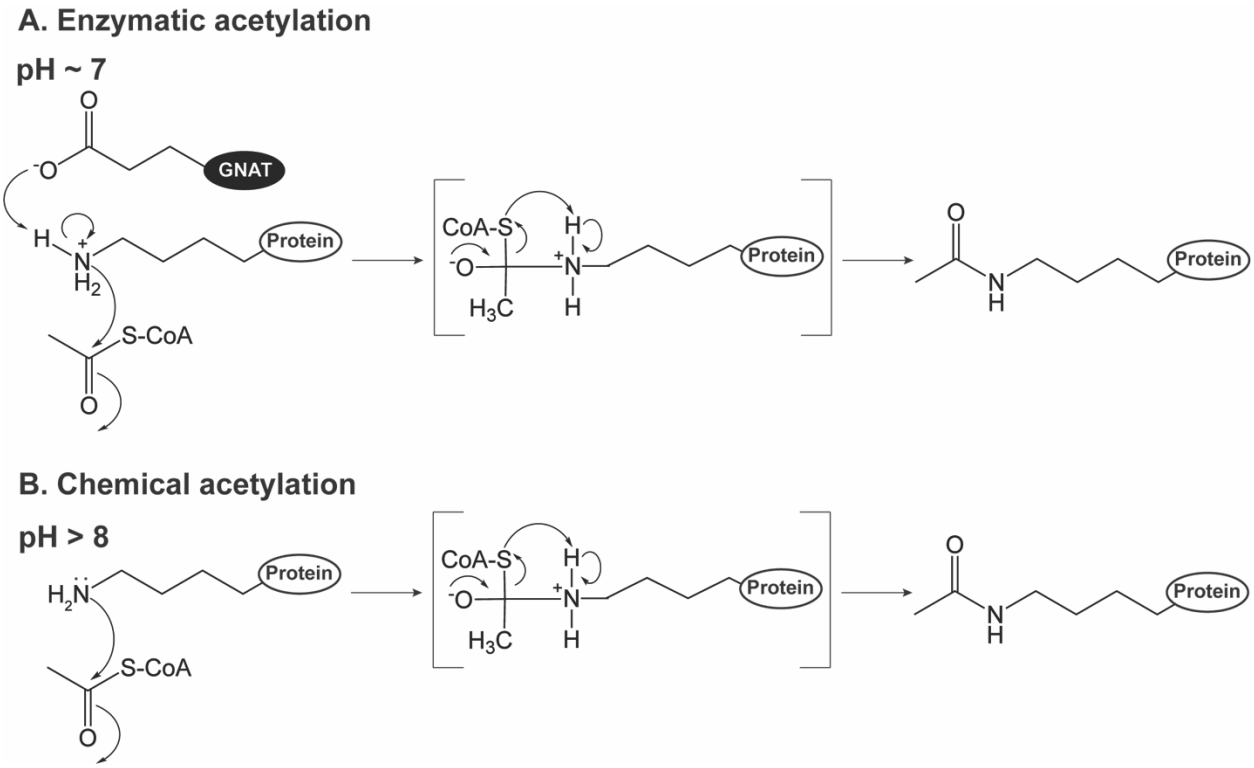


Figure 1.5. Reaction mechanism of protein acetylation. A. GNAT-mediated protein acetylation involves a catalytic glutamate residue within the acetyltransferase core domain. The pK_a of glutamate is 4.25 and at cellular pH (~7), the side chain is deprotonated. This allows the glutamate residue to act as a base that facilitates a water-mediated proton abstraction from the lysyl residue (pK_a = 10.53) of the target protein. The e-amine group of the lysyl residue then undergoes a nucleophilic attack on the carbonyl carbon of the acetyl moiety of CoA, leading to the formation of an acetylated lysyl side chain. **B.** Under high pH conditions, a larger portion of lysyl residues are deprotonated, bypassing the need for a catalytic proton abstraction from the side chain.

understood. Regardless, acetyltransferases have evolved this specificity, and the enzymes (or small molecules) they target are intentional and mechanistic. For this reason, characterizing putative acetyltransferases is of utmost importance. Many papers discussed in the remainder of this review, identify lysyl residues that are acetylated, but not the acetyltransferases carrying out this reaction. We would argue that in order to truly understand physiological relevance of acetylation reactions, we must identify the enzymes that have evolved to control protein activity.

Due to the pK_a of a lysine side chain, it is also important to point out the relevance of carrying out *in vitro* acetylation assays at a cellular pH (~7). When the pH of a reaction mixture is >8, a higher percentage of lysyl residues are deprotonated non-enzymatically, and can be chemically acetylated by either Ac-CoA or Ac-P (Fig. 1.5). While it is possible these acetylation events may be relevant, often these chemical acetylation events do not alter an enzyme's activity. It is therefore important not only to assess enzyme activity and function after an *in vitro* acetylation experiment, but almost more important to identify the cellular acetyltransferase responsible for this PTM. Therefore, this review will not focus on global protein acetylome studies, as these experiments generally lead to large levels of chemical acetylation due to high pH conditions and the targets are not verified with pure proteins *in vitro*.

CELLULAR PROCESSES UNDER RLA CONTROL

Carbon-nitrogen metabolism

Every cell must undergo carbon and nitrogen metabolism for survival. Consequently, acetylation of the proteins involved in these processes is key to understanding the regulatory and evolutionary mechanisms of acetylation. In prokaryotes, the paradigm for acetylation is post-translational control of the AMP-forming acetyl-CoA synthetase (Acs) (Fig. 1.3) (18, 42).

Acetylation of Acs is necessary for maintaining cellular homeostasis (52), because dysregulation of this enzyme is detrimental to growth (52). Since the discovery of Acs acetylation, understanding the control of this enzyme has been studied in depth.

While we know acetylation of central metabolic enzymes is necessary for adaptation, the cascade of events that trigger acetylation remain to be elucidated. Acs, like other members of the ANL superfamily of proteins (53) (https://pfam.xfam.org/clan/ANL) has AMP/ATP binding pockets that when occupied, stimulate an "open" conformation of the protein (54). This open conformation exposes the active site lysine of Acs, rendering it available for acetylation (54). It has recently been shown that acetylation is increased when an AMP analog, cyclic-AMP (cAMP), binds to the AMP pocket of Acs (55). While cAMP was bound to acetylated Acs, its deacetylation levels also decreased (55). This could imply that cAMP is a competitor for AMP, which serves as a cellular signal for protein acetylation and control. Not only was this seen with Acs, but other CoA synthetases (FadD and PrpE) of the cell.

An additional mechanism to control acetylation of Acs is through regulation of the post-translational machinery that targets this enzyme. Multiple studies have shown that central regulatory enzymes control transcription of the genes coding for acetyltransferases and deacetylases (56-58). The nitrogen response regulator, GlnR, has been shown to upregulate expression of the *Saccharopolysporea erythraea* acetyltransferase, *acuA*, and deacetylase, *srtN,* under nitrogen limiting conditions (57). In the absence of GlnR, *S. erythraea* displayed acetylation dependent Acs growth phenotypes that were related to lower levels of the *acuA* acetyltransferase. This work linked nitrogen metabolism to regulation of central metabolic acetylation machinery. Other studies have shown acetylation of GlnR alters *in vitro* DNA-binding, although these studies require further analysis to assess physiological relevance (59).

Even though acetylation of Acs is thought to be conserved in all domains of life, how this acetylation event differs amongst bacterial species has recently been investigated. In pathogenic bacteria (*i.e. S. enterica*, *E. coli*, and *S. aureus*), Acs is acetylated by a large type I protein acetyltransferase (Pat) (18) (Fig. 1.2 and 1.3). This Pat protein contains a regulatory domain of unknown function (~90 kDa) attached to the catalytic GNAT domain (~20 kDa). Pat acetylates Acs on a single conserved lysine residue, thereby abolishing its activity while maintaining cellular homeostasis (52). However, we are beginning to note differences in how Acs is recognized and acetylated, from work that highlights Acs acetylation in *Micromonospora* (38) and *Streptomyces* (17).

Work in *Micromonospora* identified an ACT-domain acetyltransferase (ACT abbreviated for the enzymes initially found with this domain, aspartate kinase, chorismate mutase and TyrA) that bound cysteine and arginine, which increased its activity for Acs acetylation (38). The homologous ACT-Pat protein was identified in *Streptomyces* (17) that also modified *Streptomyces* Acs. The most important findings of these studies were: 1) Pat enzymes do not show cross talk for Acs acetylation (*i.e Salmonella* Pat cannot recognize *Streptomyces* Acs and *Streptomyces* ACT-Pat cannot recognize *Salmonella* Acs) and 2) It was found that a serine residue within *Stretpomyces* Acs was acetylated and that this serine-modification was required for ACT-Pat mediated Acs acetylation. It seems that acetylation mechanisms have evolved to be specific within each species and interestingly, seems to be thus far coordinated with cellular environment (*i.e. Salmonella* and *Escherichia* are intestinal pathogens with identical acetylation machinery, while *Micromonospora* and *Streptomyces* are soil microorganisms with identical acetylation machinery).

While acetylation of Acs has and continues to be analyzed in depth, studies are expanding to assess PTMs of additional central metabolic enzymes. A good place to start for investigating lysine acetylation is to identify central metabolic enzymes that have active site lysines (60). As mentioned previously, the positive charge of lysyl residues coordinates a substrate into an enzyme's active site. Lysine acetylation abolishes this charge and sterically blocks catalytic pockets, which inhibits enzyme activity. Enolase, a central metabolic enzyme that converts 2-phosphoglycerate to phosphoenolpyruvate, has been found to be acetylated in many acetylome studies (61, 62). Work has been done that shows *in vitro* chemical acetylation of enolase decreased its activity and that acetylation mimics (Lys to Gln) of active site enolase lysyl residues reduced its activity (60). Other work in *Mycobacterium tuberculosis* identified two acetylation cites of isocitrate lyase (63). Acetylation of one lysine increased activity of the enzyme while acetylation of the other decreased enzyme activity. Interestingly, growth of *M. tuberculosis* on acetate and propionate increased acetylation levels of isocitrate lysase. It is possible this was due to chemical modifications from increased acetyl-CoA and acetyl-P pools, as an enzyme to carry out acetylation was not found. To fully understand the physiology of an organism related to acetylation, we must identify the putative acetyltransferases responsible for these aforementioned chemical acetylation events.

In acetylome studies, many enzymes are acetylated on multiple lysines. Therefore, verification of these acetylation events is pivotal for development of the field. One such study characterized the acetylation of phosphoenolpyruvate carboxylase (PEPC), an enzyme that converts phosphoenolpyruvate to oxaloacetate (64, 65). Metabolic flux after production of oxaloacetate leads to increased synthesis of aspartate amino acids and ensures carbon flow through the citric acid cycle for production of glutamate. *Cornyebacterium glutamicum* is a

bacterium used for industrial production of glutamate and under glutamate producing conditions, the authors found PEPC was acetylated at Lys653 (64). Although cells with PEPC acetylation mimics (K653Q) did not have growth phenotypes, glutamate production was completely abolished and *in vitro* activity was greatly reduced compared to $PEPC^{WT}$. A PEPC deacetylation mimic (K643R) produced high levels of glutamate and maintained moderate activity *in vitro*. These results showed that the charge of Lys653 is not only necessary for enzyme activity, but that by abolishing enzymatic activity in the $PEPC^{K653Q}$ variant, metabolic flux was greatly impacted. While the deacetylase of *C. glutamicum* was able to deacetylate artificially acetylated PEPC, no acetyltransferase was identified for this modification. Although acetyl-P or acetyl-CoA could be a substrate for chemical modification of PEPC, we would argue PEPC acetylation would have to be tightly regulated and is likely enzymatic.

Other work has begun to investigate PTMs of enzymes involved in overproduction of antimicrobial-producing enzymes (66), and highlight that PTMs can greatly alter the derived products. The work by Xu et al. provides an excellent example of malonylation of the above mentioned enzymes, but since the focus of this review is on acetylation, the results will not be discussed in detail. Both of these studies (64, 66) highlight the importance of understanding the effects of protein acylation on metabolic flux. For industrial production of small molecules, we cannot only be thinking of the synthesized enzymes, but how those enzymes may be post-translationally modified, and how we can use that information to better engineer microbes to produce optimal yields of products.

TRANSCRIPTION

The ways in which acetylation alters gene regulation has been studied extensively in eukaryotes through the process of histone acetylation (12). Histones are nucleoid associated

proteins that have lysine rich DNA binding domains that interact with the negatively charged DNA backbone (Fig. 1.6). The yeast Gcn5-acetyltransferase targets lysines within the DNA binding domain of histones H3 and H2B (67), while yeast EsaI acetylates histone H4 and H2A (67). While histone acetylation has been given much attention, the prokaryotic mechanisms from which it evolved have recently been investigated. Although acetylation of transcription factors had been reported in acetylome studies (61, 62), it wasn't until recently that an acetyltransferase was found to acetylate and alter DNA-binding of a prokaryotic transcription factor (68). Since then, an array of studies has been conducted to further elucidate these mechanisms.

The prokaryotic DNA-binding protein, HU, is a small 10 kDa protein that most closely resembles eukaryotic histones. HU is similar to histones, in that it causes DNA supercoiling, with the assistance of topoisomerase. HU has been found to be acetylated in multiple acetylome studies, however only in *Mycobacterium* has an acetyltransferase been identified to directly acetylate HU (69). This acetyltransferase, Eis, modified *Mt*HU and altered its DNA binding capabilities *in vitro* (69). Mass spectrometry revealed 29 possible acetylation sites in the *C*-terminal domain, but currently it is unknown if Eis targets one or multiple lysyl residues of *Mt*HU. Interestingly, overexpression of *eis* in *M. tuberculosis* caused enhanced cell survival in macrophages, highlighting that acetylation plays an important role in pathogenesis of this organism. Further work identified a deacetylase, Rv1151c, that reversed the acetylation of HU by Eis (70). Deacetylation of *Mt*HU restored its DNA binding capabilities and together this work encompassed a deacetylase/acetyltransferase system that targets a DNA binding protein.

In addition to HU, another DNA binding protein in *M. tuberculosis*, DosR, was found to be acetylated (71). A *Mtb* cAMP-domain acetyltransferase, RV0998, acetylated DosR *in vitro*. Lys182 of DosR is known to bind DNA, and under hypoxia it was found this lysyl residue had

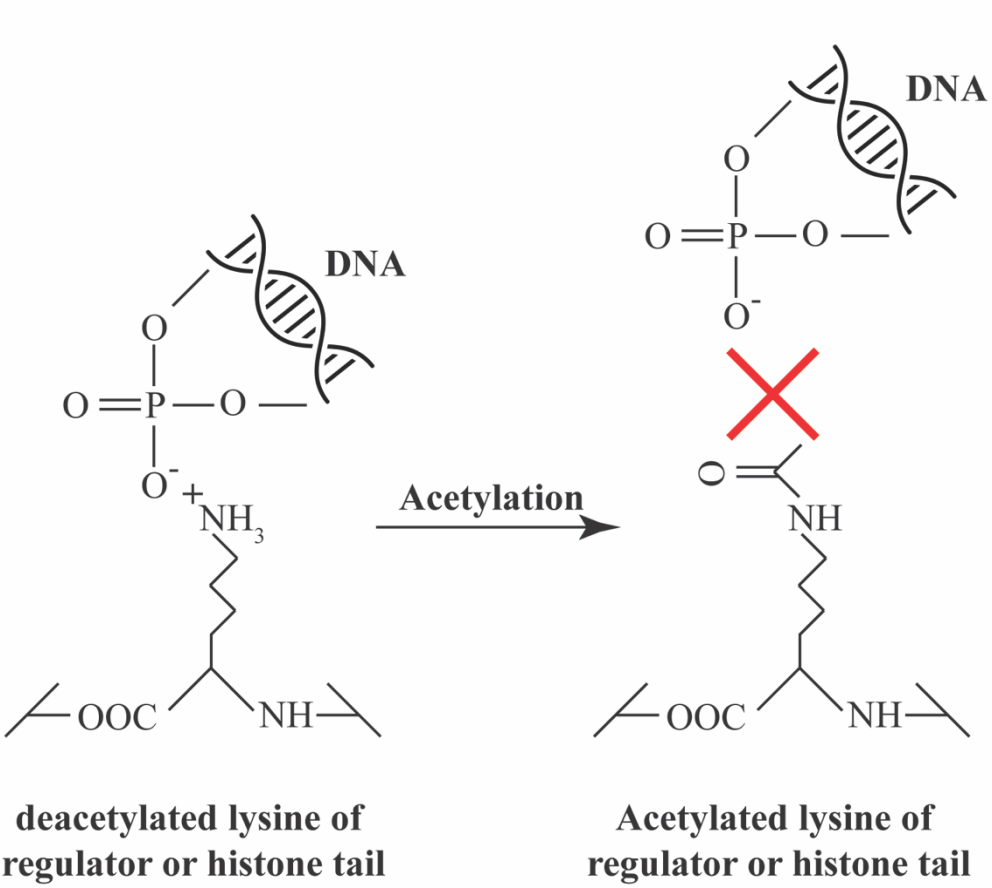


Figure 1.6 . Acetylation schematic of lysyl residues of histone tails or prokaryotic DNA binding proteins. Deacetylated lysyl residues (left) maintain a positive charge at cellular pH and interact with the negatively charged phosphate moieties on the DNA backbone. Upon acetylation of these lysyl residues by acetyltransferases (right), the charge on the side chain is abolished and can no longer interact with the DNA, causing a disruption in the interactions.

lower levels of acetylation. To replicate deacetylation of DosR under hypoxia, variants were construct to mimic deacetylated DosR ($DosR^{K182R}$). Presence of $DosR^{K182R}$ lead to a significant change in transcription of many genes (*acr*, *dosR*, *fdxA*, and others). The $DosR^{K182R}$ deacetylation mimic bound all DNA in electrophoretic mobility shift assays while an acetylation $DosR^{K182Q}$ mimic did not. Interestingly, a RV0998 acetyltransferase knockout had reduced growth in macrophages, while $DosR^{K182R}$ producing cells inhibited intracellular survival of *Mtb*. Overall, it was hypothesized that hypoxia may induce deacetylation of DosR, which increases its DNA binding ability to promote transcription of target genes needed for *Mtb* to adapt to hypoxia during virulence. While work needs to be done to understand what triggers acetylation of DosR, this work highlighted an additional reversible lysine acetylation system of *Mtb* that is necessary for pathogenesis.

Much like *Mycobacterium*, pathogenesis is affected via acetylation of transcriptional regulators in *Salmonella enterica*. Work done in this organism identified acetylation of HilD, which is a regulator that binds and induces expression of *hilA.* HilA is a regulator needed to induce transcription of *Salmonella* pathogenicity island (SPI). The protein acetyltransferase (Pat) of *Salmonella* acetylated Lys297 of HilD, while surprisingly the sirtuin deacetylase of *S. enterica*, CobB, could not reverse this modification. Lysine acetylation ($HilD^{K297Q}$) and deacetylation ($HilD^{K297R}$) mimics bound DNA to varying degrees, with the acetylation mimic binding DNA to a lesser degree. Additionally, $HilD^{Ac}$ had a 60% reduced invasion rate in *Salmonella* compared to deacetylation mimics. Lastly, $HilD^{Ac}$ reduced *Salmonella* colonization *in ceca.* Acetylation is used to fine tune HilD activity during *Salmonella* infections, since it is hypothesized that too much HilD can be detrimental.

TRANSLATION

Acetylation has been thought to be a part of regulating translation since the discovery of Rim acetyltransferases. Three acetyltransferases, RimI, RimJ, and RimL acetylate N_{α}-amine groups of ribosomal proteins S18, S5, and L12, respectively (24, 25). Since then, many indirect mechanisms of acetylation have been found to alter translation efficiency in prokaryotes.

The most recent and noteworthy research has been with toxin-antitoxin systems in bacteria, where the toxin codes for an acetyltransferase. Work in *Salmonella enterica* and *Escherichia coli* has shown multiple toxin acetyltransferases that acetylate amino-acyl tRNAs (20, 30, 72). It was found that three toxin-antitoxin acetyltransferase systems in *Salmonella enterica* acetylated specific aminoacyl-tRNAs to varying degrees (20, 30). In *E. coli*, the acetyltransferase toxin identified acetylated only formyl-methionine tRNA, although it is unclear the site of acetylation, since there is no free amine group on f_{met}-tRNA (72). It is possible that the formyl-group is lost immediately before acetylation by the toxin. Acetylation of free amine groups on aminoacyl tRNAs blocks the ability of the ribosome to form peptide bonds (Fig. 1.7). The result is a block in protein synthesis and in the case of pathogenic organisms, entrance into a persister state of infection. These persister states are of importance, because cells shut down essential cellular machinery to evade host defenses. Once the cell has avoided the threat, they are able to halt aminoacyl-tRNA acetylation and return to active growth.

It is unknown what triggers the acetylation of amino-acyl tRNAs in these organisms, but hypotheses involving acetylation of the antitoxin have been studied (21). In this work, it was shown that acetylation of the antitoxin by its cognate toxin increased the toxin's activity for acetylation of aminoacyl-tRNAs (21). The sirtuin deacetylase of *S. enterica*, CobB was able to deacetylate the antitoxin, which returned the toxin to its original activity for aminoacyl-tRNA

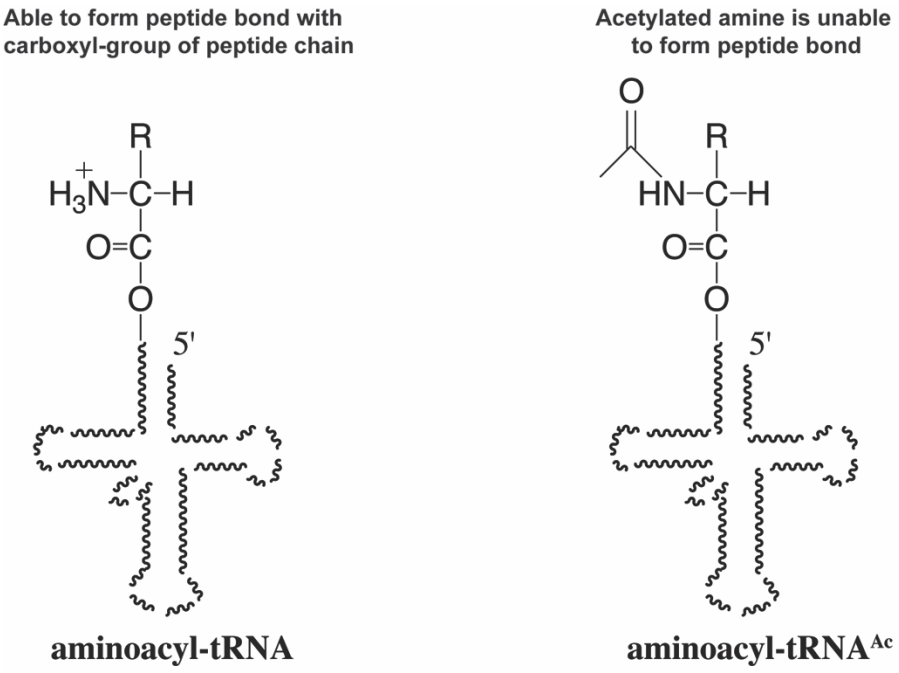


Figure 1.7. Acetylation of aminoacyl-tRNAs blocks peptide bond formation and inhibits protein synthesis. Under *S. enterica* and *E. coli* non-infectious states of growth, aminoacyl-tRNAs in the P site of the ribosome contain a free amine group that is able to undergo a nucleophilic attack with the carbonyl-group of the peptide chain in the A site. Many toxin acetyltransferases in these organisms have been found to acetylate this free amine group (right), which blocks the ability to undergo a nucleophilic attack with the carbonyl group in the P site of the ribosome, which in turn blocks protein synthesis.

acetylation. This work was intriguing because it was the first instance of an acetyltransferase targeting a protein (antitoxin) and non-protein (aminoacyl-tRNA) substrate. Additionally, this work highlighted a toxin that modulated its own activity through acetylation of its antitoxin.

In addition to the work in *Salmonella* and *Escherichia*, other groups have identified toxin-antitoxin systems coding for acetyltransferase toxins that are predicted or shown to be involved in inhibiting protein synthesis (73-76). It would be of importance to study other toxin-antitoxin acetyltransferase systems to determine if the toxins maintain their activity while in complex with their antitoxins and further determine the mechanisms in which cells enter and leave persister cell states. The current hypotheses in the field involved recycling of acetylated aminoacyl-tRNAs by peptidyl-tRNA hydrolase (pth) (20) or through acetylation/deacetylation of the antitoxin to control the toxin's aminoacyl-tRNA acetyltransferase activity (21). It is possible that it is a combination of both processes. Many reviews have been written on toxin-antitoxin systems and can be found elsewhere (77-81).

Acetylation has been shown to affect the activity of tyrosyl-tRNA synthetase (TyrRS) in *Escherichia coli* (82) and was shown through genetically coding for acetyllysines on specific positions of TyrRS. These variants had lowered catalytic efficiencies and the CobB sirtuin deacetylase was able to remove these chemical modifications, returning TyrRS to its wildtype activity. A noteworthy experiment of this work, was that authors purified 25 *E. coli* acetyltransferases and tested for *in vitro* acetylation of TyrRS. The only potential acetyltransferase candidate was Pat, but the activity of TyrRS was not tested after enzymatic acetylation. Authors suggested that perhaps acetyl-P was the source of acetylation *in vivo*, although it would be interesting to further study the effects of Pat on this system.

LIMITATIONS OF CURRENT RESEARCH

After review of the literature, there are some important limitations that should be noted and avoided in future experiments. It seems the biggest block in advancement of the field of bacterial protein acetylation lies within our ability to identify and characterize acetyltransferases. Many proteins that were found to be acetylated (as mentioned above), do not have PTM machinery that targets specific lysyl residues. It is important to note, that many of these studies generate acetylated proteins either through chemical acetylation or by genetically coding for an acetyl-lysyl residue at a specific site. Afterwards, it is common to utilize the *S. enterica* or *E. coli* CobB sirtuin NAD^+-dependent deacetylase to show removal of this acetyl group. However, due to the fact a cell's genome only codes for one to three deacetylases, these enzymes tend to be less specific for their targets and more specific for acetyl-lysine residues. We would therefore argue that deacetylation of a chemically acetylated enzyme does not imply physiological relevance in the same way of enzymatic acetylation.

Lastly, utilizing anti-acetyllysine antibodies to probe for acetylation of proteins from *in vivo* overexpression experiments isn't always informative, due to the fact that in stationary phase, lysyl residues of proteins act as sinks for excess Ac-CoA and Ac-P (83). If westerns are to be utilized, it is best to harvest in early or mid-log of growth, and to verify single acetylation sites via site directed mutagenesis of targeted lysyl residues. It is generally thought that a physiologically relevant acetylation event occurs on a single lysyl residue, and therefore acetylation of a protein should be abolished (via *in vitro* assays, *i.e.* radiolabeled acetyl-CoA assays or *in vivo* assays, *i.e.* Western Blots) when the single lysyl-residue is altered. While it is possible multiple acetylation events on a single protein can change an enzyme's activity, this has not been enzymatically shown.

OUTLOOK AND CONCLUDING REMARKS.

Within the field there are some outstanding questions and opportunities with regards to bacterial protein acetylation. What controls and triggers protein acetylation and deacetylation? Are there acetyl-P acetylation mechanisms that are physiologically relevant or are there acetyltransferase enzymes that prefer Ac-P over Ac-CoA? Or is non-enzymatic Ac-P acetylation utilized as a carbon-storage strategy? Regardless of mechanisms, there lies a huge gap within the field for validation of acetylation targets and what the fates of acetylation are on an enzyme. Many studies have shown that an enzyme's activity is not the only aspect to assess after acetylation, and perhaps we should be looking into the effects of acetylation on structure, protein-protein interaction, or cellular localization.

The information on protein acetylation is rapidly growing, and provides a solid foundation for further characterizing bacterial acetyltransferases. As the field progresses, it is likely that we will find many acetyltransferases that target not only proteins, but small molecules as well. Acetylation, paired with deacetylation, is a rapid way for a cell to adjust cellular metabolism or flux, and is involved in a variety of important cellular processes ranging from virulence, central metabolism, DNA-binding, or protein synthesis. Acetylation therefore warrants much attention with regards to bacterial physiology and will provide many insights into how cells respond and adapt to stimuli.

DISSERTATION OUTLINE

This thesis will focus on protein and small molecule acetylation in prokaryotes. Chapter 2 will outline construction of vectors used to facilitate *in vitro* acetylation studies. Chapter 3 will focus on characterizing phosphinothricin acetyltransferases from a variety of prokaryotes. Chapter 4 will highlight work done in *S. enterica* that characterized a toxin acetyltransferase and

describe its role in persistence. Chapter 5 will provide data on characterization of an ACT-domain acetyltransferase and its target, acetyl-CoA synthetase. Chapter 6 will investigate the physiology of *R. palustris* and through acetylation during photosynthetic growth on benzoate. Chapter 7 will focus on a *S. enterica* acetyltransferase that acetylated *O*-acetyl-serine, which lead to changes in cysteine regulon transcription. Chapter 8 will focus on further characterizing the ACT-domain acetyltransferase from Chapter 5. Chapter 9 will state a brief conclusion on all of these diverse topics. Of course, if you need to reference any of these chapters, it would be more useful to read the table of contents than this dissertation outline paragraph.

DISCLOSURE STATEMENT

The authors do not have any conflict of interest to declare.

ACKNOWLEDGMENTS

This works was supported by NIH grant R01 GM062203 to J.C.E.-S.

REFERENCES.

1. Rivera C, Gurard-Levin ZA, Almouzni G, Loyola A. 2014. Histone lysine methylation and chromatin replication. Biochim Biophys Acta 1839:1433-1439.
2. Kennelly PJ. 2014. Protein Ser/Thr/Tyr phosphorylation in the Archaea. J Biol Chem 289:9480-9487.
3. Cousin C, Derouiche A, Shi L, Pagot Y, Poncet S, Mijakovic I. 2013. Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation. FEMS Microbiol Lett 346:11-19.
4. Van Meter M, Mao Z, Gorbunova V, Seluanov A. 2011. Repairing split ends: SIRT6, mono-ADP ribosylation and DNA repair. Aging (Albany NY) 3:829-835.

5. Li C, Wang D, Lv X, Jing R, Bi B, Chen X, Guo J, Wang F, Sun S, Azadzoi KM, Yang JH. 2017. *Yersinia pestis* acetyltransferase-mediated dual acetylation at the serine and lysine residues enhances the auto-ubiquitination of ubiquitin ligase MARCH8 in human cells. Cell Cycle 16:649-659.

6. Weinert BT, Scholz C, Wagner SA, Iesmantavicius V, Su D, Daniel JA, Choudhary C. 2013. Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. Cell Rep 4:842-851.

7. Rieser E, Cordier SM, Walczak H. 2013. Linear ubiquitination: a newly discovered regulator of cell signalling. Trends Biochem Sci 38:94-102.

8. Muller MP, Albers MF, Itzen A, Hedberg C. 2014. Exploring adenylylation and phosphocholination as post-translational modifications. Chembiochem 15:19-26.

9. Gould N, Doulias PT, Tenopoulou M, Raju K, Ischiropoulos H. 2013. Regulation of protein function and signaling by reversible cysteine *S*-nitrosylation. J Biol Chem 288:26473-26479.

10. Zverina EA, Lamphear CL, Wright EN, Fierke CA. 2012. Recent advances in protein prenyltransferases: substrate identification, regulation, and disease interventions. Curr Opin Chem Biol 16:544-552.

11. Jarrell KF, Ding Y, Meyer BH, Albers SV, Kaminski L, Eichler J. 2014. N-Linked glycosylation in archaea: a structural, functional, and genetic analysis. Microbiol Mol Biol Rev 78:304-341.

12. Phillips DM. 1963. The presence of acetyl groups of histones. Biochem J 87:258-263.

13. Sterner DE, Berger SL. 2000. Acetylation of histones and transcription-related factors. Microbiol Mol Biol Rev 64:435-459.

14. Hentchel KL, Escalante-Semerena JC. 2015. Acylation of biomolecules in prokaryotes: a widespread strategy for the control of biological function and metabolic stress. Microbiol Mol Biol Rev 79:321-346.

15. Helbig AO, Gauci S, Raijmakers R, van Breukelen B, Slijper M, Mohammed S, Heck AJ. 2010. Profiling of *N*-acetylated protein termini provides in-depth insights into the *N*-terminal nature of the proteome. Mol Cell Proteomics 9:928-939.

16. Tucker AC, Escalante-Semerena JC. 2013. Acetoacetyl-CoA synthetase activity is controlled by a protein acetyltransferase with unique domain organization in *Streptomyces lividans*. Mol Microbiol 87:152-167.

17. VanDrisse CM, Escalante-Semerena JC. 2018. In *Streptomyces lividans*, acetyl-CoA synthetase activity is controlled by O-serine and N(epsilon)-lysine acetylation. MolMicrobiol 107:577-594.

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

19. Hentchel KL, Escalante-Semerena JC. 2015. In *Salmonella enterica*, the Gcn5-related acetyltransferase MddA (formerly YncA) acetylates methionine sulfoximine and methionine sulfone, blocking their toxic effects. J Bacteriol 197:314-325.

20. Cheverton AM, Gollan B, Przydacz M, Wong CT, Mylona A, Hare SA, Helaine S. 2016. A *Salmonella* toxin promotes persister formation through acetylation of tRNA. Mol Cell 63:86-96.

21. VanDrisse CM, Parks AR, Escalante-Semerena JC. 2017. A toxin involved in *Salmonella* persistence regulates its activity by acetylating its cognate antitoxin, a modification reversed by CobB sirtuin deacetylase. MBio 8.

22. Stuecker TN, Hodge KM, Escalante-Semerena JC. 2012. The missing link in coenzyme A biosynthesis: PanM (formerly YhhK), a yeast GCN5 acetyltransferase homologue triggers aspartate decarboxylase (PanD) maturation in *Salmonella enterica*. Mol Microbiol 84:608-619.

23. Fukuchi J, Kashiwagi K, Takio K, Igarashi K. 1994. Properties and structure of spermidine acetyltransferase in *Escherichia coli*. J Biol Chem 269:22581-22585.

24. Yoshikawa A, Isono S, Sheback A, Isono K. 1987. Cloning and nucleotide sequencing of the genes *rimI* and *rimJ* which encode enzymes acetylating ribosomal proteins S18 and S5 of Escherichia coli K12. Mol Gen Genet 209:481-488.

25. Tanaka S, Matsushita Y, Yoshikawa A, Isono K. 1989. Cloning and molecular characterization of the gene *rimL* which encodes an enzyme acetylating ribosomal protein L12 of *Escherichia coli* K12. Mol Gen Genet 217:289-293.

26. Leibowitz MJ, Soffer RL. 1970. Enzymatic modification of proteins. 3. Purification and properties of a leucyl, phenylalanyl transfer ribonucleic acid protein transferase from *Escherichia coli*. J Biol Chem 245:2066-2073.

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

28. Ikeuchi Y, Kitahara K, Suzuki T. 2008. The RNA acetyltransferase driven by ATP hydrolysis synthesizes N^4-acetylcytidine of tRNA anticodon. EMBO J 27:2194-2203.

29. Hung MN, Rangarajan E, Munger C, Nadeau G, Sulea T, Matte A. 2006. Crystal structure of TDP-fucosamine acetyltransferase (WecD) from *Escherichia coli*, an enzyme required for enterobacterial common antigen synthesis. J Bacteriol 188:5606-5617.

30. Rycroft JA, Gollan B, Grabe GJ, Hall A, Cheverton AM, Larrouy-Maumus G, Hare SA, Helaine S. 2018. Activity of acetyltransferase toxins involved in *Salmonella persister f*ormation during macrophage infection. Nat Commun 9:1-11.

31. Vetting MW, Carvalho LPSd, Yu M, Hegde SS, Magnet S, Roderick SL, Blanchard JS. 2005. Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys 433:212-226.

32. Gardner JG, Grundy FJ, Henkin TM, Escalante-Semerena JC. 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD^+ involvement in *Bacillus subtilis*. J Bacteriol 188:5460-5468.

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

34. Thao S, Escalante-Semerena JC. 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.

35. Tucker AC, Taylor KC, Rank KC, Rayment I, Escalante-Semerena JC. 2014. Insights into the specificity of lysine acetyltransferases. J Biol Chem 289:36249-36262.

36. Noy T, Xu H, Blanchard JS. 2014. Acetylation of acetyl-CoA synthetase from Mycobacterium tuberculosis leads to specific inactivation of the adenylation reaction. Arch Biochem Biophys 550-551:42-49.

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

38. Xu JY, You D, Leng PQ, Ye BC. 2014. Allosteric regulation of a protein acetyltransferase in *Micromonospora aurantiaca* by the amino acids cysteine and arginine. J Biol Chem 289:27034-27045.

39. Hodawadekar SC, Marmorstein R. 2007. Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. Oncogene 26:5528-5540.

40. Sauve AA, Wolberger C, Schramm VL, Boeke JD. 2006. The biochemistry of sirtuins. Annu Rev Biochem 75:435-465.

41. Tsang AW, Escalante-Semerena JC. 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. Starai VJ, Celic I, Cole RN, Boeke JD, Escalante-Semerena JC. 2002. Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science 298:2390-2392.

43. Gardner JG, Escalante-Semerena JC. 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.

44. Rack JG, Morra R, Barkauskaite E, Kraehenbuehl R, Ariza A, Qu Y, Ortmayer M, Leidecker O, Cameron DR, Matic I, Peleg AY, Leys D, Traven A, Ahel I. 2015.

Identification of a class of protein ADP-Ribosylating sirtuins in microbial pathogens. Mol Cell 59:309-320.

45. Starai VJ, Gardner JG, Escalante-Semerena JC. 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.

46. Starai VJ, Escalante-Semerena JC. 2004. Acetyl-coenzyme A synthetase (AMP forming). Cell Mol Life Sci 61:2020-2030.

47. Starai VJ, Takahashi H, Boeke JD, Escalante-Semerena JC. 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.

48. Starai VJ, Takahashi H, Boeke JD, Escalante-Semerena JC. 2003. Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae*. Genetics 163:545-55.

49. Crosby HA, Rank KC, Rayment I, Escalante-Semerena JC. 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.

50. Brinsmade SR, Escalante-Semerena JC. 2007. In vivo and in vitro analyses of single-amino acid variants of the *Salmonella enterica* phosphotransacetylase enzyme provide insights into the function of its N-terminal domain. J Biol Chem 282:12629-12640.

51. Chittori S, Savithri HS, Murthy MR. 2012. Structural and mechanistic investigations on *Salmonella typhimurium* acetate kinase (AckA): identification of a putative ligand binding pocket at the dimeric interface. BMC Struct Biol 12:24.

52. Chan CH, Garrity J, Crosby HA, Escalante-Semerena JC. 2011. In *Salmonella enterica*, the sirtuin-dependent protein acylation/deacylation system (SDPADS) maintains energy homeostasis during growth on low concentrations of acetate. Mol Microbiol 80:168-183.

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

54. Crosby HA, Rank KC, Rayment I, Escalante-Semerena JC. 2012. Structure-guided expansion of the substrate range of methylmalonyl-CoA synthetase (MatB) of *Rhodopseudomonas palustris*. Appl Environ Microbiol 78:6619-6629.

55. Han X, Shen L, Wang Q, Cen X, Wang J, Wu M, Li P, Zhao W, Zhang Y, Zhao G. 2017. Cyclic AMP Inhibits the Activity and Promotes the Acetylation of Acetyl-CoA Synthetase through Competitive Binding to the ATP/AMP Pocket. J Biol Chem 292:1374-1384.

56. Hentchel KL, Escalante-Semerena JC. 2015. Complex regulation of the sirtuin-dependent reversible lysine acetylation system of *Salmonella enterica*. Microbial Cell 2:451-453.

57. You D, Wang MM, Ye BC. 2017. Acetyl-CoA synthetases of *Saccharopolyspora erythraea* are regulated by the nitrogen response regulator GlnR at both transcriptional and post-translational levels. Mol Microbiol 103:845-859.

58. You D, Yin BC, Li ZH, Zhou Y, Yu WB, Zuo P, Ye BC. 2016. Sirtuin-dependent reversible lysine acetylation of glutamine synthetases reveals an autofeedback loop in nitrogen metabolism. Proc Natl Acad Sci U S A 113:6653-6658.

59. Amin R, Franz-Wachtel M, Tiffert Y, Heberer M, Meky M, Ahmed Y, Matthews A, Krysenko S, Jakobi M, Hinder M, Moore J, Okoniewski N, Macek B, Wohlleben W,

Bera A. 2016. Post-translational serine/threonine phosphorylation and lysine acetylation: A novel regulatory aspect of the global nitrogen response regulator GlnR in *S. coelicolor* M145. Front Mol Biosci 3:38.

60. Nakayasu ES, Burnet MC, Walukiewicz HE, Wilkins CS, Shukla AK, Brooks S, Plutz MJ, Lee BD, Schilling B, Wolfe AJ, Muller S, Kirby JR, Rao CV, Cort JR, Payne SH. 2017. Ancient Regulatory Role of Lysine Acetylation in Central Metabolism. mBio doi:10.1128/mBio.01894-17:e01894-1917.

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

62. Yu BJ, Kim JA, Moon JH, Ryu SE, Pan JG. 2008. The diversity of lysine-acetylated proteins in *Escherichia coli*. J Microbiol Biotechnol 18:1529-1536.

63. Bi J, Wang Y, Yu H, Qian X, Wang H, Liu J, Zhang X. 2017. Modulation of central carbon metabolism by acetylation of isocitrate lyase in *Mycobacterium tuberculosis*. Sci Rep 7:44826.

64. Nagano-Shoji M, Hamamoto Y, Mizuno Y, Yamada A, Kikuchi M, Shirouzu M, Umehara T, Yoshida M, Nishiyama M, Kosono S. 2017. Characterization of lysine acetylation of a phosphoenolpyruvate carboxylase involved in glutamate overproduction in Corynebacterium glutamicum. Mol Microbiol 104:677-689.

65. Mizuno Y, Nagano-Shoji M, Kubo S, Kawamura Y, Yoshida A, Kawasaki H, Nishiyama M, Yoshida M, Kosono S. 2016. Altered acetylation and succinylation profiles in *Corynebacterium glutamicum* in response to conditions inducing glutamate overproduction. Microbiologyopen 5:152-173.

66. Xu JY, Xu Z, Liu X, Tan M, Ye BC. 2018. Protein Acetylation and Butyrylation Regulate the Phenotype and Metabolic Shifts of the Endospore-forming Clostridium acetobutylicum. Mol Cell Proteomics 17:1156-1169.

67. Suka N, Suka Y, Carmen AA, Wu J, Grunstein M. 2001. Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. Mol Cell 8:473-479.

68. Thao S, Chen CS, Zhu H, Escalante-Semerena JC. 2010. N(epsilon)-Lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. PLoS One 5:e15123.

69. Ghosh S, Padmanabhan B, Anand C, Nagaraja V. 2016. Lysine acetylation of the *Mycobacterium tuberculosis* HU protein modulates its DNA binding and genome organization. Mol Microbiol 100:577-588.

70. Anand C, Garg R, Ghosh S, Nagaraja V. 2017. A Sir2 family protein Rv1151c deacetylates HU to alter its DNA binding mode in *Mycobacterium tuberculosis*. Biochem Biophys Res Commun 493:1204-1209.

71. Yang H, Sha W, Liu Z, Tang T, Liu H, Qin L, Cui Z, Chen J, Liu F, Zheng R, Huang X, Wang J, Feng Y, Ge B. 2018. Lysine acetylation of DosR regulates the hypoxia response of *Mycobacterium tuberculosis*. Emerg Microbes Infect 7:1-14.

72. Jurenas D, Chatterjee S, Konijnenberg A, Sobott F, Droogmans L, Garcia-Pino A, Van Melderen L. 2017. AtaT blocks translation initiation by N-acetylation of the initiator tRNA(fMet). Nat Chem Biol 13:640-646.

73. Qian H, Yao Q, Tai C, Deng Z, Gan J, Ou HY. 2018. Identification and characterization of acetyltransferase-type toxin-antitoxin locus in *Klebsiella pneumoniae*. Mol Microbiol doi:10.1111/mmi.13934.

74. McVicker G, Tang CM. 2016. Deletion of toxin-antitoxin systems in the evolution of *Shigella sonnei* as a host-adapted pathogen. Nat Microbiol 2:1-8.

75. Iqbal N, Guerout AM, Krin E, Le Roux F, Mazel D. 2015. Comprehensive Functional Analysis of the 18 *Vibrio cholerae* N16961 Toxin-Antitoxin Systems Substantiates Their Role in Stabilizing the Superintegron. J Bacteriol 197:2150-2159.

76. Lobato-Marquez D, Moreno-Cordoba I, Figueroa V, Diaz-Orejas R, Garcia-del Portillo F. 2015. Distinct type I and type II toxin-antitoxin modules control Salmonella lifestyle inside eukaryotic cells. Sci Rep 5:9374.

77. Yeo CC. 2018. GNAT toxins of bacterial toxin-antitoxin systems: acetylation of charged tRNAs to inhibit translation. Mol Microbiol doi:10.1111/mmi.13958.

78. Van Melderen L, Jurenas D, Garcia-Pino A. 2017. Messing up translation from the start: how AtaT inhibits translation initiation in *E. coli*. RNA Biol doi:10.1080/15476286.2017.1391439:0.

79. Jurenas D, Garcia-Pino A, Van Melderen L. 2017. Novel toxins from type II toxin-antitoxin systems with acetyltransferase activity. Plasmid 93:30-35.

80. Lobato-Marquez D, Diaz-Orejas R, Garcia-Del Portillo F. 2016. Toxin-antitoxins and bacterial virulence. FEMS Microbiol Rev 40:592-609.

81. Hall AM, Gollan B, Helaine S. 2017. Toxin-antitoxin systems: reversible toxicity. Curr Opin Microbiol 36:102-110.

82. Venkat S, Gregory C, Gan Q, Fan C. 2017. Biochemical characterization of the lysine acetylation of tyrosyl-tRNA synthetase in *Escherichia coli*. Chembiochem 18:1928-1934.

83. Schilling B, Christensen D, Davis R, Sahu AK, Hu LI, Walker-Peddakotla A, Sorensen DJ, Zemaitaitis B, Gibson BW, Wolfe AJ. 2015. Protein acetylation dynamics in response to carbon overflow in *Escherichia coli*. Mol Microbiol 98:847-863.

CHAPTER 2

NEW HIGH-CLONING-EFFICIENCY VECTORS FOR COMPLEMENTATION STUDIES AND RECOMBINANT PROTEIN OVERPRODUCTION IN *ESCHERICHIA COLI* AND *SALMONELLA ENTERICA*[1]

[1]VanDrisse C.M. and Escalante-Semerena J.C. 2016. *Plasmid.* 86:1-6.

Reprinted here with permission from the publisher.

ABSTRACT

Galloway et al. recently described a method to alter vectors to include Type IIS restriction enzymes for high efficiency cloning. Utilizing this method, the multiple cloning sites of complementation and overexpression vectors commonly used in our laboratory were altered to contain recognition sequences of the Type IIS restriction enzyme, BspQI. Use of this enzyme increased the rate of cloning success to >97% efficiency. L(+)-Arabinose-inducible complementation vectors and overexpression vectors encoding *N*-terminal recombinant tobacco etch virus protease (rTEV)-cleavable H_6-tags were altered to contain BspQI sites that allowed for cloning into all vectors using identical primer overhangs. Additionally, a vector used for directing the synthesis of proteins with a *C*-terminal, rTEV-cleavable H_6-tag was engineered to contain BspQI sites, albeit with different overhangs from that of the previously mentioned vectors. Here we apply a method used to engineer cloning vectors to contain BspQI sites and the use of each vector in either *in vivo* complementation studies or *in vitro* protein purifications.

INTRODUCTION

Increasing cloning efficiency is desirable for the rapid completion of experiments where complementation of function or protein overproduction is the objective. Engler et al. developed a method termed Golden Gate cloning; in which a Type IIS restriction enzyme (BsaI) is utilized for high efficiency cloning into expression constructs (1). Additionally, Galloway et al. (2) reported an approach for the introduction of BspQI sites into a vector's multiple-cloning site (MCS). Type IIS restriction enzymes (*e.g.,* BspQI, BsaI, BpiI) have been used in cloning methods for the construction of gene expression reporters (3), recombinant protein expression (1, 2), assembly of multiple DNA fragments (4) or gene fusions and promoter shuffling (5).

Type IIS restriction enzymes cleave at least one strand of DNA located outside of the recognition sequence, with BspQI cleaving both strands of DNA (6). Type IIS restriction enzymes act as dimers, with one domain binding to a 4-7 base pair recognition sequence and another domain interacting with a cleavage site 1-20 nucleotides away from the recognition site (7, 8). Having a separate cut site from the recognition sequence is valuable for vector design because the cleavage site can be engineered to have its own unique overhangs. An enzyme, such as BspQI, has a recognition sequence [GCTCTTC (1/4)] where the first number in parentheses corresponds to the position of cleavage on the coding strand and the second number to the cleavage site on the complementary strand (6). This creates a three base pair overhang, resulting in 81 different possible overhangs (Fig. 2.1).

The method described by Galloway et al. simplifies cloning efforts because digestion and ligation occur in the same reaction mixture, substantially reducing the time needed to clone genes of interest (2). Additionally, because Type IIS restriction enzymes cut outside of their recognition sequence, ligated insert-vector products no longer contain the recognition site, and any empty vector remaining in the reaction mixture can be linearized by addition of the Type IIS restriction enzyme being used, substantially reducing false-positive background (1).

In this study, we used the method described by Galloway *et al.* to introduce BspQI sites into the MCSs of plasmids used for arabinose-inducible expression of genes of interest, or for the overproduction of proteins. Here, we modified the MCSs of three complementation vectors (pBAD24, pBAD30, and pBAD33-SD1 (9)). Specifically, each MCS was modified to contain two BspQI sites. We also modified two overexpression vectors to contain BspQI sites matching those present in the complementation vectors. The resulting vectors directed the synthesis of proteins of interest fused to an *N*-terminal, recombinant tobacco etch virus (rTEV)-cleavable hexahistidine

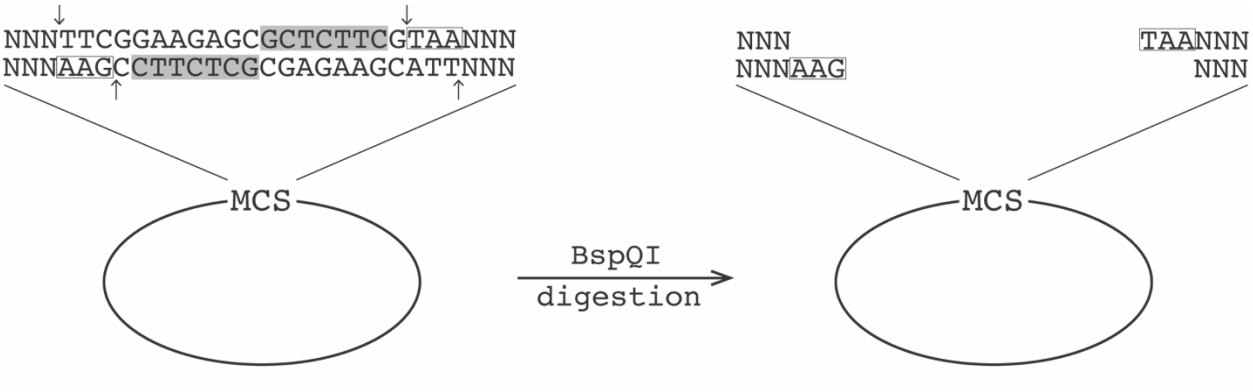


Figure 2.1. Visual representation of BspQI vector digestion. BspQI binds to the recognition site [GCTCTTC (1/4), highlighted in grey] and will cut both strands at indicated arrows. This creates a three base pair overhang. The cut site does not interfere with recognition and can be custom designed. Shown is an example of vectors constructed in this study (pCV1-3, pTEV18, 19).

(H_6-tag (pTEV5, (10)) or an *N*-terminal, rTEV-cleavable H_6, maltose binding protein tag (pTEV6, (10)). Conveniently, cloning into all of the above-mentioned vectors was done using the same primer overhangs. Lastly, an overexpression vector with a pTEV backbone was also altered to contain BspQI sites. In this case, however, the protein whose synthesis was directed by this vector was fused to a *C*-terminal, rTEV-cleavable H_6-tag. Cloning into all of the BspQI-containing vectors was $\geq$97% efficient. The main goal of this work is to provide investigators with information regarding the specific aforementioned vectors.

MATERIALS AND METHODS

Bacterial strains, culture media, chemicals, and sequencing methods.

All strains used in this study are listed in Table 2.1. *Escherichia coli* C41 (λDE3) (11) and DH5α (Biolabs) strains were grown at 37°C in lysogeny broth (LB, Difco). Strains used for growth analysis were derivatives of *Salmonella enterica* sv Typhimurium LT2 and grown at 37°C in nutrient broth (NB, Difco) containing NaCl (85 mM), or no-carbon essential (NCE) minimal medium (12) supplemented with sodium acetate (10 mM), L-methionine (0.5 mM), $MgSO_4$ (1 mM), and trace minerals (13). Antibiotics were used at the following concentrations: ampicillin, 100 μg mL^{-1}; chloramphenicol, 20 μg mL^{1}. All chemicals were purchased from Fischer unless noted otherwise; chloramphenicol, L(+)-arabinose (Sigma-Aldrich); ethylenediaminetetraacetic acid (EDTA, VWR); and isopropyl β -D-1-thiogalactopyranoside (IPTG, IBI Scientific). All restriction enzymes were purchased from Thermo Scientific™ with the exception of BspQI (New England Biolabs). DNA sequencing was performed using Big Dye® Terminator v3.1 protocols (Applied Biosystems). DNA sequencing was performed at the Georgia Genomics Facility.

Table 2.1. Bacterial Strains and plasmids used in this study[a]		
Strain or Plasmid	**Description**	**Source/reference**
E. coli strains		
E. coli DH5α	Φ80d*lacZ*ΔM15 *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 (r_k^-, m_k^+) *supE*44 *relA*1 *deoR* Δ(*lacZYA-argF)* U169 *phoA*	NEB
E. coli C41 (λDE3)	*pka12*::*kan*$^+$ *ompT hsdS* (r_Bm_B) *gal λ (DE3)*	Laboratory collection
S. enterica strains		
JE6583	*metE205 ara-9*	K. Sanderson via J. Roth
Derivatives of JE6583		
JE7758	Δ*acs2*	Laboratory collection
JE22070	pCV1	
JE22071	pCV2	
JE22072	pCV3	
JE22073	Δ*acs2* / pCV1	
JE22074	Δ*acs2* / pCV2	
JE22075	Δ*acs2* / pCV3	
JE22085	Δ*acs2* / pACS62	
JE22086	Δ*acs2* / pACS63	
JE22087	Δ*acs2* / pACS64	

[a] Unless otherwise stated, all plasmids were constructed during the course of this work.

Construction of vectors containing BspQI sites

Native BspQI sites (GCTCTTC) in the backbone of pBAD24, pBAD30, pBAD33-SD1, pTEV5, and pTEV6 were eliminated via site-directed mutagenesis and new BspQI sites were introduced into the MCS as described previously (2). Native BspQI sites (GCTCTTC) in the backbone of pBAD24, pBAD30, pBAD33-SD1, pTEV5, and pTEV6 were mutated (GCTCTTC → WCTWTTC) by site-directed mutagenesis using the QuikChange site-directed mutagenesis protocol (Stratagene). Polymerase chain reaction was performed using *Pfu* Ultra II DNA polymerase (Agilent Technologies) as per manufacturer's instructions, using primers listed in Table 2.2. Modifications included an anneal time of 60 s, an extension temperature of 68°C, and an extension time of 2.5 min kb^{-1}. All vectors contained one native BspQI site with the exception of plasmid pBAD24, which contained two sites that were mutated sequentially. PCR products were digested with DpnI (Fermentas) for 1 h and transformed into *E. coli* strain DH5α chemically competent cells (14). Plasmids were isolated from antibiotic–resistant cells using Wizard® Plus SV Miniprep DNA Purification System (Promega). Disruption of BspQI sites was determined by restriction analysis with Thermos Scientific™ Fast Digest™ enzymes using BspQI and a second restriction enzyme that cut the plasmid backbone. For example, plasmids pTEV5 and pTEV6 were cut with NotI; plasmids pBAD24 and pBAD30 were cut with EcoRV; and plasmid pBAD33 was cut with HindIII. Products were analyzed electrophoretically on a 1% agarose gel with Tris/Borate/EDTA (TBE) buffer for 40 min at 115 V. BspQI sites were engineered into the multiple cloning site (MCS) using *Pfu* Ultra II DNA polymerase with primers that extended out from the MCS with BspQI overhangs and 5' phosphates (Table 2.2). PCR products were purified using Wizard® SV DNA Clean-Up System (Promega). Linear blunt-end fragments were ligated using the Fast-Link-DNA Ligation Kit (Epicentre) as per

Table 2.2. Primers used in this study.	
Primer Name	**Primer Sequence 5' → 3'**
Eliminating BspQI from plasmid backbone	
5' pBAD BspQI knockout	GCGTCTTTTACTGGCWCWTCTCGCTAACCAAACC [a]
3' pBAD BspQI knockout	GGTTTGGTTAGCGAGAWGWGCCAGTAAAAGACGC [a]
5' pBAD24 BspQI knockout2	TGAGCGAGGAAGCGGAWGWCGCCTGATGCGGTAT [a]
3' pBAD24 BspQI knockout2	ATACCGCATCAGGCGCWCWTCCGCTTCCTCGCTCA [a]
5' pTEV BspQI knockout	AGCGAGGAAGCGGAWGWGCGCCTGATGCGGTA [a]
3' pTEV BspQI knockout	TACCGCATCAGGCGCWCWTCCGCTTCCTCGCT [a]
Engineering BspQI into MCS	
5' BspQI MCS into pBAD	/5Phos/GCTCTTCNTAAAAGCTTGGCTGTTTTGGCGGATGAGAGAAG
3' BspQI MCS into pBAD24	/5Phos/GCTCTTCNGAATTCCTCCTGCTAGCCCAAAAAAACGGG
3' BspQI MCS into pBAD30	/5Phos/GCTCTTCNGAATTCGCTAGCCCAAAAAAACGGGTATGG
3' BspQI MCS into pBAD33	/5Phos/GCTCTTCNGAATTCGCTAGCCCTCCTCCAAAAAAACGG
5' BspQI MCS into pTEV5	/5Phos/GCTCTTCNTAACCTGCAGGCAAGCTTGCGGCCGCACTCGAG
3' BspQI MCS into pTEV5	/5Phos/ GCTCTTCNGGTACCTTGGAAGTAGAGATTCTCTGACGTGG
5' BspQI MCS into pTEV6	/5Phos/GCTCTTCNTAAGAATTCTCGAGCTCCCGGGATCCGCGGCCGCT
3' BspQI MCS into pTEV6	/5Phos/ GCTCTTCNGCTAGCGCCCTGAAAATACAGGTTTTCAC
5' BspQI MCS and c-term His into pTEV	/5Phos/GCTCTTCNGCGGCCGCTGAGCAATAACTAGCAGAAAACCTGTATTTTCAGGGCCATCACCATCACCATCACTGA
3' BspQI MCS and c-term His into pTEV	/5Phos/GCTCTTCNGTAGTACGACATATATGTATATCTCCTTCTTAAA
Site-Directed Mutagenesis	
5' pTEV16 SDM	AGAGCGCTCTTCGGAAAGCGCCCTGAAAATACAGG
3' pTEV16 SDM	CCTGTATTTTCAGGGCGCTTTCCGAAGAGCGCTCT
5' pTEV17 SDM	AAGGAAGAGCGCTCTTCTGAAACCTTGGAAGTAGAGATTC
3' pTEV17 SDM	GAATCTCTACTTCCAAGGTTTCAGAAGAGCGCTCTTCCTT
Cloning [b]	
5' Acs pTEV16	NNGCTCTTCN**AGC**ATGAGCCAAACACATAAACACGCC
5' Acs pTEV17	NNGCTCTTCN**ACC**ATGAGCCAAACACATAAACACGCC
5' Acs pCV1_3_pTEV18_19	NNGCTCTTCN**TTC**ATGAGCCAAACACATAAACACGCC
5' Acs pCV2_RBS	NNGCTCTTCN**TTC**CCTACAAGGAGAACAACAGCATGAGC
5' Acs pTEV20	NNGCTCTTCN**TAC**ATGAGCCAAACACATAAACACGCC
3' Acs pCV1-3 pTEV16-19	NNGCTCTTCN**TTA**TTATGACGGCATCGCGATGGCCT
3' Acs pTEV20	NNGCTCTTCN**TTC**TTATGACGGCATCGCGATGGCCT
Checking primers	
T7 Forward	TAATACGACTCACTATAGGG
T7 Reverse	GCTAGTTATTGCTCAGCGG
pBAD sequencing primer 1	CTGTTTCTCCATACCCGTT
pBAD sequencing primer 2	GGCTGAAAATCTTCTCT

[a] W= A or T

[b]Cloning primers, 3 base overhang in bold is generated after BspQI digest of PCR product

manufacturer's instructions, and transformed into *E. coli* strain DH5α competent cells. Cells were plated on LB + ampicillin or chloramphenicol and plasmids were isolated from an overnight culture grown in LB containing antibiotic; plasmids were isolated as indicated above. Introduction of BspQI sites into the MCS was confirmed by digestion with BspQI and an enzyme outside of the MCS [*i.e.,* pTEV5, pTEV6 (ScaI); pBAD24 (PstI), pBAD30 (EcoRV); pBAD33 (NcoI)], and by DNA sequencing.

Newly engineered vectors are listed in Table 2.3 and are hereafter referred to as pTEV16 (derived from pTEV5), pTEV17 (derived from pTEV6), and pCV1 (for complementation vector, derived from pBAD24), pCV2 (derived from pBAD30), and pCV3 (derived pBAD33-SD1).

Site-Directed mutagenesis of pTEV16 and pTEV17

To reduce the number of primers and to facilitate cloning into overexpression and complementation vectors, DNA containing the cut site (pTEV16: AGC, pTEV17: ACC) was mutated to the pCV cut site (*i.e.,* TTC) by site-directed mutagenesis. Polymerase chain reaction was performed using *Pfu* Ultra II DNA polymerase using primers listed in Table 2.2. Modifications included an anneal time of 60 s, an extension temperature of 68 °C, and an extension time of 2.5 min kb^{-1}. DNA changes were confirmed by sequencing.

Construction of *C*-terminal rTEV cleavable H_6-tag BspQI vector

The pTEV18 vector was amplified with *Pfu* Ultra II DNA polymerase using primers outside of the *N*-terminal H_6-tag and MCS (Table 2.2). PCR product was purified using Wizard® SV DNA Clean-Up System (Promega). Linear blunt-end fragments were ligated using the Fast-Link-DNA Ligation Kit (Epicentre), and transformed into DH5α chemically competent cells (14). Cells were plated on LB + ampicillin, incubated overnight at 37ºC, and 10 individual Ap^R colonies were used to inoculate fresh LB + ampicillin medium. Plasmids were isolated from each overnight

Table 2.3. Plasmids used in this study[a]			
Plasmid	**Genotype**	**Description**	**Source**
pBAD24	*araC*$^+$ *bla*$^+$	Arabinose[b] inducible complementation vector with pBR322 *ori*	(9)
pBAD30	*araC*$^+$ *bla*$^+$	Arabinose inducible complementation vector with p15A *ori*	(9)
pBAD33-SD1	*araC*$^+$ *cat*$^+$	Arabinose inducible complementation vector with p15A *ori*	(9), Laboratory collection
pTEV5	*lacI*$^+$ *bla*$^+$	IPTG[c] inducible *N*-terminal rTEV[d] cleavable H_6-tag	(10)
pTEV6	*lacI*$^+$ *malE*$^+$ *bla*$^+$	IPTG inducible *N*-terminal rTEV cleavable H_6-tag-MBP fusion	(10)
pCV1	*araC*$^+$ *bla*$^+$	pBAD24 with BspQI MCS[e]	
pCV2	*araC*$^+$ *bla*$^+$	pBAD30 with BspQI MCS	
pCV3	*araC*$^+$ *cat*$^+$	pBAD33-SD1 with BspQI MCS	
pTEV16	*lacI*$^+$ *bla*$^+$	pTEV5 with BspQI MCS	
pTEV17	*lacI*$^+$ *malE*$^+$ *bla*$^+$	pTEV6 with BspQI MCS	
pTEV18	*lacI*$^+$ *bla*$^+$	pTEV16 with altered BspQI digested overhang	
pTEV19	*lacI*$^+$ *malE*$^+$ *bla*$^+$	pTEV17 with altered BspQI digested overhang	
pTEV20	*lacI*$^+$ *bla*$^+$	pTEV18 backbone with BspQI MCS and *C*-terminal rTEV cleavable H_6-tag	
pACS62	*acs*$^+$ *bla*$^+$	*S.e. acs*$^+$ in pCV1	
pACS63	*acs*$^+$ *bla*$^+$	*S.e. acs*$^+$ in pCV2	
pACS64	*acs*$^+$ *cat*$^+$	*S.e. acs*$^+$ in pCV3	
pACS65	*acs*$^+$ *bla*$^+$	*S.e. acs*$^+$ in pTEV16	
pACS66	*acs*$^+$ *bla*$^+$	*S.e. acs*$^+$ in pTEV17	
pACS67	*acs*$^+$ *bla*$^+$	*S.e. acs*$^+$ in pTEV18	
pACS68	*acs*$^+$ *bla*$^+$	*S.e. acs*$^+$ in pTEV19	
pACS69	*acs*$^+$ *bla*$^+$	*S.e. acs*$^+$ in pTEV20	

[a] Unless otherwise stated, all plasmids were constructed during the course of this work
[b] L(+)-arabinose was used as inducer
[c] IPTG, isopropyl β-D-1-thiogalactopyranoside
[d] rTEV, recombinant tobacco etch virus protease
[e] MCS, multiple cloning site

culture using the Wizard® Plus SV Miniprep DNA Purification System (Promega). The presence of BspQI sites in the MCS was determined by cutting with BspQI and an enzyme outside of the MCS (*e.g.,* ScaI). The resulting MCS was sequenced to confirm the presence of the BspQI site and *C*-terminal H_6 rTEV cleavable tag.

Cloning of *Se*Acs into newly designed BspQI vectors

To illustrate the efficiency of the new vectors, the *S. enterica acs*$^+$ gene encoding acetyl-CoA synthetase was cloned. For this purpose, primers were designed with BspQI sites on the 5'-ends such that when amplicons were cut, a three base pair overhang was complementary to overhangs in the corresponding digested vector (Table 2.2). The *acs*$^+$ gene was PCR amplified from *S. enterica* LT2 genomic DNA using *Pfu* Ultra II DNA polymerase. PCR products corresponding to the correct size were verified via gel electrophoresis and extracted as described above. Gene products were cloned into pCV1, pCV2, pCV3 pTEV16, pTEV17, pTEV18, pTEV19 and pTEV20 as described elsewhere (2).

Primers were designed with BspQI sites on the 5'-ends such that when amplicons were cut, a three base pair overhang was complementary to overhangs in the corresponding digested vector (Table 2.2). The *acs*$^+$ gene was PCR amplified from *S. enterica* LT2 genomic DNA using *Pfu* Ultra II DNA polymerase and PCR products corresponding were verified via gel electrophoresis and gel extracted (Promega). In a 1.5-mL microfuge tube the following reagents were combined: plasmid containing BspQI sites in the MCS (1 μL, 100 ng μL^{-1}), cleaned PCR product (2 μL, ~100-200 ng DNA), 10X Fast-Link ligase buffer [3 μL, Tris-acetate buffer (0.3 M, pH 7.5) containing potassium acetate (0.66 M), magnesium acetate (0.1 M), and DTT (0.005 M)], BspQI (1 μL, 10 U μL^{-1}, NEB), and nuclease-free H_2O (16 μL, Promega). It was necessary to utilize BspQI purchased from NEB due to its active temperature being at 50°C. This allows for the initial digestion to occur at 50°C,

the ligation at 25°C, and the final back-cutting digestion again at 50°C, all within the same tube. Therefore, the reaction was initially incubated at 50°C for 1 h. After incubation, the reaction was cooled to room temperature and 11.5 μL was removed from the tube and ATP (1.5 μL, 10 mM), Fast-Link ligase (1 μL, 2 U μL^{-1}) and BspQI (1 μL, 10 U μL^{-1}) were added. The reaction was incubated at room temperature for 30 min to allow for ligation, followed by a 1 h incubation period at 50°C to cut any re-ligated empty vector. A representation of each incubation step is visualized in Figure 2.2.

A sample of the ligation reaction (1 μL) was used to transform *E. coli* DH5α competent cells (14), and cells were plated on LB containing the appropriate antibiotic. The presence of inserts in the cloning vectors was confirmed by colony PCR with Go Taq® Green Master Mix (Promega) using primers that annealed to the plasmid outside of the MCS (Table 2.2). PCR products were analyzed on a 1% agarose gel with Tris/Borate/EDTA buffer for 40 min at 115V. Cells harboring plasmids containing the expected inserts were grown overnight in LB supplemented with the appropriate antibiotic, and plasmids were isolated as described above. The presence of *acs*$^+$ in each plasmid was confirmed by DNA sequencing.

Growth studies

Complementation vectors carrying the *acs*$^+$ allele or empty vector were electroporated into *acs*$^+$ and Δ*acs* strains (15) (Table 2.1). Starter cultures of each strain were grown overnight at 37°C with shaking in NB containing the appropriate antibiotic. Fresh minimal medium (198 μl) containing sodium acetate (10 mM), appropriate antibiotics, and L(+)arabinose (250 μM) was dispensed into each well of a 96-well microtiter dish. Each well was inoculated with 2 μl (1% v/v) of the aforementioned overnight cultures. Arabinose was included in the medium to induce expression of *acs*$^+$ from the P_{araBAD} promoter. Growth was monitored at 630 nm with shaking at

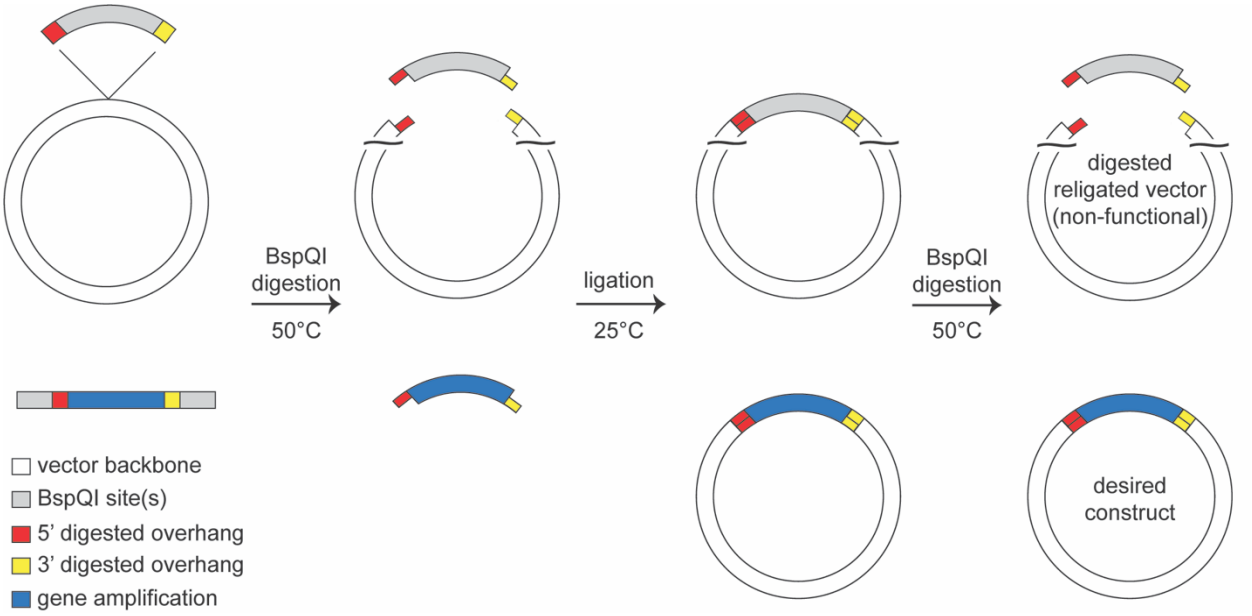


Figure 2.2. Visual representation of BspQI cloning. The initial reaction contains undigested vector, amplified gene product, buffer, and NEB BspQI (panel 1). After the first BspQI digestion, the MCS containing BspQI sites (grey) is cut out of the vector and the amplified gene product (blue) comprises of overhangs that are complementary to the vector's digested overhangs (red and yellow, panel 2). Gene product overhangs are obtained via primer design with additional sequences added to the 5' of each primer, as listed in Table 2.2. To the same tube, ATP and ligase are added at room temperature to allow for ligation. This tube will contain a mixed population of re-ligated empty vector, and desired constructs (panel 3). The final 50°C digestion will linearize empty vector, increasing the presence of the desired product. Schematic above the curvy lines are not at scale.

37°C for 24 h (BioTek ELx808-1 Ultra microplate reader). Three technical and three biological replicates were analyzed; a representative growth curve is shown. Data were analyzed using Prism v6 software (GraphPad) and error bars represent standard deviation.

Purification of *Se*Acs from pTEV16-20

Plasmids encoding H_6-*Se*Acs (pACS65 and pACS67), H_6-MBP-tagged *Se*Acs (pACS66 and pACS68), and *Se*Acs-H_6 (pACS69) were electroporated using a protocol described elsewhere (16) into *Escherichia coli* strain C41 (λDE3) (11) Δ*pat* (strain JE9314). Cultures of cells containing plasmids were grown to stationary phase (OD_{650} ~1.3) and sub-cultured (1:100 v/v) into 1 L of LB + ampicillin. Cultures were grown shaking at 37°C to an OD_{650} of 0.5, after which ectopic gene expression was induced with IPTG (0.5 mM). Cultures were grown overnight at 25°C, cells were harvested by centrifugation at 6,000 x *g* for 15 min at 4°C, and cell pellets were stored at -80°C until used . Acs proteins were purified from cell pellets as described (17). The amount of Acs in fractions was quantified on a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific) using the molecular weight (72.15 kDa) and extinction coefficient (138,770 M^{-1} cm^{-1}; ExPASy ProtParam) of the protein. Percent purity was calculated using ImageQuant™ TL.

RESULTS

Cloning efficiencies of newly designed vectors

The *S. enterica acs*$^+$ gene was cloned into each vector (Materials and Methods section 2.6) to test complementation (pCV1-3) and overexpression (pTEV16-20) using primers listed in Table 2.2. For each vector, 20 colonies were screened using colony PCR and cloning efficiencies were calculated (Table 2.4). A 2.2-kb band was indicative of a positive result, as shown in Figure 2.3.

Table 2.4. Cloning efficiencies of newly designed vectors			
Vector	**GenBank Accession #**	**No. of positive clones (20 screened)[a]**	**Efficiency (%)[b]**
pCV1	KU974153	19	95
pCV2	KU974154	20	100
pCV3	KU974155	19	95
pTEV16	KU974156	20	100
pTEV17	KU974157	20	100
pTEV18	KU974158	18	90
pTEV19	KU974159	19	95
pTEV20	KU974160	20	100

[a] Colony PCR was used to screen 20 colonies for each vector
[b] Cloning efficiencies were calculated as described under Materials & Methods section 2.6. A positive result was indicated by a 2.2 kb-band as shown in Figure 2.3.

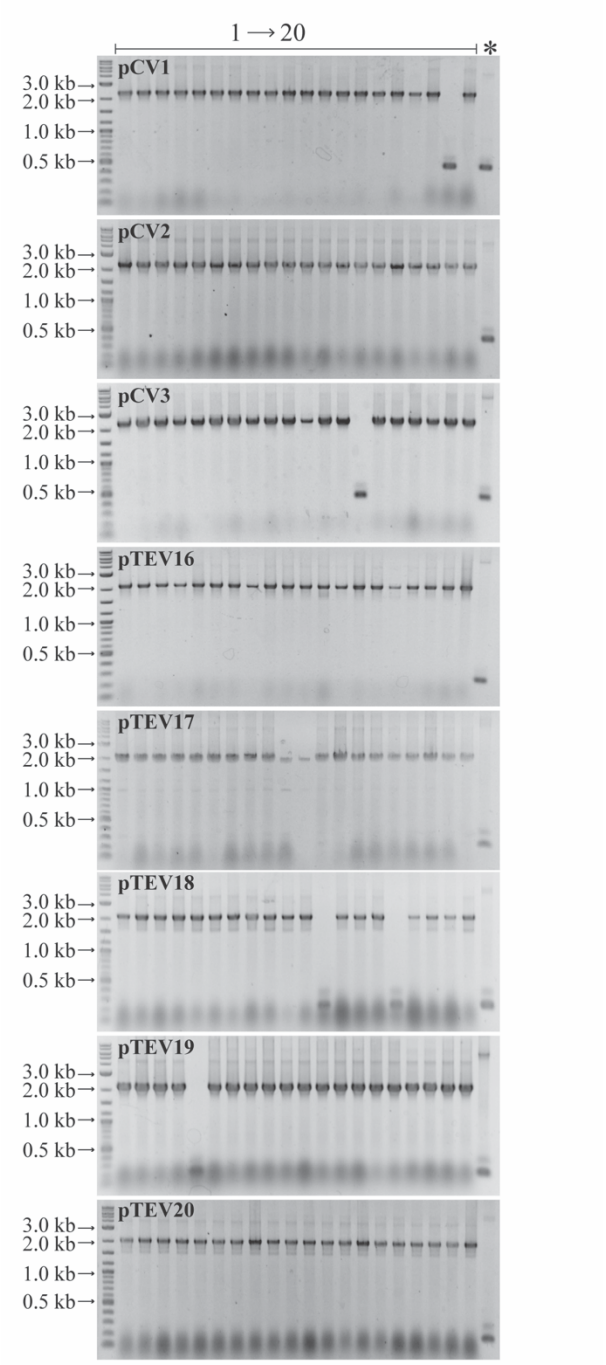


Figure 2.3. Gel electrophoresis of PCR amplicons. An agarose gel (1%, w/v) was developed to verify the presence of *S.e. acs*$^{+}$ into pCV1, pCV2, pCV3, pTEV16, pTEV17, pTEV18, pTEV19, and pTEV20. A positive result was indicated by a 2.2 kb band, with an empty vector amplification designated by a 250-500-bp band. Each vector was screened for a total of 20 replicates from the same plate (1-20) with an empty vector (EV) control. NEB Quick Load® 2-log DNA ladder was used as a reference to visualize the mobility of markers of known molecular mass.

The cloning efficiency varied between 95-100%. Information on troubleshooting cloning issues can be found in Table 2.6 of the Conclusions section.

Description of complementation vectors (pCVs)

Plasmids pCV1, pCV2, and pCV3 contained the promoter region of the *E. coli araBAD* operon (P_{araBAD}) for the purpose of inducing expression of genes of interest by the addition of L(+)-arabinose (9, 18). All vectors contained an *rrnB* terminator, and either a bla^+ gene for ampicillin resistance (pCV1, pCV2) or a cat^+ gene for chloramphenicol resistance (pCV3) (9). Plasmid pCV1 contained a pBR322 origin of replication, while pCV2 and pCV3 contain a p15A origin of replication, making pCV1 compatible *in vivo* with either pCV2 or pCV3. All plasmids contained M13 intergenic regions for ssDNA packaging into phage capsids. The MCS and plasmid maps of each complementation vector are shown in Figure 2.4. Noteworthy is the presence of the Shine Dalgarno (SD) sequence in plasmids pCV1 and pCV3. Plasmid pCV2 did not contain a SD sequence, allowing for the cloning of a gene and its native ribosome-binding site. A stop codon was added after the second BspQI site and when combined with a gene's native stop codon, the error of read through was reduced. NheI, EcoRI and HindIII restriction sites surrounded the BspQI sites for verification of inserts by restriction analysis. It was necessary to add BspQI recognition and cut sites to both forward and reverse primers. This allowed for cutting of amplified gene products so they could anneal to the corresponding digested overhangs for each vector. In the case of pCV plasmids, the primer overhangs were identical for all three vectors. A simplified list of overhangs to be added to primers for each vector can be found in Table 2.5.

Use of complementation vectors for growth behavior analysis

S. enterica Δ*acs* strains grow poorly on 10 mM acetate as the sole source of carbon and energy, and such growth defect can be corrected by the ectopic expression of the acs^+ allele (19).

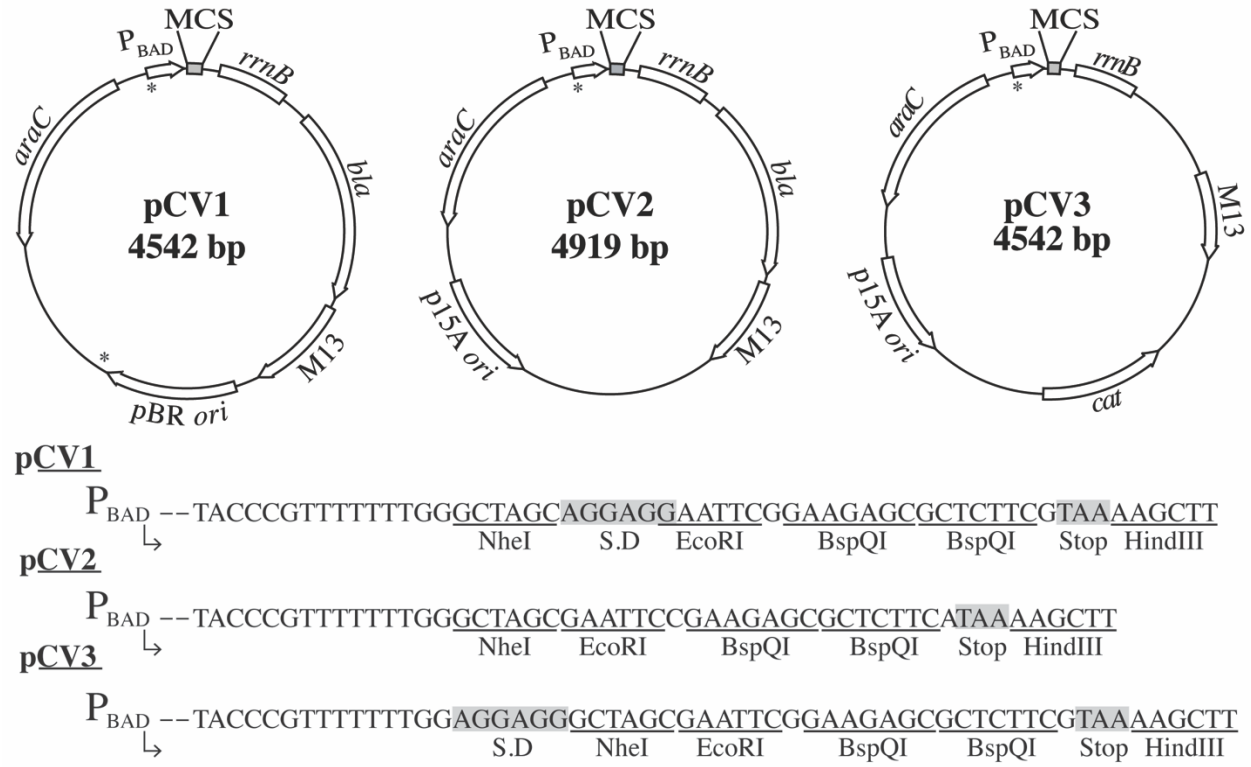


Figure 2. VanDrisse and Escalante-Semerena , 2 column fitting

Figure 2.4. Plasmid maps and MCSs of complementation vectors. Plasmid maps show notable genetic features. Plasmid MCS sequences of each vector are shown between NheI and HindIII restriction sites. S.D. in MCSs represents Shine-Dalgarno (ribosome-binding site) sequence. Asterisks identify location of BspQI sites in original vectors that were mutated from GCTCTTC to WCTWTTC, as explained in Materials and Methods.

Table 2.5. Designing primers for vectors used in this study	
Plasmid	**Sequence to add to 5' of primer**
pTEV16 forward primer	5' NNGCTCTTCNAGC
pTEV17 forward primer	5' NNGCTCTTCNACC
pCV1, pCV2, pCV3, pTEV18, pTEV19 forward primer	5' NNGCTCTTCNTTC
pCV1, pCV2, pCV3, pTEV16, pTEV17, pTEV18, pTEV19 reverse primer	5' NNGCTCTTCNTAA
pTEV20 forward primer	5' NNGCTCTTCNTAC
pTEV20 reverse primer	5' NNGCTCTTCNTTC

We used the above-mentioned phenotype of *S. enterica acs* strains to verify the functionality of the *acs*$^+$ allele in the newly described complementation vectors as described in the Materials and Methods section 2.7. Plasmids pCV1, pCV2, or pCV3 encoding *S.e. acs*$^+$ (pACS62-pACS64) were induced with L(+)-arabinose (250 μM), a concentration of inducer that effectively compensated for the absence of the Acs protein in a *S. enterica* Δ*acs* strain (Figure 2.5). This result indicated that introduction of BspQI sites into the MCS did not disrupt the intended function of these plasmids.

Overexpression vectors containing rTEV protease cleaving sites

All of the TEV vectors described herein contained the *E. coli lacI*$^+$ allele, whose protein product acts as a repressor to the T_7 promoter of the plasmids. Addition of IPTG had the dual effect of relieving LacI repression and inducing transcription of genome-encoded T_7 RNA polymerase. All vectors included a *bla*$^+$ gene for the synthesis of β-lactamase, which provided resistance to ampicillin, an f1 origin for ssDNA packaging into phage capsids, and a pBR322 origin of replication (20, 21). The plasmid maps and MCSs of plasmids pTEV16-20 are shown in Figure 2.6.

Plasmids pTEV16, pTEV18

Genes expressed from plasmids pTEV16 and pTEV18 resulted in proteins with an *N*-terminal, rTEV-cleavable H_6-tag. Plasmid pTEV18 was constructed by mutagenizing the NheI site of plasmid pTEV16 to GCTTTC, rendering this site inactive, but allowed for the design of one set of primers that could be used for cloning into the complementation and overexpression vectors. Tag removal resulted in proteins with three additional residues (*i.e.,* Gly-Ala-Ser for pTEV16; or Gly-Ala-Phe for pTEV18) on the *N*-terminus. A stop codon was engineered immediately after the BspQI sites to reduce read through when combined with the stop codon of the gene of interest.

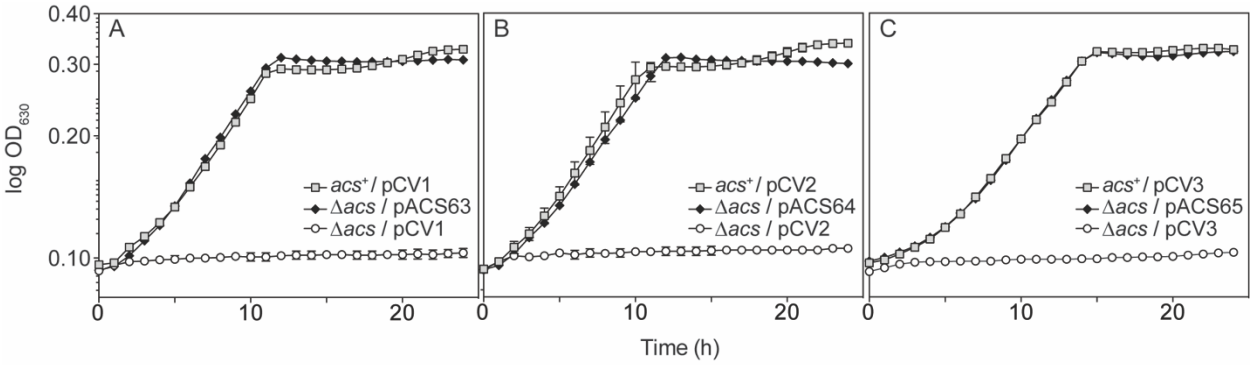


Figure 2.5. Growth analysis and complementation of *S. enterica* Δ*acs* phenotype. *S.e. acs*$^+$ was cloned into pCV1-pCV3 and transformed into a Δ*acs* strain to assess the effectiveness of the vectors. Cells were grown in NCE medium with acetate (10 mM) and transcription of *acs*$^+$ in each vector was induced with L(+)-arabinose (250 μM). Growth curves were obtained using a microplate reader (BioTek Instruments).

Plasmids pTEV16 and pTEV18 also contained multiple restriction sites in the MCS for alternative cloning methods or verification via restriction analysis. An XbaI site in plasmid pTEV18 is present upstream of the H_6-tag, if needed for insert verification by restriction analysis.

Plasmids pTEV17, pTEV19

Genes expressed from plasmids pTEV17 and pTEV19 directed the synthesis of proteins with an *N*-terminal rTEV-cleavable H_6-tag fused to a maltose binding protein (MBP). The MBP-tag was included with the goal of increasing protein solubility. Plasmid pTEV19 was constructed by mutagenizing the KpnI site of pTEV17 to GGTTTC, rendering this site inactive but allowing for the design of primers that could be used with complementation and overexpression vectors in this study. A stop codon was also engineered in these vectors to reduce read through. Tag removal resulted in proteins with two additional residues (*i.e.,* Gly-Thr for pTEV17; or Gly-Phe for pTEV19). Additional restriction sites different from those found in plasmids pTEV16 and pTEV18 were present in plasmids pTEV17 and pTEV19 for alternative cloning options or verification via restriction analysis. An XbaI site in pTEV19 lies upstream of the H_6-tag, if needed for insert verification by restriction analysis.

Vectors engineered for the synthesis of C-terminally tagged proteins

Plasmid pTEV20 was also engineered for the overproduction of proteins. However, pTEV20 differed from the aforementioned overexpression vectors by the location of the rTEV-cleavage site and the primer overhangs. Proteins overproduced in cells harboring plasmid pTEV20 have a *C*-terminal H_6 -tag that can be removed by rTEV protease. Plasmid pTEV20 was constructed by amplifying plasmid pTEV18 outside of the H_6-tag and MCS; consequently pTEV20 maintained the backbone features of plasmids pTEV16 and pTEV18. Differences in the MCS of pTEV20 can be seen in Figure 2.6. Designed primer overhangs also differed for this vector due to

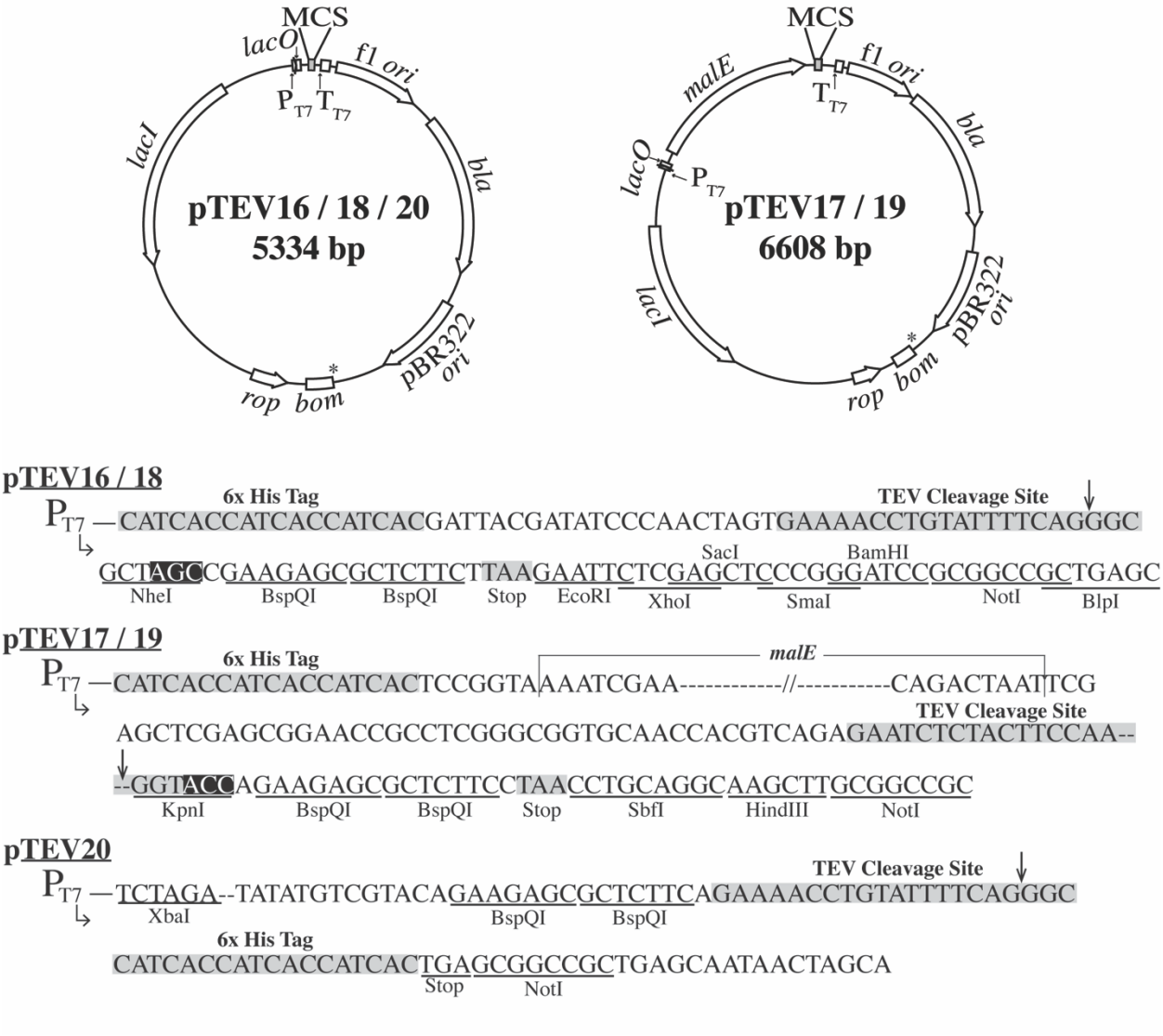


Figure 2.6. Plasmid maps and MCSs of overexpression vectors. Plasmid maps show notable genetic features. Plasmid MCS sequences of each vector are shown. Nucleotides inside black boxes of pTEV16 were changed to TTC resulting in pTEV18 and nucleotides in black boxes in pTEV17 were changed to TTC resulting in pTEV19. Therefore pTEV18 lacks an NheI site and pTEV19 lacks a KpnI site. An XbaI site lies upstream of the H_6-tag. Plasmid pTEV20 was constructed by amplifying plasmid pTEV18 outside of the H_6-tag and NotI restriction site, resulting in an rTEV-cleavable *C*-terminal H_6-tag. Arrows mark the rTEV cleavage site. Asterisks identify location of BspQI sites in original vectors that were mutated from GCTCTTC to WCTWTTC, as explained in Materials and Methods.

the fact that there is no stop codon engineered immediately downstream of the cloned gene. When designing primers for pTEV20, it was necessary to remove the stop codon of the native gene to ensure read through of the rTEV cleavage site and the H_6-tag. After cleavage, proteins obtained from pTEV20 retained six additional amino acids (*i.e.,* Glu-Asn-Leu-Tyr-Phe-Gln) at the *C*-terminus. A simplified list of overhangs to add to primers for each TEV-vector can be seen in Table 2.5.

Isolation of *Salmonella enterica* acetyl-CoA synthetase (Acs) from pTEV16-20

To assess the efficacy of the overexpression vectors, *S.e. acs*$^+$ was cloned into pTEV16-20, overexpressed in *E. coli* strain C41 (λDE3), and proteins were purified as described in the Materials and Methods. Purity of protein purified from cells overexpressing *acs*$^+$ from pTEV16, pTEV17, pTEV18, pTEV19, and pTEV20 was assessed and final protein products are shown in Figure 2.7. Yields of *Se*Acs protein purified from overexpression from each vector were 8.35 mg L^{-1} (pTEV16), 0.9 mg L^{-1} (pTEV17), 8.55 mg L^{-1} (pTEV18), 0.9 mg L^{-1} (pTEV19), 5.32 mg L^{-1} (pTEV20).

CONCLUSIONS

Here we describe modified cloning vectors containing BspQI restriction sites that allow for rapid and efficient cloning of genes of interest. Common issues that may arise during cloning may be resolved by solutions provided in Table 2.6. Using these vectors paired with a previously described cloning method (2), the number of cells carrying empty vectors is greatly reduced. We also have designed the complementation vectors (pCV1-3) and *N*-terminal rTEV cleavable H_6-tag vectors (pTEV18-19) to carry the same primer overhangs so that only one primer set is needed for *in vivo* and *in vitro* studies. Because MBP is a 42-kDA protein, yields of the cleaved protein once

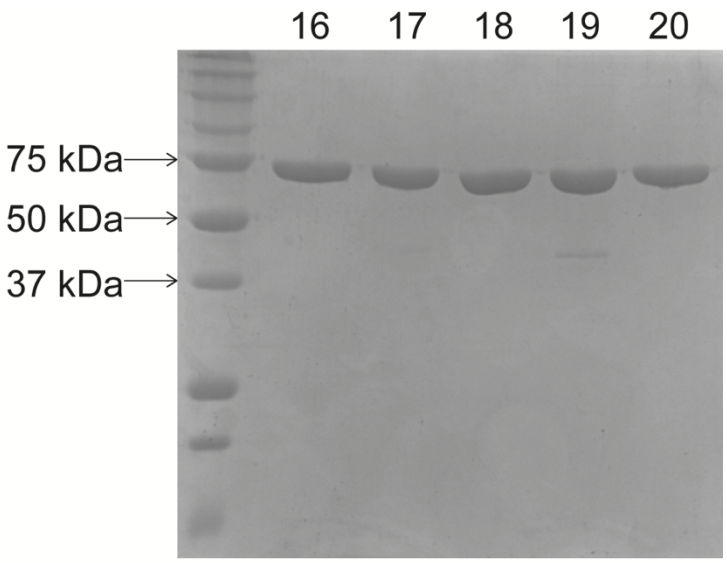


Figure 2.7. SDS-PAGE gel of purified Acs. A 12.5% SDS-PAGE gel shows molecular weight standards (BioRad Precision Plus Protein™, left) and Acs (72 kDa) purified from cells harboring pACS65 (pTEV16), pACS66 (pTEV17), pACS67 (pTEV18), pACS68 (pTEV19), and pACS69 (pTEV20). Numbers above each well correspond to protein purified from cells harboring respective pTEV vectors (*e.g.,* well labeled 16 corresponds to pTEV16).

Table 2.6. Troubleshooting	
Problem	**Potential explanations or solutions**
Cloning yielding all empty vector	1. Potential primer dimers cloning into vectors, gel extract PCR reaction 2. Check activity of BspQI
Cloning yielding incorrect insert sizes	1. Verify competent cells are DH5α and not contaminant 2. Contaminants in PCR reaction, gel extract correct band size
Cloning yielding zero colonies	1. Confirm primer overhangs correspond correctly to vector overhangs 2. Check for BspQI sites in gene, if present eliminate second BspQI digestion step and inactivate ligase at 70°C before transformation 3. Re-PCR amplify insert to rule out mutations in overhangs 4. Verify activity of ligase or stability of ATP

fused to MBP were lower. For this reason, it may be necessary to increase expression volumes to obtain similar yields compared to the pTEV16/18 vectors. It is important to note that when purifying proteins encoded by plasmids pTEV17 or pTEV19, some MBP contamination may occur, which if necessary may be removed using an MBPTrap HP column (GE Healthcare Life Sciences). Additionally, an *N*-terminal tag may be problematic for some proteins due to folding issues or interference with the protein's activity. For this reason we have provided a *C*-terminal rTEV-cleavable H_6-tag vector (pTEV20). Due to the nature of the rTEV cleavage site, recombinant proteins purified and cleaved from the cells harboring this vector will contain additional amino acids compared to the *N*-terminally tagged vectors. Overall, we have established a set of vectors that should facilitate complementation studies or protein production.

ACKNOWLEDGEMENTS

This work was supported by USPHS grant R01 GM062203 to J.C.E.-S. The authors thank Nickolaus Galloway in the laboratory of Dr. Cory Momany for technical assistance.

REFERENCES

1. Engler C, Kandzia R, Marillonnet S. 2008. A one pot, one step, precision cloning method with high throughput capability. PLoS One 3:e3647.
2. Galloway NR, Toutkoushian H, Nune M, Bose N, Momany C. 2013. Rapid cloning for protein crystallography using Type IIS restriction enzymes. Crystal Growth & Design 13:2833-2839.
3. Oster CJ, Phillips GJ. 2011. Vectors for ligation-independent construction of *lacZ* gene fusions and cloning of PCR products using a nicking endonuclease. Plasmid 66:180-185.

4. Engler C, Gruetzner R, Kandzia R, Marillonnet S. 2009. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS One 4:e5553.

5. Engler C, Marillonnet S. 2014. Golden Gate cloning. Methods Mol Biol 1116:119-131.

6. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev SK, Dryden DT, Dybvig K, Firman K, Gromova ES, Gumport RI, Halford SE, Hattman S, Heitman J, Hornby DP, Janulaitis A, Jeltsch A, Josephsen J, Kiss A, Klaenhammer TR, Kobayashi I, Kong H, Kruger DH, Lacks S, Marinus MG, Miyahara M, Morgan RD, Murray NE, Nagaraja V, Piekarowicz A, Pingoud A, Raleigh E, Rao DN, Reich N, Repin VE, Selker EU, Shaw PC, Stein DC, Stoddard BL, Szybalski W, Trautner TA, Van Etten JL, Vitor JM, Wilson GG, Xu SY. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. Nucleic Acids Res 31:1805-1812.

7. Szybalski W, Kim SC, Hasan N, Podhajska AJ. 1991. Class-IIS restriction enzymes--a review. Gene 100:13-26.

8. Pingoud A, Fuxreiter M, Pingoud V, Wende W. 2005. Type II restriction endonucleases: structure and mechanism. Cell Mol Life Sci 62:685-707.

9. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177:4121-30.

10. Rocco CJ, Dennison KL, Klenchin VA, Rayment I, Escalante-Semerena JC. 2008. Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*. Plasmid 59:231-237.

11. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289-298.

12. Berkowitz D, Hushon JM, Whitfield HJ, Jr., Roth J, Ames BN. 1968. Procedure for identifying nonsense mutations. J Bacteriol 96:215-220.

13. Atlas R. 1995. Handbook of Media for Environmental Microbiology, p 6. CRC Press, Boca Raton

14. Maniatis T, Fritsch EF, Sambrook J. 1982. Introduction of plasmid and bacteriophage lambda into *Escherichia coli*, p 250-251. *In* Maniatis T, Fritsch EF, Sambrook J (ed), Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

15. O'Toole GA, Rondon MR, Escalante-Semerena JC. 1993. Analysis of mutants of defective in the synthesis of the nucleotide loop of cobalamin. J Bacteriol 175:3317-3326.

16. Seidman CG, Struhl K, Sheen J, Jessen T. 1997. Introduction of plasmid DNA into cells, p 1.8.-1.8.10. *In* Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (ed), Curr Protoc Mol Biol, vol 1. Wiley Interscience, New York.

17. Hentchel KL, Escalante-Semerena JC. 2015. In *Salmonella enterica*, the Gcn5-related acetyltransferase MddA (formerly YncA) acetylates methionine sulfoximine and methionine sulfone, blocking their toxic effects. J Bacteriol 197:314-325.

18. Cronan JE. 2006. A family of arabinose-inducible *Escherichia coli* expression vectors having pBR322 copy control. Plasmid 55:152-157.

19. Chan CH, Garrity J, Crosby HA, Escalante-Semerena JC. 2011. In *Salmonella enterica*, the sirtuin-dependent protein acylation/deacylation system (SDPADS) maintains energy homeostasis during growth on low concentrations of acetate. Mol Microbiol 80:168-183.

20. Bolivar F. 1978. Construction and characterization of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique Eco RI sites for selection of Eco RI generated recombinant DNA molecules. Gene 4:121-136.

21. Soberon X, Covarrubias L, Bolivar F. 1980. Construction and characterization of new cloning vehicles. IV. Deletion derivatives of pBR322 and pBR325. Gene 9:287-305.

CHAPTER 3

IDENTIFICATION OF PHOSPHINOTHRICIN ACETYLTRANSFERASES USING *IN VIVO*, *IN VITRO,* AND BIOINFORMATICS ANALYSES [1]

[1]VanDrisse C.M, Hentchel K.L., . and Escalante-Semerena J.C. 2016. *Appl. Environ. Microbiol.* 82:7041-7051.

Reprinted here with permission from the publisher.

ABSTRACT

Acetylation of small molecules is widespread in nature, and in some cases, cells use this process to detoxify harmful chemicals. *Streptomyces* species utilize a Gcn5 *N*-acetyltransferase (GNAT), known as Bar, to acetylate and detoxify a self-produced toxin, phosphinothricin (PPT), a glutamate analogue. Bar homologues, such as MddA from *Salmonella enterica* acetylate methionine analogues such as methionine sulfoximine (MSX) and methionine sulfone (MSO), but not PPT, even though Bar homologues are annotated as PPT acetyltransferases. *S. enterica* was used as a heterologous host to determine whether or not putative PPT acetyltransferases from various sources could acetylate PPT, MSX, and MSO. *In vitro* and *in vivo* analysis identified substrates acetylated by putative PPT acetyltransferases from *Deinococcus radiodurans* (DR_1057, DR_1182) and *Geobacillus kaustophilus* (GK0593, GK2920). *In vivo*, synthesis of DR_1182, GK0593, and GK2920 blocked the inhibitory effect of PPT, MSX, and MSO. In contrast, DR_1057 did not detoxify any of the above substrates. Results of *in vitro* studies were consistent with the *in vivo* results. In addition, phylogenetic analyses were used to predict functionality of annotated PPT acetyltransferases in *Burkholderia xenovorans*, *Bacillus subtilis, Staphylococcus aureus, Acinetobacter baylyi,* and *Escherichia coli*.

IMPORTANCE

The work reported herein provides an example of the use of a heterologous system for the identification of enzyme function. Many members of this superfamily of proteins do not have a known function, or it has been annotated solely on the basis of sequence homology to previously characterized enzymes. The critical role of GNATs in the modulation of central metabolic processes, and in controlling metabolic stress, necessitates approaches that can reveal their

physiological role. The combination of *in vivo, in vitro*, and bioinformatics approaches reported herein identified Gcn5 *N*-acetyltransferases (GNATs) that can acetylate and detoxify phosphinothricin.

INTRODUCTION

The Gcn5 *N*-acetyltransferase (GNAT, PF00583) superfamily of proteins is present in all domains of life and catalyzes the transfer of the acetyl group of acetyl-CoA to proteins or small molecules (1). These enzymes were first shown to acetylate and inactivate aminoglycoside antibiotics (2-6) and protect against cellular stressors (1, 7-9), providing a precedent for the role of GNAT-mediated detoxification via acetylation. Previous work identified a subset of GNAT enzymes annotated as phosphinothricin (PPT) acetyltransferases. These enzymes protect cells against the toxic effects of PPT (a glutamate analogue), and oxidized forms of methionine (*e.g.*, methionine sulfoximine (MSX) and methionine sulfone (MSO) (7, 10-14) (Fig. 3.1).

PPT is a component of bialaphos, a tripeptide (phosphinothricyl-alanyl-alanine) with toxic properties produced by *Streptomyces* species (15). The toxic effect occurs when the tripeptide is cleaved, releasing PPT, which inhibits growth by irreversibly binding to glutamine synthetase (GlnA) (16, 17). A GNAT of *Streptomyces* spp, known as Bar (a.k.a. Pat for phosphinothricin acetyltransferase), is responsible for the acetylation and detoxification of the endogenously produced PPT toxin *in vivo* (14). Once acetylated, acetyl-PPT can no longer bind to and inhibit GlnA. Bialaphos is a potent natural herbicide and plants have been genetically engineered to be resistant by encoding the *bar* gene (14, 18).

A subgroup of GNATs has been classified as PPT acetyltransferases based on protein sequence similarity to the Bar protein of *Streptomyces*. Some Bar homologues have been

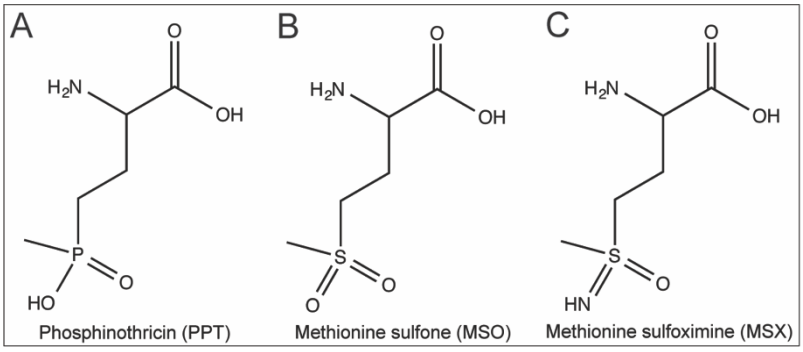


Figure 3.1. Chemical structure of PPT analogues. (A) Phosphinothricin (PPT); (B) Methionine sulfone (MSO); and (C) Methionine sulfoximine (MSX).

incorrectly annotated as PPT acetyltransferases, including *Pseudomonas aeurginosa* PITA (gene locus PA4866) and *Salmonella enterica* MddA (gene locus STM1590) (7). Instead, these enzymes acetylate the structurally related toxic analogues MSX and MSO (Fig. 3.1), which are also inhibitors of glutamine synthetase.

Here we examine the selectivity of putative PPT acetyltransferases from *Geobacillus kaustophilus* and *Deinococcus radiodurans*. Using *S. enterica* as a heterologous host, we assayed for function of the predicted PPT acetyltransferases in *S. enterica mddA1*::*cat*$^{+}$ and *mddA*$^{+}$ strains in culture medium containing MSX, MSO, or PPT. *In vitro* measurements of enzyme activity were consistent with the observed growth phenotypes. The data show diverse specificity of these enzymes for MSX, MSO, and PPT. Lastly, we investigated whether or not putative PPT acetyltransferases from *Burkholderia xenovorans*, *Bacillus subtilis, Escherichia coli, Staphylococus aureus,* and *Acinetobacter baylyi* could acetylate PPT, MSX, and MSO. The data show that the above enzymes acetylate MSX, MSO, and PPT *in vivo,* with the exception of *Ec*YncA, which did not acetylate PPT; and *Bxe*_B1787, which did not acetylate MSO. Our results confirmed bioinformatics-based predictions and provided a reliable *in vivo* approach using *S. enterica* to identify PPT acetyltransferases.

MATERIALS AND METHODS

Culture media and chemicals.

Nutrient broth (NB, Difco) was used as rich medium to grow inocula. The minimal medium used in this study was no-carbon essential (NCE) minimal medium (19) containing $MgSO_4$ (1 mM), sodium, ammonium phosphate (17 mM), Wolfe's trace minerals (1x) (20), and glycerol (22 mM) as the sole source of carbon and energy. When used, antibiotics were added at the following

concentrations: chloramphenicol (20 μg ml^{-1}), ampicillin (100 μg□ml^{-1}). All chemicals were purchased from Sigma-Aldrich unless noted otherwise; Ampicillin, NaCl, and 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, Fischer Scientific); isopropyl β-D-1-thiogalactopyranoside (IPTG, IBI Scientific); and dithiothreitol (DTT, Gold BioTechnology). L-methionine sulfoximine (MSX), L-methionine sulfone (MSO), and DL-phosphinothricin (PPT) were all purchased from Sigma-Aldrich.

Bacterial strains.

All *S. enterica* strains are derivatives of serovar Typhimurium strain LT2 (unless specified), and are listed in Table 3.1.

Plasmid construction for complementation and overexpression.

All plasmids used in this work are listed in Table 3.1. All primers used in this study were synthesized by IDT (Coralville, Iowa) and are listed in Table 3.2. We used cloning technology reported by Galloway *et al* (21) to construct all plasmids in this study. Genes of interest were amplified from genomic DNA from *D. radiodurans* R1, *G. kaustophilus* HTA426, *B. xenovorans* LB400, *E. coli* MG1655, *S. aureus* USA3000, *Acinetobacter baylyi* ADP1, or *B. subtilis* SMY. *D. radiodurans* R1 genomic DNA was a gift from John Batista (Louisiana State University), *A. baylyi* ADP1 genomic DNA was a gift from Cory Momany (University of Georgia), and *S. aureus* USA3000 genomic DNA a gift from Alexander Horswill (University of Iowa). All other genomic DNA was obtained from strains available in our laboratory collection. DNA sequences were confirmed using BigDye® (ABI PRISM) protocols, and sequencing reactions were resolved and analyzed at the Georgia Genomics Facility, UGA.

Plasmid pTEV16 (24), which directs the synthesis of the protein with a cleavable *N*-terminal hexahistidine tag was used for overexpression. The resulting plasmids are referred to as

Table 3.1. Strains and plasmids used in this study.		
***S. enterica* strains**	**Relevant genotype**	**Reference/source[a]**
JE10079	*ara-9 mddA*$^+$	Laboratory strain
Derivatives of JE10079		
JE18333	*mddA*::*cat*$^+$	(10)
JE18961	*mddA*::*cat*$^+$ / pMDD8	(10)
JE20780	pDR_1182-2	
JE20781	*mddA*::*cat*$^+$ / pDR_1182-2	
JE20782	pGK0593-2	
JE20783	*mddA*::*cat*$^+$ / pGK0593-2	
JE21642	pDR_1057-7	
JE21643	*mddA*::*cat*$^+$ / pDR1057-2	
JE20864	*mddA*::*cat*$^+$ / pCV1	
JE20865	*mddA*::*cat*$^+$ / pGK2920-2	
JE20866	pGK2920-2	
JE20973	pCV1	
JE21098	*mddA*::*cat*$^+$ / pBXE_A2261-2	
JE21597	*mddA*::*cat*$^+$ / pBXE_B1787-2	
JE21232	*mddA*::*cat*$^+$ / pBSYwnH-1	
JE21596	*mddA*::*cat*$^+$ / pYNCA	
JE21595	*mddA*::*cat*$^+$ / pSAYwnH-1	
***E. coli* strains**		
E. coli C41(λDE3)	*ompT hsdS* (r_B m_B) *gal* λ (DE3) including at least one non-characterized mutation	(22, 23)
Plasmids		
pMDD7	*S. enterica mddA*$^+$ cloned into pTEV16[b]	(10)
pMDD8	*S. enmterica mddA*$^+$ cloned into pCV1[c]	(10)
pDR1057-8	*D. radiodurans* DR_1057$^+$ cloned into pTEV16[b]	
pDR1057-7	*D. radiodurans* DR_1057$^+$ cloned into pCV1[c]	
pDR1182-1	*D. radiodurans* DR_1182$^+$ cloned into pTEV16[b]	
pDR1182-2	*D. radiodurans* DR_1182$^+$ cloned into pCV1[c]	
pGK0593-3	*G. kaustophilus* GK0593$^+$ cloned into pTEV16[b]	
pGK0593-2	*G. kaustophilus* GK0593$^+$ cloned into pCV1[c]	
pGK2920-1	*G. kaustophilus* GK2920$^+$ cloned into pTEV16[b]	
pGK2920-2	*G. kaustophilus* GK2920$^+$ cloned into pCV1[c]	
pBXE_A2261-2	*B. xenovorans* Bxe_A2261$^+$ cloned into pCV1[c]	
pBXE_B1787-2	*B. xenovorans* Bxe_B1787$^+$ cloned into pCV1[c]	
pBSYwnH-1	*B. subtilis ywnH*$^+$ cloned into pCV1[c]	
pSAYwnH-1	*S. aureus* SAUSA300_2468$^+$ cloned into pCV1[c]	
pYNCA	*E. coli yncA*$^+$ cloned into pCV1[c]	
pCV1	Cloning vector	(24)

[a] All strains and plasmids were constructed during the course of this work, unless otherwise stated
[b,c] (24)

Table 3.2. Primers used in this study.	
Primer Name	**Primer Sequence**
Cloning	
5' DR_1057 pCV1	NNGCTCTTCNTTCATGCCGAGAGCGGCAACGCGCCCAT
5' DR_1057 pTEV16	NNGCTCTTCNAGCATGCCGAGAGCGGCAACGCGCCCAT
3' DR_1057 pVector	NNGCTCTTCNTTATCAGTCCTGTGCCAACGTCCGAACC
5' DR_1182 pCV1	NNGCTCTTCNTTCATGACCTCTGTCATTCGCCCCGCTG
5' DR_1182 pTEV16	NNGCTCTTCNAGCATGACCTCTGTCATTCGCCCCGCTG
3' DR_1182 pVector	NNGCTCTTCNTTATTAGCTCTCCTCGTCCAGCAGCAGT
5' GK0593 pCV1	NNGCTCTTCNTTCATGAACATTCGTAGCTTTCGCAAAG
5' GK0593 pTEV16	NNGCTCTTCNAGCATGAACATTCGTAGCTTTCGCAAAG
3' GK0593 pVector	NNGCTCTTCNTTATCAGTCTATACCTACAACTGGACTG
5' GK2920 pCV1	NNGCTCTTCNTTCTTGCGTAAACGGGCCGAGAAACGCG
5' GK2920 pTEV16	NNGCTCTTCNAGCTTGCGTAAACGGGCCGAGAAACGCG
5' GK2920 M55 pTEV16	NNGCTCTTCNAGCATGGAATCATGTTTGATCCGCGATG
3' GK2920 pVector	NNGCTCTTCNTTACTATACAAGCCGTTTGCCGACAATG
5' *Bxe*_A2261 pCV1	NNGCTCTTCNTTCATGAGCCTTTTCTACCGCGATGCCA
3' *Bxe*_A2261 pVector	NNGCTCTTCNTTCATGAGCCTTTTCTACCGCGATGCCA
5' *Bxe*_B1787 pCV1	NNGCTCTTCNTTCATGAGCACCACCCGCCCCGCGCCCT
3' *Bxe*_B1787 pVector	NNGCTCTTCNTTCATGAGCACCACCCGCCCCGCGCCCT
5' *BsYwnH* pCV1	NNGCTCTTCNTTCATGACATTGCGTCTTGCTGAACATA
3' *BsYwnH* pVector	NNGCTCTTCNTTATCATGAAAGCTCTCTCCCTAAAATT
5' SAUSA300_2468 pCV1	NNGCTCTTCNTTCATGATTAGATACGCTAAAAAAGAGGAT
3' SAUSA300_2468 pVector	NNGCTCTTCNTTACTAGTCTTGTAAATCTAATTCGTAAAATGCTA
5' *EcYncA* pCV1	NNNNNNGAATTCATGTCCATCCGTTTTGCC
3' *EcYncA* pVector	NNNNNNAAGCTTTCATCCAATCGCGTCCGGTT
5' ACIAD_1637 pCV1	NNGCTCTTCNTTCATGTTTTCTCCATCCACTACAACC
3' ACIAD_1637 pVector	NNGCTCTTCNTTATTAATCGTCTTGAGGGTGTAATGGTG
Mutagenesis	
5' DR_1057 N114E	CCCGACCGCTACGAAGTCACGGTCAC
3' DR_1057 N114E	GTGACCGTGACTTCGTAGCGGTCGGG

pDR_1057-8, pDR_1182-1, pGK0593-3, and pGK2920-1. For complementation purposes, each gene of interest was cloned into the L-(+)-arabinose inducible vector pCV1 (24). The names and genotypes of the resulting plasmids are shown in Table 3.1. Site-directed mutagenesis was performed using primers designed from PrimerX (available at http://www.bioinformatics.org/primerx/) to mutate the asparagine (N114) of DR_1057^{WT} to a glutamate residue (E114) to construct a variant (DR_1057^{N114E}) in the pCV1 complementation vector (pDR_1057-3).

Growth behavior analyses.

Starter cultures were grown overnight at 37°C with shaking in NB containing the appropriate antibiotic, 2 μl of which was used to inoculate 198 μl of fresh medium in each well of a 96-well plate. Plates were incubated at 37°C with shaking for 20-48 h in a Powerwave Microplate Reader (Bio-Tek Instruments). Growth studies were performed in triplicate in three independent experiments, with a representative growth curve shown. Data were analyzed using Prism v6 (GraphPad) analytical software. Error bars represent standard deviations. Plasmid-encoded genes of interest were induced with varying concentrations of L-(+)-arabinose, as described in figures and figure legends. Additional chemicals such as PPT, MSX, or MSO were added at concentrations indicated in figures and figure legends.

Protein overproduction and purification.

Plasmids encoding for each protein of interest **(**DR_1182, DR_1057, GK0593, or GK2920$^{M55\text{-}Stop}$) were transformed into *E. coli* C41 (λDE3). Overnight cultures of the transformants were sub-cultured [1:100 (v/v, inoculum:medium)] into 1 L of lysogeny broth (LB) containing ampicillin (100 μg ml^{-1}). Cultures were grown at 37°C with shaking to an OD_{600} of 0.6, induced with IPTG (1 mM), and shaken overnight at ~28°C. Cells were harvested by centrifugation

at 6,000 x *g* in a Beckman Coulter Avanti J-20 XOI refrigerated centrifuge with a JLA-8.1000 rotor for 15 min at 4°C. The collected cell paste was re-suspended in binding buffer A [HEPES buffer (50 mM, pH 7.2) containing NaCl (500 mM) and imidazole (20 mM)] plus lysozyme (1 mg ml^{-1}), DNAse I (25 μg ml^{-1}) and protease inhibitor phenylmethanesulfonyl fluoride (PMSF, 0.5 mM)]. Cells were lysed by sonication for 1 min (2 s, 50% duty) for two rounds on ice using a 550 Sonic Dismembrator (Fisher Scientific) at setting 4. Clarified cell lysates were obtained after centrifugation for 45 min at 4°C at 43,667 x *g* in a Beckman Coulter Avanti J-25I refrigerated centrifuge with a JA–25.50 rotor followed by filtration of the supernatant through a 0.45 μm filter (Millipore). Samples were loaded onto a 1-ml HisPur™ Ni-NTA resin column (Thermo Scientific) at 4°C, pre-equilibrated with binding buffer. The Ni^{+} column was washed first with buffer B [HEPES buffer (50 mM, pH 7.2) with NaCl (500 mM)] that contained imidazole (40 mM) to remove non-specifically bound proteins. Subsequently, the His_6-tagged proteins eluted in the same buffer system that contained a high concentration of imidazole (500 mM). Proteins were dialyzed at 4°C and stored in HEPES buffer (50 mM, pH 7.2) containing NaCl (100 mM), *tris*-(2-carboxyethyl)phosphine hydrochloride (TCEP, 0.5 mM)) and glycerol (10% v/v), drop-frozen in liquid nitrogen, and stored at -80°C. Proteins (DR_1182^{WT}, DR_1057^{WT} $GK0593^{WT}$ and $GK2920^{M55\text{-}Stop}$) were purified and homogeneity ranged from 70-96%, as determined by densitometry calculations using Total Lab v2005 software (Fig. 3.2).

Mass spectrometry analysis of proteins.

Proteins or peptides that required further analysis by mass spectrometry were excised from SDS-PAGE gels and analyzed by the Proteomics and Mass Spectrometry Core Facility at UGA. Peptide sequences were matched to proteins using the MASCOT database (http://www.matrixscience.com), with protein coverage ranging from 55-77%. Gel bands were

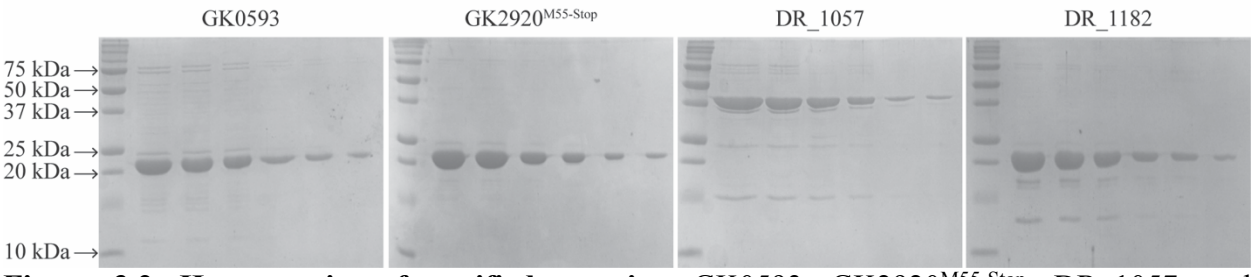


Figure 3.2. Homogeneity of purified proteins. GK0593, GK2920$^{M55\text{-}Stop}$, DR_1057, and DR_1182 were purified on a nickel affinity column. A 15% SDS-PAGE gel shows the molecular mass standards (left), with each gel displaying 4 μg, 3 μg, 2 μg, 1 μg, 0.5 μg, and 0.25 μg of loaded protein sample. Percent purity was calculated using Total Lab v2005 software with GK0593 at 82%, GK2920$^{M55\text{-}Stop}$ at 96%, DR_1057 at 70%, and DR_1182 at 80%.

sliced into small pieces and rinsed with 50% acetonitrile / 20 mM ammonium bicarbonate (~pH 7.5-8) twice. Proteins were reduced and alkylated by dithiothreitol and iodoacetamide, followed by two more rinses with 50% acetonitrile / 20 mM ammonium bicarbonate. The gel pieces were dehydrated by adding 100% acetonitrile. After removing the solvent, the remaining acetonitrile/water was dried out by placing the Eppendorf tubes with gel pieces on a heat block at 60°C, or an incubator at 37°C. A various amount of Trypsin solution (0.1μg/μL in 20 mM ammonium bicarbonate) was added until the gel pieces absorbed the trypsin solution. The tubes were placed in an incubator at 37°C overnight. The tryptic peptides were extracted from gel pieces by incubating with 50% (v/v) acetonitrile / 0.1% (v/v) formic acid twice. The extracts were dried down using a SpeedVac.

The mass spectrometry analyses were performed on either a Bruker Autoflex Matrix-Assisted Laser Desorption Ionization (MALDI) Time-of-Flight mass spectrometer or a Thermo-Fisher LTQ Orbitrap Elite Mass Spectrometer coupled with a Proxeon Easy NanoLC system (Waltham, MA) located at Proteomics and Mass Spectrometry Facility, University of Georgia. For MALDI analysis, the tryptic peptides were re-suspended in 0.1% (v/v) formic acid / 5% (v/v) acetonitrile. Mixing with 1 μL of 10 mg/mL 2,5 dihydroxybenzoic acid in 50% (v/v) methanol, the tryptic peptides were loaded onto the MALDI plate and air dried.

The enzymatic peptides were loaded into a reversed-phase column (Dionex PepMap 100 C8, or ~ self-packed column/emitter with 200 Å 5 μM Bruker MagicAQ C18 resin), then directly eluted into the mass spectrometer. Briefly, the two-buffer gradient elution (0.1% formic acid as buffer A and 99.9% acetonitrile with 0.1% (v/v) formic acid as buffer B) ran with 5% B for 2 min, then increased to 25% B over 60 min, to 40% B over 10 min, and to 95% B over 10 min.

The data-dependent acquisition (DDA) method was used to acquire MS data. A survey MS scan was acquired first, and the top 5 ions in the MS scan were selected for following CID and HCD MS/MS analysis. MS and MS/MS scans were acquired at the resolutions of 120,000 and 30,000, respectively. Data were acquired using Xcalibur software (version 2.2, Thermo Fisher Scientific). Proteins identification and modification characterization were performed using Thermo Proteome Discoverer (version 1.3/1.4) with Mascot (Matrix Science) or SEQUEST (Thermo) programs. The spectra of modified peptides were inspected further to verify the accuracy of the assignments.

Kinetic analyses.

Apparent kinetic parameters (k_{cat}, K_m, k_{cat}/K_m) for DR_1057, DR_1182, GK0593, or GK2920$^{M55\text{-}Stop}$ when MSX, MSO, or PPT was used as substrate were determined using a continuous spectrophotometric assay that employed 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent) as described (10). Briefly, assays were performed at 30°C in 100-µl reaction mixtures in 96-well plates and reactions were monitored at 412 nm (25, 26). Reaction mixtures contained HEPES buffer (50 mM, pH 7.2), DTNB (0.3 mM), saturating levels of acetyl-CoA (1 mM), protein (DR_1057, DR_1182, GK0593, or GK2920$^{M55\text{-}Stop}$) (100-200 nM), and substrate (PPT, MSO, or MSX) (0-2000 µM). Data were acquired every 10 s over a 5-min time period and were reported using the average of technical triplicates. Absorbance (415 nm) measurements of reaction mixtures was monitored over time (s) using the Soft Max Pro 6.2 Spectramax software and absorbance readings were corrected for path lengths for each well. Initial velocity was recorded from the slope of the linear range (ΔA_{415} s^{-1}). ΔA_{415} s^{-1} was converted to µM s^{-1} using Beer's Law (Eq. 1) with the molar extinction coefficient of the TNB^{2-} thiolate anion (14,150 M^{-1} cm^{-1}) (27). Because A_{415} readings were corrected for path lengths, a value of 1 cm was used for *l*

and equation 1 was solved for concentration (c).

$$A = \varepsilon l c \ldots\ldots\ldots \text{(Eq. 1)}$$

Graphs of initial velocity ($\mu M\ s^{-1}$) versus substrate concentration (μM) were plotted using Prism v6 (GraphPad) software. Curves were fit to the model of Michaelis-Menton kinetics to determine K_m and $V_{max.}$ The turnover number (k_{cat}), was determined by equation 2, where [E] was the concentration of protein. The parameters and the curves are shown in Figure 3.3 and Table 3.3.

$$V_{max} = k_{cat}\,[E] \ldots\ldots\ldots \text{(Eq. 2)}$$

Bioinformatics analyses.

An alignment of the primary amino acid sequence of various annotated PPT acetyltransferases was generated using the NCBI COBALT multiple alignment tool (http://www.st-va.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi). A phylogenetic tree was generated using FigTree software (http://tree.bio.ed.ac.uk/software/figtree/).

RESULTS

Phosphinothricin (PPT) delays the onset of exponential growth of *S. enterica.*

Previous work from our research group characterized the MddA (formerly YncA, STM1590) enzyme from *S. enterica* as a MSX and MSO acetyltransferase (10). Although the addition of PPT delayed the onset of exponential growth the effect was not dependent on the presence or absence of MddA (Fig. 3.3D). As shown in figures 3.3C and 2D, PPT only affected growth of *S. enterica* when present at a high concentration (*e.g.,* 100 μM), delaying the onset of exponential growth by ~15 h. The onset of growth was not due to the acquisition of mutations since re-inoculation of fresh medium with cells from cultures that experienced the lag phase did not result in any reduction of the lag time (data not shown). We used the phenotypes in Figure 3.3

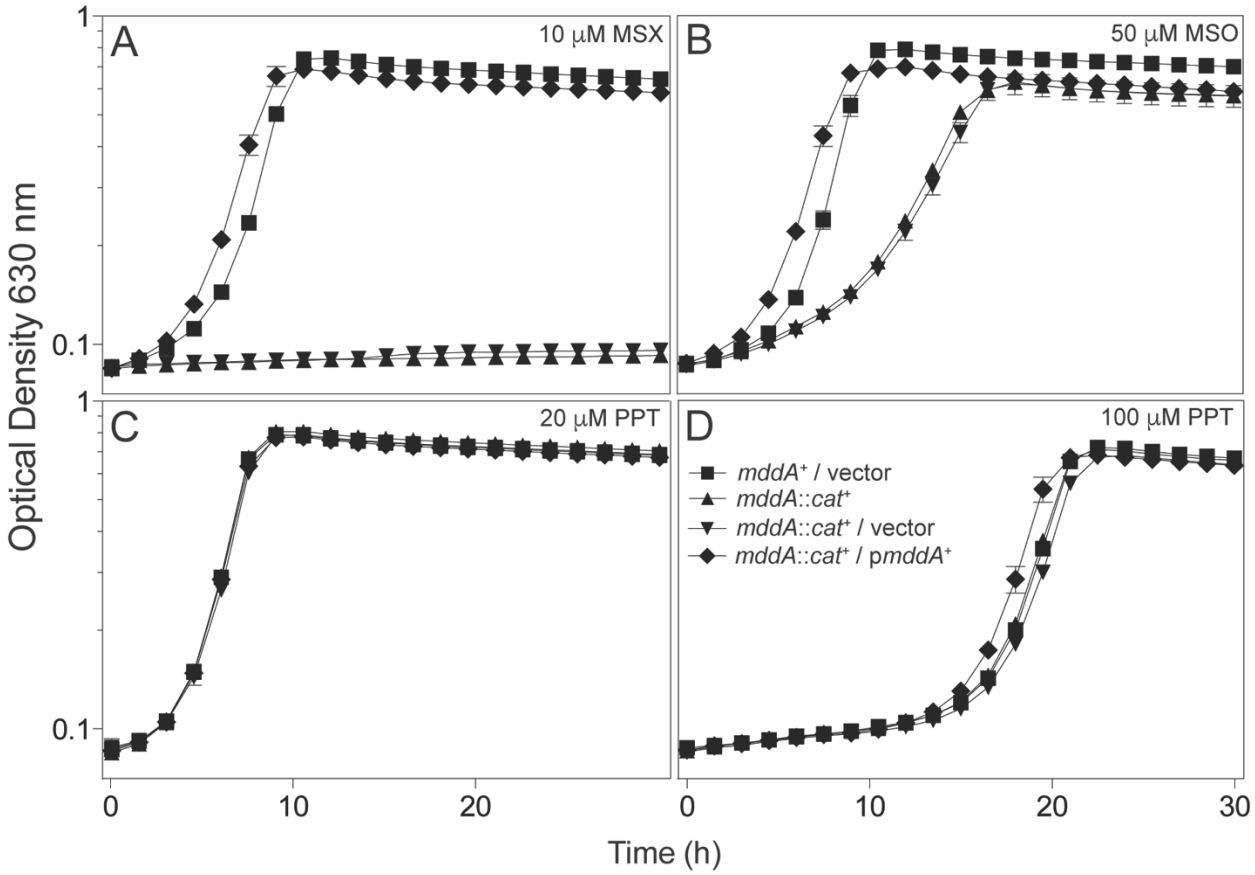


Figure 3.3. Growth of the *S. enterica mddA1*::*cat*$^+$ strain in conditions containing MSX, MSO, or PPT. Growth and complementation of the *S. enterica mddA1*::*cat*$^+$ strain in NCE minimal medium (glycerol, 22 mM) was examined in the presence of MSX (panel A, 10 μM), MSO (panel B, 50 μM), or PPT (panel C, 20 μM; panel D, 100 μM). Vectors were induced with 10 μM L-(+)-arabinose. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under *Materials and Methods*. Strains analyzed: *ara-9* / vector (JE20973, squares), *ara-9 mddA1*::*cat*$^+$ (JE18333, triangles), *ara-9 mddA1*::*cat*$^+$ / vector (JE20864, inverted triangles), *ara-9 mddA1*::*cat*$^+$ / *mddA*$^+$ (pMDD8; JE18961, diamonds). Error bars represent standard deviation. Symbols in panel D apply to all panels.

Table 3.3. Kinetic parameters[a] of *G. kaustophilus* and *D. radiodurans* acetyltransferases				
Enzyme	Substrate	$K_{m(app)}$ (μM)	k_{cat} (s^{-1})	k_{cat} / $K_{m(app)}$ (M^{-1} s^{-1})
GK2920$^{M55\text{-}Stop}$	MSX	121 ± 26	9 ± 1	7 x 10^4
	MSO	89 ± 16	10 ± 1	1 x 10^5
	PPT	NDA[b]	NDA	NDA
GK0593	MSX	17 ± 4	1 ± 0.1	6 x 10^4
	MSO	4 ± 2	1 ± 0.1	3 x 10^5
	PPT	5 ± 3	1 ± 0.1	2 x 10^5
DR_1182	MSX	316 ± 23	11 ± 0.3	3 X 10^4
	MSO	198 ± 30	11 ± 1	6 X 10^4
	PPT	721 ± 79	5 ± 0.2	7 X 10^3

to identify PPT acetyltransferases from other microorganisms.

***Deinococcus radiodurans* and *Geobacillus kaustophilus* encode PPT acetyltransferases with different substrate preferences**.

To simplify the interpretation of our results, the activities of putative PPT acetyltransferases were assessed using a *S. enterica* strain devoid of MddA (*mddA1*::*cat*$^+$, JE18333). Ectopic synthesis of *Deinococcus radiodurans* DR_1182 or *Geobacillus kaustophilus* GK0593 restored growth of the *S. enterica mddA1*::*cat*$^+$ strain to wild-type levels at low levels of induction (10 μM L-(+)-arabinose) in medium supplemented with either MSX or MSO (Fig. 3.4). In contrast, expression of GK2920 restored growth at concentrations of inducer ranging between 0.2-1 mM, suggesting that MSX and MSO were poor substrates for the enzyme. Synthesis of DR_1057 protein failed to correct the phenotype of the *mddA1*::*cat*$^+$ strain caused by MSX or MSO, even when the medium contained high levels of inducer (*i.e.,* 1 mM L-(+) arabinose) (open squares, Fig. 3.4A, 3.4C, respectively).

Direct correlation between enzyme level and degree of protection.

Because *S. enterica* MddA does not acetylate PPT, we used the wild-type *S. enterica* strain to screen for phenotypes at higher concentrations of MSX, MSO, and PPT. Growth of wild-type *S. enterica* was negatively affected but not abolished by MSX (20 μM) and MSO (500 μM), and PPT (100 μM) (10). Plasmids carrying DR_1057$^+$, DR_1182$^+$, GK0593$^+$ or GK2920$^+$ placed under the control of an L-(+)-arabinose inducible promoter were each introduced into an *S. enterica mddA*$^+$ strain. The resulting strains were grown in minimal medium containing MSX (20 μM), MSO (500 μM), or PPT (100 μM), and varying levels of inducer (10-1000 μM L-(+)-arabinose). When induced, cultures expressing DR_1182$^+$ or GK2920$^+$ grew with substantially shorter lag times when compared to the *mddA*$^+$ *S. enterica* strain carrying the vector only (VOC, vector-only

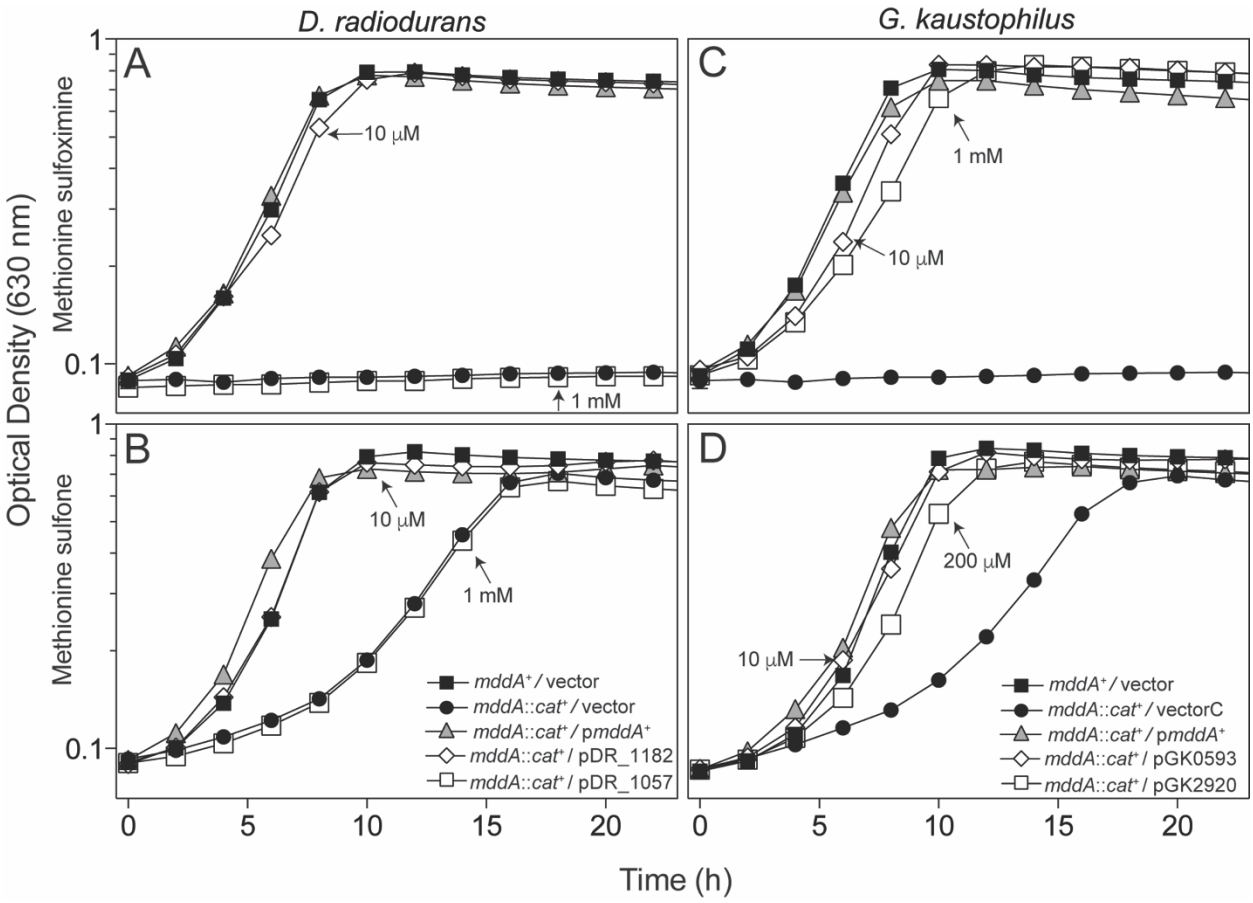


Figure 3.4. Complementation studies using genes encoding putative PPT acetyltransferases from *D. radiodurans* and *G. kaustophilus* in *S. enterica*. Growth of an *S. enterica mddA1*::*cat*$^+$ strain carrying an plasmid encoding DR_1057, DR_1182, GK0593, or GK2920 under the control of an L-(+)-arabinose inducible promoter was examined in the presence of MSX (10 μM, panels A and C) or MSO (50 μM, panels B and D) in glycerol (22 mM) minimal medium. Plasmids were induced with varying concentrations of L-(+)-arabinose (0.01-1 mM), as indicated. Strains analyzed: *ara-9* / vector (JE20973, squares), *ara-9 mddA1*::*cat*$^+$ / vector (JE20864, circles), *ara-9 mddA1*::*cat*$^+$ / p*mddA*$^+$ (JE18961, triangles), *ara-9 mddA1*::*cat*$^+$ / pDR_1057-2 (JE21642, open squares in panels A and C), *ara-9 mddA1*::*cat*$^+$ / pDR_1182-2 (JE20781, open diamonds in panels A and C), *ara-9 mddA1*::*cat*$^+$ / pGK0593-2 (JE20783, open diamonds in panels B and D), *ara-9 mddA1*::*cat*$^+$ / pGK2920-2 (JE20865, open squares in panels B and D). Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under *Materials and Methods*. Error bars represent standard deviation. Symbols in B apply to panels A and B; symbols in D apply to panels C and D.

control) (Fig. 3.5). These data were consistent with the idea that increased protein levels provided greater protection against the toxic effects of MSX, MSO, and PPT. Low levels of induction (*i.e.,* 10 μM L-(+) arabinose) generated sufficient GK0593 enzyme to completely reduce the growth lag of the *S. enterica mddA*$^+$ strain in medium containing 100 μM PPT (Fig. 3.5E). In contrast, the same level of induction of GK0593$^+$ provided limited protection against MSX (20 μM) or MSO (500 μM). A 10- and 100-fold higher level of inducer was needed to express sufficient GK0593 protein to reduce the lag phase in the presence of MSX or MSO, respectively (Fig. 3.5D, 4.5F). The DR_1057 enzyme did not improve growth of the *mddA*$^+$ *S. enterica* strain in medium containing MSX, MSO, or PPT even at high induction (1 mM L-(+) arabinose), suggesting that this enzyme may not recognize MSX, MSO, or PPT as substrates (Fig. 3.5A-C). The GK2920$^+$ and DR_1182$^+$ alleles corrected all phenotypes at low induction (10 μM *L*-(+) arabinose) (Fig. 3.5A-F).

***In vitro* activity analyses of putative PPT acetyltransferases.**

To study the substrate preference of the annotated PPT acetyltransferases in *D. radiodurans* [DR_1057, DR_1182] and *G. kaustophilus* [GK0593, GK2920] the proteins were isolated to 70-95% homogeneity (Fig. 3.2), and enzyme kinetic parameters were determined to gain insights into their substrate preference (Table 3.3, Fig. 3.6). The catalytic efficiencies of *Deinococcus* enzymes correlated with *in vivo* data, with DR1182 having similar k_{cat} / K_m values for MSX, MSO, and PPT as substrates (Table 3.3). Given that DR_1057 failed to acetylate any of the substrates *in vivo,* it was not surprising that this enzyme did not acetylate MSX, MSO, or PPT under the *in vitro* conditions used (data not shown).

Acetyltransferases from *Geobacillus* showed similar catalytic efficiencies for MSX and MSO as substrates, with GK0593 having a higher catalytic efficiency for PPT than GK2920. We

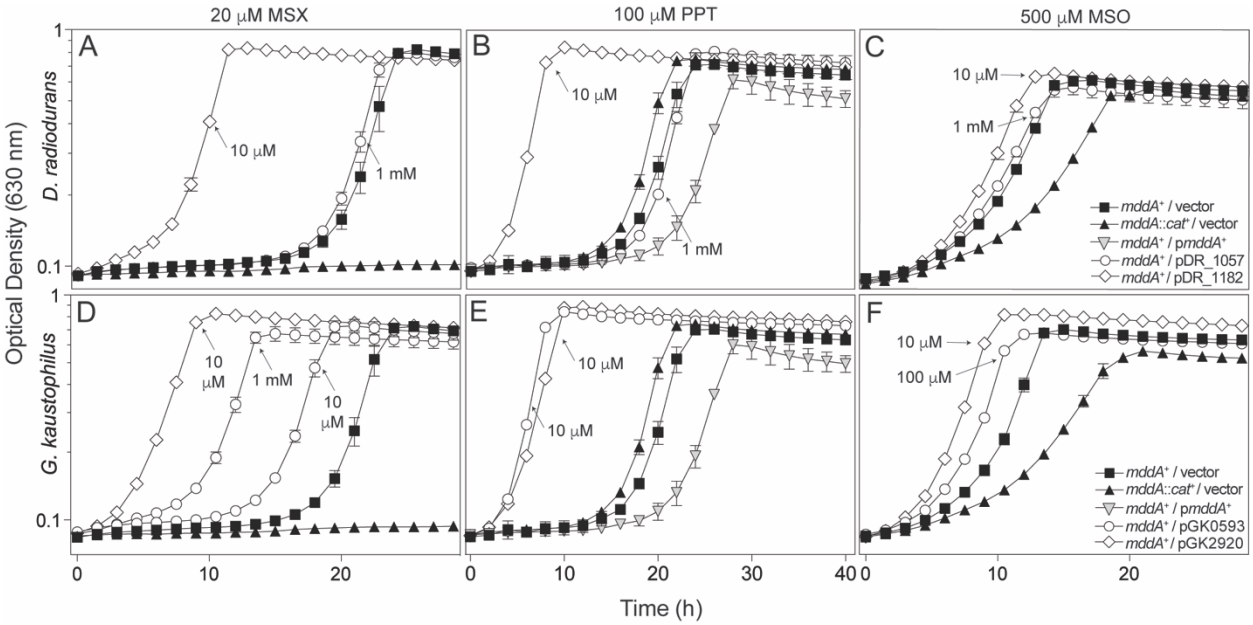


Figure 3.5. Overexpression provides resistance to higher levels of PPT, MSX, and MSO. Growth of wild-type *S. enterica* (JE10079) carrying a plasmid encoding DR_1057, DR_1182, GK0593, or GK2920 under the control of an L-(+)-arabinose inducible was examined in presence of PPT (100 μM, B and E), MSX (20 μM, A and D), or MSO (500 μM, C and F), in glycerol (22 mM) minimal medium with glycerol with varying concentrations of L-(+)-arabinose (0.01-1 mM), as indicated. Strains analyzed: *ara-9* / vector (JE20973, squares), *ara-9 mddA1*::*cat*$^+$ / vector (JE20864, triangles), *ara-9 mddA1*::*cat*$^+$ / p*mddA*$^+$ (JE18961, inverted triangles) *ara-9* / pDR_1057-2 (JE21642, circles, panels A-C), *ara-9* / pDR_1182-2 (JE20780, diamonds, panels A-C), *ara-9* / pGK0593-2 (JE20782, circles, panels D-F), *ara-9* / pGK2920-2 (JE20866, diamonds, panels D-F). Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under *Materials and Methods*. Error bars represent standard deviation. Symbols in panel C apply to panels A-C; symbols in panel F apply to panels D-F.

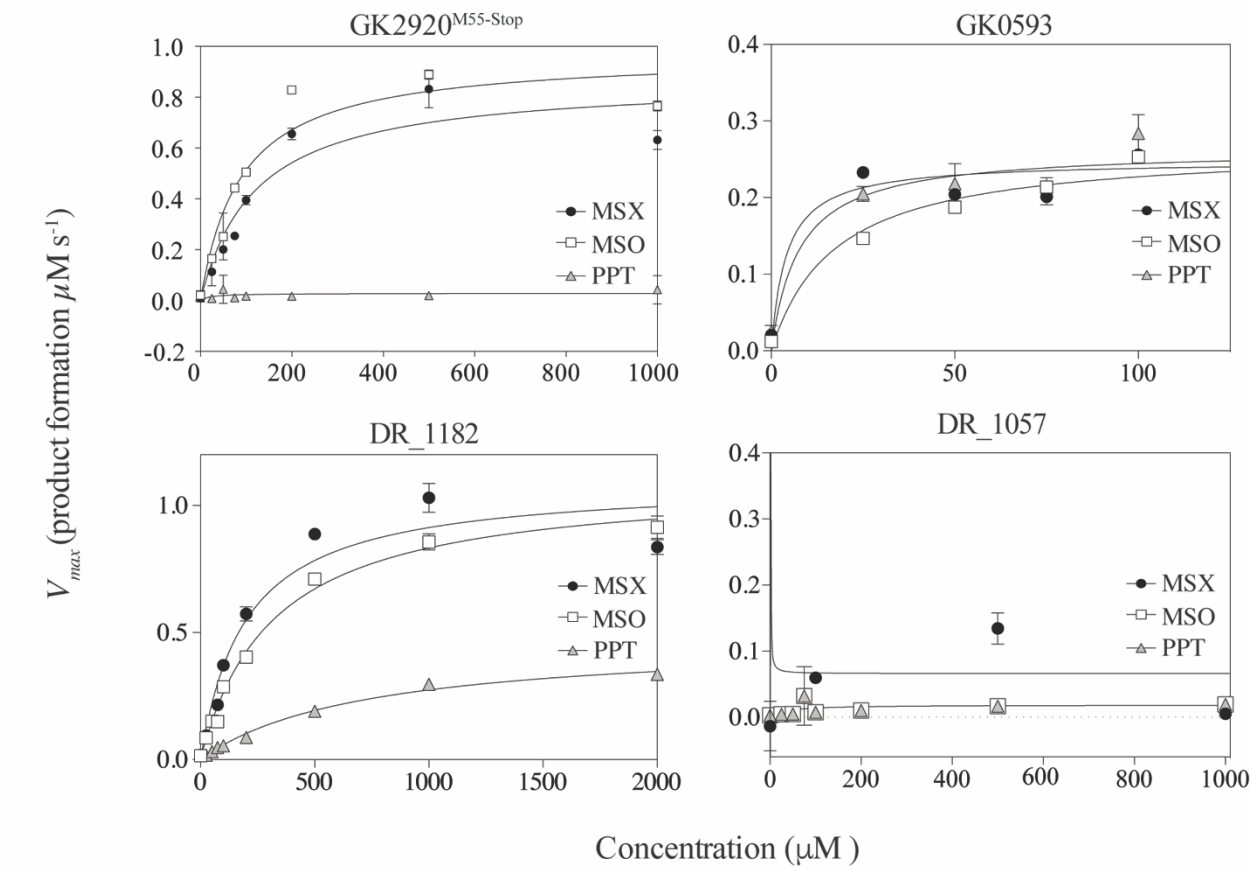


Figure 3.6. Michaelis Menten saturation curves. Phosphinothricin acetyltransferases (100 nM of GK2920$^{M55\text{-}Stop}$, DR_1057, and DR_1182 and 200 nM of GK0593) were incubated with MSX, MSO, or PPT (0-2000 μM), Acetyl-CoA (1 mM), and DTNB (0.3 mM) at 30°C. Each data point represents the mean ± S.D. of three replicates from a single experiment. The data were fit to a first-order rate equation as indicated by black lines. The protocol for the activity assay is described in the manuscript *Materials & Methods* section.

noted that during the purification of GK2920, high amounts of protein were collected in flow-through fractions. We posited that during overproduction, the protein was cleaved into two peptides, resulting in one cleavage product unable to interact with the Ni-NTA resin column. The GK2920 protein fragments were excised from the gel and analyzed by LC-MS/MS, which revealed that the first 55 amino acids of the *N*-terminus (including the His_6-tag) of GK2920 were cleaved by an unknown mechanism (Fig. 3.7). The allele encoding $GK2920^{M55\text{-}Stop}$ was cloned into an overexpression vector and subsequently purified. Results from kinetic analyses showed that $GK2920^{M55\text{-}Stop}$ acetylated MSX and MSO, but had no activity when PPT was the substrate (Table 3.3). This result was unexpected as the *in vivo* data shown in Figure 3.5E (diamonds) demonstrated that low levels of induction of GK2920 substantially improved growth of the *S. enterica mddA*$^+$ strain in medium containing PPT. Our interpretation of this discrepancy is considered below in the discussion section.

Bioinformatics analyses of putative PPT acetyltransferases.

An alignment was generated for the active sites of putative PPT acetyltransferases across a range of bacterial species (NCBI COBALT Multiple Alignment Tool). From this alignment we noted that DR_1057 from *D. radiodurans* lacked the putative catalytic glutamate of GNATs, and encoded an asparagine at that position (Fig. 3.8A). To determine whether or not the lack of *in vivo* function against MSX, MSO, and PPT was due to the absence of the glutamate residue, we engineered a variant of DR_1057 in which the asparagine residue was replaced by a glutamate (DR_1057^{N114E}). The resulting variant did not restore growth of an *S. enterica mddA*::*cat*$^+$ strain in medium containing MSX or MSO, and did not improve growth of the *S. enterica mddA*$^+$ strain in medium containing PPT (data not shown). Therefore it was concluded that the absence of alluded glutamate residue in DR_1057 was not responsible for the lack of enzyme activity.

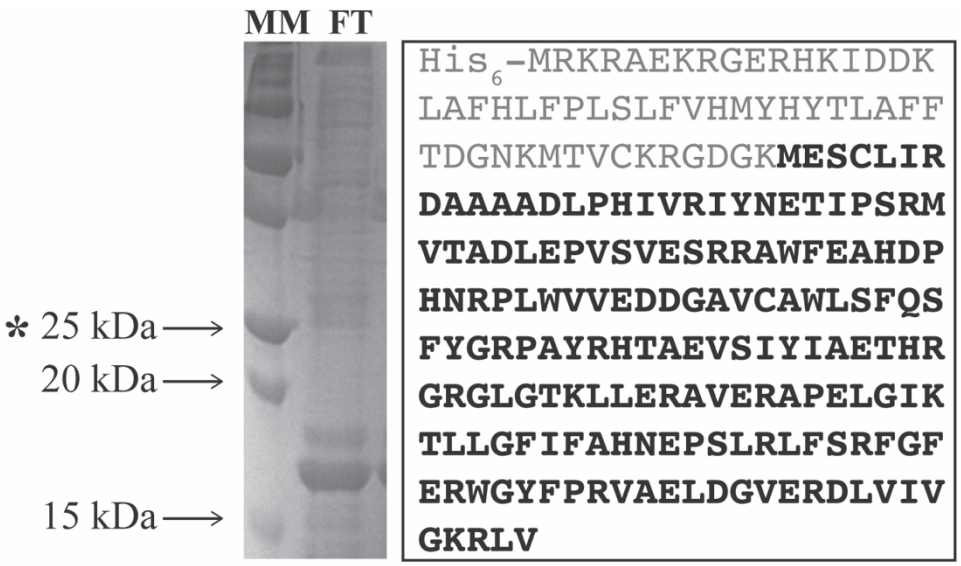


Figure 3.7. Isolation of a truncated form of GK2920. Left panel shows an SDS-PAGE gel of the flow-through (FT) fraction from the Ni-NTA affinity chromatography column used to purify GK2920. Arrows indicate relevant mass markers (MM, BioRad Precision Plus Blue Standard). The asterisk indicates the predicted size of $GK2920^{WT}$. LC/MS of the 17-kDa protein revealed the amino acid sequence highlighted in bold in the right panel.

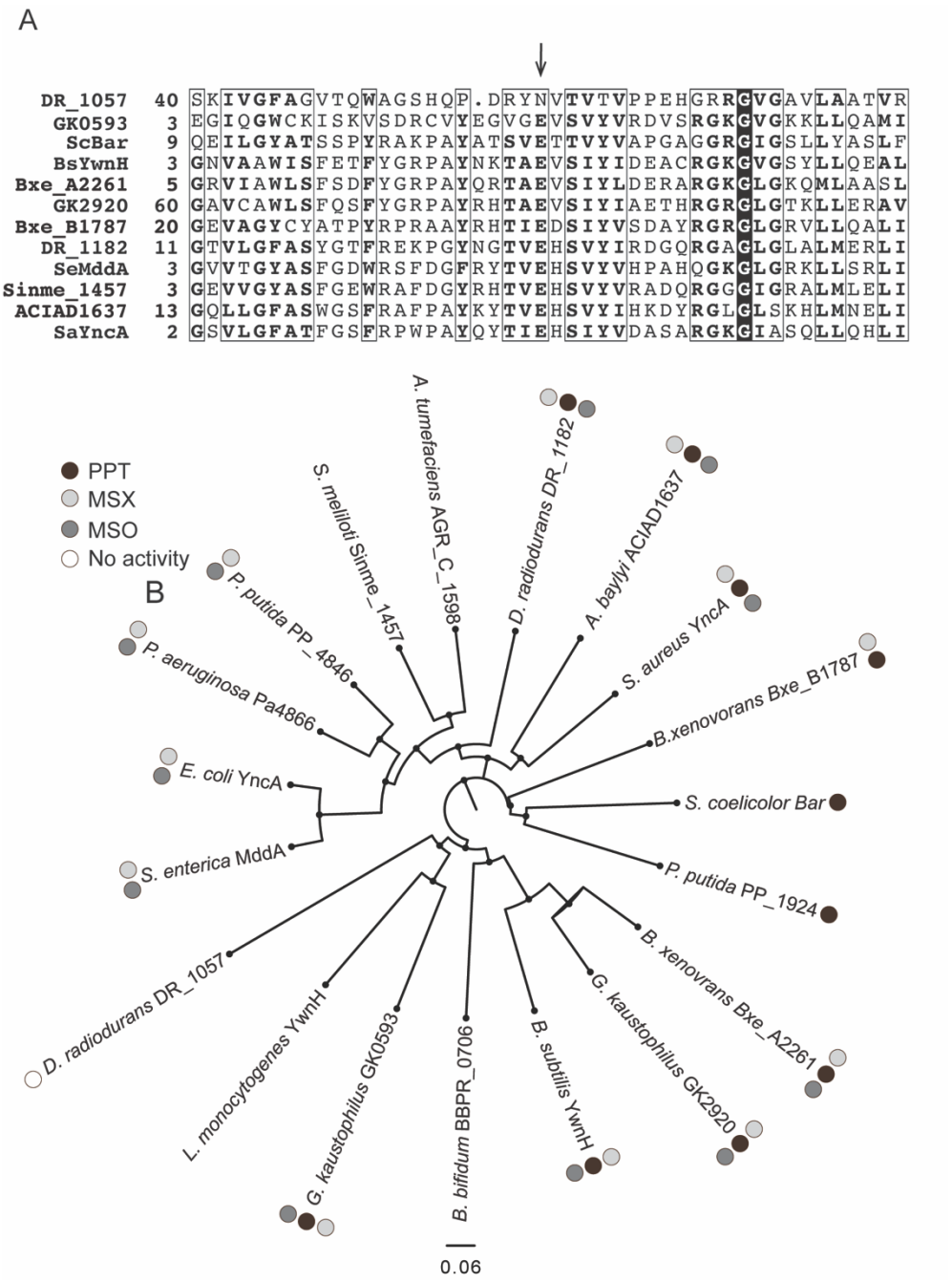


Figure 3.8. Bioinformatics analyses of annotated PPT acetyltransferases. (A) Alignments were generated using NCBI COBALT Multiple Alignment Tool. The arrow points at the predicted active-site glutamate of acetyltransferases. Numbers on the left of the sequences indicate the location of the first residue in the native protein; boxed residues show regions of at least 70% similarity; a conserved glycyl residue is identified by white letters in a black background. (B) A phylogenetic tree of 19 annotated PPT acetyltransferases was generated using FigTree software (http://tree.bio.ed.ac.uk/software/figtree/). For DR_1057 only the sequence of the GNAT domain that aligned to the other putative PPT acetyltransferases (residues 1-180) was used for the phylogenetic tree. When gene names were unavailable, locus tags were used (*i.e* for *P. putida,* PP_1924); 0.06 designates scale for tree length.

A phylogenetic tree was generated by comparing 19 PPT acetyltransferase homologues (FigTree, Fig. 3.8B). Clustering of enzymes that acetylated MSX and MSO but not PPT was observed with the exception of DR_1182, which clustered more closely with the non-PPT utilizing enzymes. Organisms containing two annotated PPT acetyltransferases clustered separately in different nodes of the tree (*e.g., D. radiodurans, B. xenovorans,* and *G. kaustophilus*). A recent report examining the specificity of the two annotated PPT acetyltransferases from *Pseudomonas putida,* which were also clustered in separate nodes of the tree, demonstrated that each enzyme was specific for either PPT or MSX and MSO (28). From *in vivo* data reported herein for *G. kaustophilus*, we see that enzyme function overlaps and each of the two enzymes can acetylate PPT, MSX, and MSO, although why an organism would encode two enzymes with overlapping function is still not clear.

Can PPT acetyltransferases be identified by computational approaches?

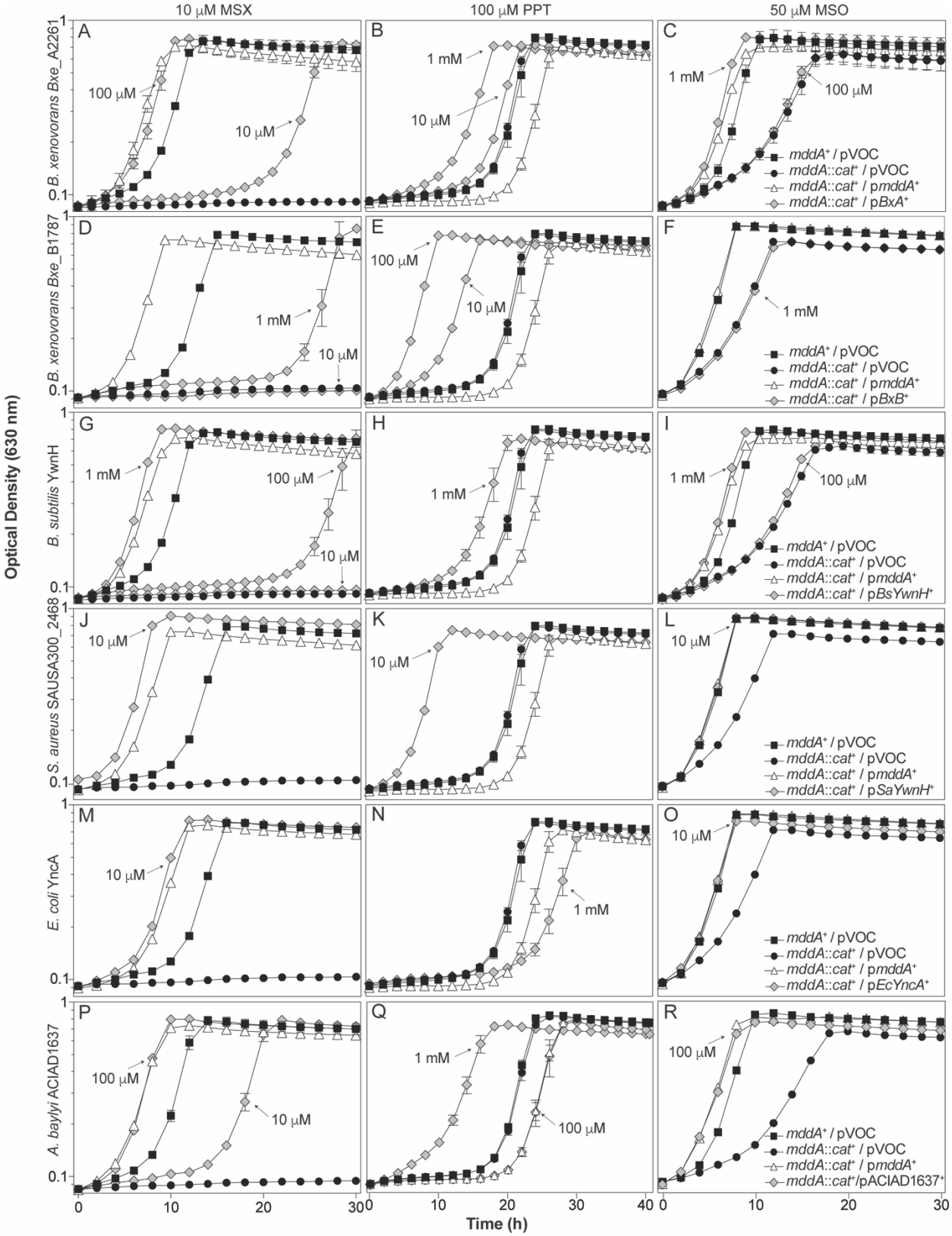
The goal of this work was to expedite the validation of putative of PPT acetyltransferases by using the above-mentioned *in vivo*, *in vitro*, and bioinformatics approaches. As proof of principle we focused on putative PPT acetyltransferases from *B. xenovorans* (*Bxe*_A2261, *Bxe*_B1787), *B. subtilis* (YwnH), *Staphyloccocus aureus* (SAUSA300_2468), and *E. coli* (YncA). We wished to determine whether or not there was divergence in the phylogeny that correlated with the ability of a particular organism to acetylate PPT. Based on the phylogenetic tree and proximity to known PPT-utilizing enzymes, the following predictions were made. We predicted that *Bxe*_A2661, *Bxe*_B1787, and *Bs*YwnH would acetylate PPT, while *Sa*YwnH and *Ec*YncA would not.

B. xenovorans Bxe_A2261 and *Bxe_B1781*, *B. subtilis ywnH*, *S. aureus* SAUSA300_2468, and *E. coli yncA* genes were cloned under the control of an L-(+)-arabinose-inducible promoter and

expressed *in trans* in an *S. enterica mddA1*::*cat*$^+$ strain grown in the presence of MSX (10 μM), MSO (50 μM), or PPT (100 μM) (Table 3.3). Each gene was induced with varying concentrations of L-(+)-arabinose (10 μM, 100 μM, or 1 mM) (Fig. 3.9). Acetylation of MSO or MSX correlated with the growth profile of the *S. enterica mddA*::*cat*$^+$ / p*mddA*$^+$ strain (JE18961), relative to that of the *mddA1*::*cat*$^+$ / pVOC strain (JE20864). Acetylation of PPT correlated with a faster growth rate and faster entry into stationary phase than the *mddA*$^+$ / pVOC strain.

As predicted, *Bxe*_A2661 acetylated PPT *in vivo*, as seen by the reduction in lag phase at high concentrations of inducer; *Bxe*_A2661 also acetylated MSX and MSO (Fig. 3.9A-C). *Bxe*_B1787 acetylated PPT *in vivo* and MSX acetylation was limited (as demonstrated by the extended lag phase compared to *mddA*::*cat*$^+$ / p*mddA*$^+$) and was only observed at high concentration of inducer, whilst MSO was not acetylated even at high inducer concentration (Fig. 3.9D-F). *Bs*YwnH acetylated MSX, MSO, and PPT as predicted, but only at high inducer concentration (Fig. 3.9G-I). Surprisingly, *Sa*YwnH acetylated PPT, MSX, and MSO even at low inducer concentrations (Fig. 3.9J-L). Lastly, *Ec*YncA acetylated MSX and MSO, but not PPT, as predicted (Fig. 3.9M-O).

It was intriguing that SAUSA300_2468 (encodes *Sa*YwnH) clustered closely with *A. baylyi* ACIAD1637, yet the latter only acetylated MSX and MSO, whereas *Sa*YwnH acetylated all three substrates. Studies that reported the crystallization and characterization of ACIAD1637 revealed that ACIAD1637 preferred MSX and MSO and PPT was not considered as a valid substrate for ACIAD1637 (12). To resolve the discrepancies of the bioinformatics prediction, we sought to examine substrates of ACIAD1637 *in vivo*. ACAID1637 was cloned into pCV1 and was induced with varying levels of arabinose in an *mddA1*::*cat*$^+$ strain to determine whether or not its substrate preference *in vivo* was similar to that SAUSA300_2468. *In vivo* analysis showed that synthesis of



legend on next page

Figure 3.9. *In vivo* functional analysis of putative PPT acetyltransferases from *B. xenovorans*, *B. subtilis*, *S. aureus, E. coli, and A. baylyi.* Growth behaviors of an *S. enterica mddA1*::*cat*$^+$ strain carrying a plasmid encoding *Bxe*_A2261, *Bxe*_B1787, *Bs*YwnH, *Sa*YwnH, *Ec*YncA, or ACIAD1637 under the control of an L-(+)-arabinose inducible promoter. Growth was examined in the presence of MSX (10 μM, panels A, D, G, J, M, P), PPT (100 μM, panels B, E, H, K, N, Q), or MSO (50 μM, panels C, F, I, L, O, R), in glycerol (22 mM) minimal medium with varying concentrations of L-(+)-arabinose (0.01-1 μM), as indicated. Strains analyzed: *ara-9* / vector (JE20973, squares), *ara-9 mddA1*::*cat*$^+$ / vector (JE20864, circles), *ara-9 mddA1*::*cat*$^+$ / pMDD8 (JE18961, diamonds), *ara-9 mddA1*::*cat*$^+$ / p*Bxe*_A2261 (JE21098), *ara-9 mddA1*::*cat*$^+$ / p*Bxe*_B1787 (JE21597), *ara-9 mddA1*::*cat*$^+$ / p*Bs*YwnH (JE21232), *ara-9 mddA1*::*cat*$^+$ / p*Sa*YwnH (JE21595), *ara-9 mddA1*::*cat*$^+$ / p*Ec*YncA (JE21596), and *ara-9 mddA1*::*cat*$^+$ / pACIAD1637 (JE21982). All strains that detoxified PPT are identified by grey diamonds. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under *Materials and Methods*. Error bars represent standard deviation. Symbols in panel C apply to panels A-C; symbols in panel F apply to panels D-F; symbols in panel I apply to panels G-I, symbols in panel L apply to panels J-L; symbols in panel O apply to panels M-O; symbols in panel R apply to panels P-Q.

ACIAD1637 blocked the deleterious effects of MSX, MSO, and PPT (Fig. 3.9P-R), in support of the prediction that *Sa*YwnH was a PPT acetyltransferase.

DISCUSSION

The combination of *in vivo* and *in vitro* approaches used in this study allowed us to validate the function of putative PPT acetyltransferases identified by bioinformatics and phylogenetic analyses. From their position in the phylogenetic tree (Fig. 3.8) several predictions were made about putative PPT acetyltransferases that were experimentally confirmed.

The combination of phenotypic analyses using a heterologous system and kinetic analyses offers the means to assign PPT acetyltransferase function.

Results obtained with the *G. kaustophilus* GK2920 protein were intriguing. For as-yet-unknown reasons, GK2920 was isolated as a cleaved protein ($GK2920^{M55\text{-}stop}$, Fig. 3.7) that retained biochemical activity when MSO or MSX were used as substrates, but not when PPT was the substrate. The results of *in vitro* assays with MSO and MSX indicate that the active site of the $GK2920^{M55\text{-}stop}$ enzyme remained functional, but appeared to be sufficiently altered to prevent PPT acetylation. One explanation for the lack of activity with PPT may be that the *N*-terminal 54 residues of GK2920 are needed for PPT binding. Clearly the GK2920 protein efficiently detoxifies PPT *in vivo* (Fig. 3.5E), whether or not the protein remains intact or is truncated is unclear at this time. We also do not yet understand why GK2920 is cleaved during isolation given that a protease inhibitor cocktail is included during cell breakage to decrease the possibility of degradation. Purified $GK2920^{M55\text{-}stop}$ showed higher catalytic efficiency for MSO than for MSX (Table 3.3), a result that was consistent with those obtained *in vivo* (Fig. 3.4C,D). Taken together, these data suggest that MSO may be the preferred substrate for GK2920.

Kinetic analyses of GK0593 were also consistent with *in vivo* data, in that this enzyme acetylated all three substrates *in vitro*, while also complementing *Salmonella* phenotypes associated with toxic levels of MSX, MSO, and PPT. The catalytic efficiency of GK0593 for MSO and PPT was higher than that for MSX, which may reflect the levels of these substrates found in the environment of *Geobacillus*. When comparing GK2920 and GK0593, the catalytic efficiencies were similar for both MSX and MSO, raising questions about why this organism encodes two acetyltransferases with overlapping functions. Because GK2920 could not acetylate PPT *in vitro*, no conclusions can be drawn when comparing kinetic parameters to GK0593.

Synthesis of the *D. radiodurans* DR_1057 failed to prevent the inhibitory effect of PPT even when the gene encoding it was expressed with high levels of inducer (Fig. 3.5B). These results may suggest that DR_1057 may have diverted from the protein conformation that allows it to acetylate PPT. We note, that the *D. radiodurans* genome contains one additional gene encoding a putative PPT acetyltransferase (*i.e.,* DR_1182). In contrast to DR_1057, synthesis of DR_1182 blocked the inhibitory effect of PPT (Fig. 3.5B). Additionally, DR_1182 showed higher catalytic efficiencies for MSX and MSO as substrates when compared to PPT. *D. radiodurans* encodes two putative phosphinothricin acetyltransferases, although only one of these enzymes (DR_1182) carries out its annotated function. Either DR_1057 lost the ability to acetylate MSX, MSO, and PPT, or this enzyme is an acetyltransferase with alternative target(s). Further analysis of this enzyme is needed.

Bioinformatics analyses provide additional means for the identification of PPT acetyltransferases.

From the phylogenetic tree in Figure 3.8 we can determine where the limit of an enzymes ability to acetylate PPT may lie. It appears as though enzymes clustering near *S. coelicolor* Bar

PPT acetyltransferase can acetylate PPT, while enzymes clustering near *S. enterica* MddA cannot. *In vivo* analyses validated PPT acetyltransferase activities of several proteins of interest. That is, synthesis of genes clustered near *bar* (*B. xenovorans Bxe*_A2261 and *Bxe*_B1787, *B. subtilis* YwnH, and *S. aureus* SAUSA300_2468) clearly blocked the deleterious effect of PPT on the indicator strain (Figs. 3.9 B, E, H, K, respectively). The lack of protection against PPT by *E. coli* YncA was not surprising, given the phylogenetic relatedness of this protein to the *S. enterica* MddA enzyme, which has been shown to lack PPT acetyltransferase activity (10).

Correction of a functional assignment in the literature.

Our results with *A.b.* ACIAD1637 (Fig. 3.9Q) were inconsistent with those reported by Davies *et al*, who concluded that *A. baylyi* ACIAD1637 did not have PPT acetyltransferase activity. These authors reached their conclusion because wild-type and ACIAD1637-devoid strains of *A. baylyi* were equally sensitive to PPT (12). We note, however, that although *A. baylyi* ACIAD1637 blocked PPT inhibition of the indicator strains, high level of inducer (1 mM) was needed to afford the observed protection, suggesting that *A. baylyi* ACIAD1637 may have low level of PPT acetylation activity. Regardless, these results highlight the usefulness of the *in vivo* system we report herein for the assessment of PPT acetyltransferase function.

Usefulness of heterologous systems in the assignment of gene function.

The use of model organisms with sophisticated genetic systems that allow investigators to rapidly construct strains in diverse genetic backgrounds is a powerful, yet simple, way to probe for gene function. Results obtained with such a system guides the use of phylogenetic, bioinformatics, and biochemical approaches to further cement the assignment of protein function. One factor to consider, however, is that a heterologous system could produce false negatives depending on protein solubility or degradation in a heterologous environment. It is important to

note that this heterologous system will not identify substrate preferences, but will provide a tool to narrow down enzymes to be kinetically characterized. Results from this study demonstrate the effectiveness of utilizing *S. enterica mddA*::cat^{+} strains in the assessment PPT acetyltransferase function from a variety of bacteria. With the rapidly increasing number of genomes available in databases, the need for experimental validation of predicted gene function becomes more important than ever. Especially since in many cases homologous proteins share a great deal of identity, yet lack the predicted function. For example, the MddA protein of *S. enterica* and *E. coli* are currently annotated as PPT acetyltransferases and they lack such an activity.

ACKNOWLEDGEMENTS

The authors do not have any conflict of interest. The authors acknowledge Dr. Dennis Phillips and Dr. Chau-Wen Chou at the Proteomics and Mass Spectrometry Core Facility (UGA) for assistance with mass spectrometry.

REFERENCES

1. Hentchel KL, Escalante-Semerena JC. 2015. Acylation of biomolecules in prokaryotes: a widespread strategy for the control of biological function and metabolic stress. Microbiol Mol Biol Rev 79:321-346.
2. Davies J, Wright GD. 1997. Bacterial resistance to aminoglycoside antibiotics. Trends Microbiol 5:234-240.
3. Wright GD, Ladak P. 1997. Overexpression and characterization of the chromosomal aminoglycoside 6'-*N*-acetyltransferase from *Enterococcus faecium*. Antimicrob Agents Chemother 41:956-960.

4. Wolf E, Vassilev A, Makino Y, Sali A, Nakatani Y, Burley SK. 1998. Crystal structure of a GCN5-related N-acetyltransferase: *Serratia marcescens* aminoglycoside 3-N-acetyltransferase. Cell 94:439-449.

5. Draker KA, Northrop DB, Wright GD. 2003. Kinetic mechanism of the GCN5-related chromosomal aminoglycoside acetyltransferase AAC(6')-Ii from *Enterococcus faecium*: evidence of dimer subunit cooperativity. Biochemistry 42:6565-6574.

6. Vetting MW, Magnet S, Nieves E, Roderick SL, Blanchard JS. 2004. A bacterial acetyltransferase capable of regioselective *N*-acetylation of antibiotics and histones. Chem Biol 11:565-573.

7. Davies AM, Tata R, Beavil RL, Sutton BJ, Brown PR. 2007. L-Methionine sulfoximine, but not phosphinothricin, is a substrate for an acetyltransferase (gene PA4866) from *Pseudomonas aeruginosa*: structural and functional studies. Biochemistry 46:1829-1839.

8. Carper SW, Willis DG, Manning KA, Gerner EW. 1991. Spermidine acetylation in response to a variety of stresses in Escherichia coli. J Biol Chem 266:12439-41.

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

10. Hentchel KL, Escalante-Semerena JC. 2015. In *Salmonella enterica*, the Gcn5-related acetyltransferase MddA (formerly YncA) acetylates methionine sulfoximine and methionine sulfone, blocking their toxic effects. J Bacteriol 197:314-325.

11. Davies AM, Tata R, Chauviac FX, Sutton BJ, Brown PR. 2008. Structure of a putative acetyltransferase (PA1377) from Pseudomonas aeruginosa. Acta Crystallogr Sect F Struct Biol Cryst Commun 64:338-342.

12. Davies AM, Tata R, Snape A, Sutton BJ, Brown PR. 2009. Structure and substrate specificity of acetyltransferase ACIAD1637 from *Acinetobacter baylyi* ADP1. Biochimie 91:484-489.

13. Wu GB, Yuan MR, Wei L, Zhang Y, Lin YJ, Zhang LL, Liu ZD. 2014. Characterization of a novel cold-adapted phosphinothricin *N*-acetyltransferase from the marine bacterium *Rhodococcus sp* strain YM12. J Mol Catal B-Enzym 104:23-28.

14. Deblock M, Botterman J, Vandewiele M, Dockx J, Thoen C, Gossele V, Movva NR, Thompson C, Vanmontagu M, Leemans J. 1987. Engineering herbicide resistance in plants by expression of a detoxifying enzyme. EMBO J 6:2513-2518.

15. Ogawa Y, Tsuruoka T, Inoue S, Niida T. 1973. Studies on a new antibiotic SF-1293. II. Chemical structure of antibiotic SF-1293. Sci Rep Meijii Kaisha 13:42-48.

16. Circello BT, Eliot AC, Lee JH, van der Donk WA, Metcalf WW. 2010. Molecular cloning and heterologous expression of the dehydrophos biosynthetic gene cluster. Chem Biol 17:402-411.

17. Maughan SC, Cobbett CS. 2003. Methionine sulfoximine, an alternative selection for the *bar* marker in plants. J Biotechnol 102:125-128.

18. Murakami T, Anzai H, Imai S, Satoh A, Nagaoka K, Thompson CJ. 1986. The bialaphos biosynthetic genes of *Streptomyces-Hygroscopicus* - Molecular-cloning and characterization of the gene-cluster. Mol Gen Genet 205:42-50.

19. Berkowitz D, Hushon JM, Whitfield HJ, Jr., Roth J, Ames BN. 1968. Procedure for identifying nonsense mutations. J Bacteriol 96:215-220.

20. Balch WE, Wolfe RS. 1976. New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. Appl Environ Microbiol 32:781-791.

21. Galloway NR, Toutkoushian H, Nune M, Bose N, Momany C. 2013. Rapid cloning for protein crystallography using Type IIS restriction enzymes. Crystal Growth & Design 13:2833-2839.

22. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289-298.

23. Rocco CJ, Dennison KL, Klenchin VA, Rayment I, Escalante-Semerena JC. 2008. Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*. Plasmid 59:231-237.

24. VanDrisse CM, Escalante-Semerena JC. 2016. New high-cloning-efficiency vectors for complementation studies and recombinant protein overproduction in Escherichia coli and Salmonella enterica. Plasmid 86:1-6.

25. Ellman GL, Courtney KD, Andres V, Jr., Feather-Stone RM. 1961. A new and rapid colorimetric determination of acetylcholinesterase activity. Biochem Pharmacol 7:88-95.

26. Griffith KL, Wolf RE, 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.

27. Riddles PW, Blakeley RL, Zerner B. 1983. Reassessment of Ellman's reagent. Methods Enzymol 91:49-60.

28. Paez-Espino AD, Chavarria M, de Lorenzo V. 2015. The two paralogue *phoN* (phosphinothricin acetyl transferase) genes of *Pseudomonas putida* encode functionally different proteins. Environ Microbiol doi:10.1111/1462-2920.12798.

CHAPTER 4

A TOXIN INVOLVED IN *SALMONELLA* PERSISTENCE REGULATES ITS ACTIVITY BY ACETYLATING ITS COGNATE ANTITOXIN, A MODIFICATION REVERSED BY COBB SIRTUIN DEACETYLASE[1]

[1]VanDrisse C.M.[a], Parks A.R.[a], and Escalante-Semerena J.C. 2017. *mBio.* 8:e00708-17.

Reprinted here with permission from the publisher.

[a]Co-first authors

ABSTRACT

Bacterial toxin-antitoxin systems trigger the onset of a persister state by inhibiting essential cellular processes. The TacT toxin of *Salmonella enterica* is known to induce a persister state in macrophages through the acetylation of aminoacyl-tRNAs. Here, we show that the TacT toxin and the TacA antitoxin work as a complex that modulates TacT activity via the acetylation state of TacA. TacT acetylates TacA at residue K44, a modification that is removed by the NAD^+-dependent CobB sirtuin deacetylase. TacA acetylation increases the activity of TacT, down regulating protein synthesis. TacA acetylation altered binding to its own promoter, although this did not change *tacAT* expression levels. These claims are supported by results from *in vitro* protein synthesis experiments used to monitor TacT activity, *in vivo* growth analyses, electrophoretic mobility shift assays and qRT-PCR analysis. TacT is the first example of a Gcn-5 related *N*-acetyltransferase that modifies non-protein and protein substrates.

INTRODUCTION

Acetylation of proteins and small molecules is a conserved mechanism that regulates cellular processes in cells from all domains of life. Members of the Gcn5-related *N*-acetyltransferase (GNAT) protein super family (PF00583) acetylate proteins and non-protein substrates at the expense of acetyl-coenzyme A (AcCoA) (1). In some instances, protein acetylation is reversed by class III, NAD^+-dependent deacetylases, a.k.a. sirtuins (2).

The genome of the human pathogen *Salmonella enterica* subsp. *enterica* sv Typhimurium LT2 (hereafter *S. enterica*) encodes 26 putative GNATs, three of which are annotated as toxin acetyltransferases. These putative toxins are part of type II toxin-antitoxin systems (3, 4), and the genes encoding these proteins appear to comprise separate operons for each toxin-antitoxin pair.

Toxin-antitoxin systems (TA systems) have different physiological functions, with some of them contributing to plasmid stabilization as addiction modules (5-9), while others serve as survival management systems under different stress conditions (10, 11). TA systems have been shown to down regulate essential functions that can trigger the onset of a persister state which allows cells to survive unfavorable conditions without the need to acquire mutations (12-17). Importantly, the slowing of essential metabolic processes by TA module expression may allow for tolerance to antimicrobials, which can lead to recalcitrant infections (18, 19).

Type II TA systems are in most cases a part of two-gene operons that encode a protein toxin and protein antitoxin (20). Typically, the antitoxin neutralizes the toxin until a signal induces selective degradation of the antitoxin, often by stress-induced proteases (16, 21-24). Antitoxin degradation creates a stoichiometric imbalance that releases the toxin, thereby increasing its activity and upregulating its transcription (25, 26). With some exceptions, antitoxins possess a DNA-binding domain that recognizes the operator site of its own promoter resulting in modulation of its own synthesis (11, 27, 28). The *S. enterica* genome encodes three TA systems that include toxins that are homologous to Gcn5-related *N*-acetyltransferases (GNATs). All three of these systems contribute to the onset of a *S. enterica* persister state inside macrophages (29, 30). In addition, Helaine and co-workers showed that the TacT protein of a TA type II system comprised of proteins TacT (STM3651) and TacA (STM3652) acetylates the aminoacyl moiety of several charged tRNAs arresting translation and triggering a persister state (31). The authors suggested that exiting TacT-induced persistence was due to replenishment of tRNA pools through the hydrolysis of the acetylated-amino group off of the aminoacyl-tRNA by the peptidyl-tRNA hydrolase Pth (31).

Here we show that, in addition to acetylating aminoacyl-tRNAs, TacT acetylates its cognate TacA antitoxin with a concomitant increase in TacT activity without complex dissociation, a unique attribute not seen with other TA modules. We suggest that the acetylation state of the TacA antitoxin, not its degradation, rapidly modulates TacT activity. We also show that the NAD^+-dependent CobB sirtuin deacetylase reverses the effect that acetylated TacA has on TacT activity, and suggest that reversible TacA acetylation plays a key role in exiting the persister state. In addition, we present evidence that the TacA antitoxin binds to the *tacA-tacT* promoter when when in complex with TacT. The involvement of sirtuin-dependent reversible lysine acetylation (sRLA) frames the persister state of *S. enterica* within the carbon and energy statuses of the cell.

RESULTS

The absence of TacA antitoxin extends the lag time before the onset of exponential growth in cells with higher levels of *tacT*$^+$ expression.

Consistent with published data (31), ectopic expression of *tacT*$^+$ in *S. enterica tacA*::*cat*$^+$ or *tacAT*::*cat*$^+$ mutant strains delayed the onset of exponential growth when cells were grown on minimal medium (Fig. 4.1) due to increased TacT-mediated aminoacyl-tRNA acetylation. This delay was shortened when cells were grown in rich medium, suggesting that nutrient availability played a role in the observed phenotype. Notably, the extended lag phase did not have an effect on the final cell density or growth rate of the cultures (Fig. 4.1). The phenotype of *tacA*::*cat*$^+$ strains in which *tacT*$^+$ was overexpressed (*i.e., tacA*::*cat*$^+$ / pTacT^{WT} or *tacAT*::*cat*$^+$ / pTacT^{WT}) was corrected by *in trans* expression of *tacA*$^+$ (Fig. 4.1), indicating that the observed phenotype was dependent on the absence of the TacA antitoxin.

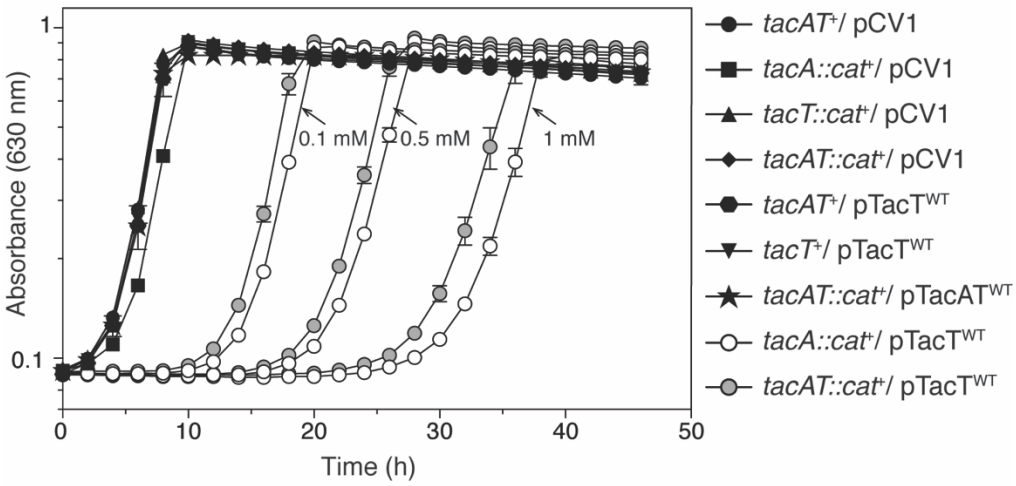


Figure 4.1. Growth of *S. enterica* cells lacking *tacA* and harboring *tacT*$^+$ plasmids have growth delays associated with increased expression of *tacT*. Genes coding for TacT (pTacT) or TacAT (pTacAT) were cloned into a complementation vector and introduced into strains with genetic backgrounds identified by the symbols in the graph. Overnight cultures were grown in NB rich medium. Cells were sub-cultured (1% v/v) and grown in NCE minimal medium supplemented with glycerol (22 mM), ampicillin, and transcription of plasmids containing *tacT*$^+$ were induced using levels of arabinose designated by arrows (grey and white circles). All strains with solid black symbols contained 0.15 mM arabinose. Growth curves were obtained in technical triplicates of biological triplicates and was repeated three times using a microplate reader (BioTek instruments). Error bars represent standard deviation of technical triplicates.

TacT acetylates residue K44 of TacA.

As shown by others, TacA co-purified with TacT, forming a complex. TacA and TacT proteins were synthesized from a vector containing the coding sequence for both proteins. Overexpression of both genes allowed for simultaneous synthesis and folding of each protein, leading to a stable TacAT complex. All toxin-antitoxin complexes (wild-type and variant proteins) reported herein, were co-purified as described under Materials and Methods. Our assessment of the purity of the proteins used in this study can be found in supplemental Fig. 4.2 (lanes 7, 8). Results of gel permeation chromatography experiments revealed the TacAT complex was a dimer of dimers (Fig. 4.3). When purified TacA and TacT proteins were incubated with [1-^{14}C]AcCoA, TacA was acetylated (Fig. 4.4A, lane 2), a modification that required TacT (Fig. 4.4A, lane 2 *vs* 3). LC/MS/MS peptide fingerprinting analysis of acetylated TacA (TacA^{Ac}) showed that residues K12, K44, and K83 were acetylated to varying degrees (Fig. 4.5). Figure 4.5 shows peptide masses with different *m/z* ratios used search the database using MASCOT software. Analyses of areas under the peaks for peptides corresponding to acetylated K12, K44, and K83 showed a 2.64-fold increase for K12, 18.86-fold increase for K44, and a 3.12-fold increase for K83.

Correction of the translation initiation codon of TacA.

In the *Salmonella* genome, *tacAT* comprise an operon with *tacA* being promoter proximal, with the last 13 bp of *tacA* overlapping with the *tacT* coding sequence. The 5' region of *tacA* contains the annotated start methionine and an additional methionine at position 8. This information was relevant, because the above-mentioned peptide finger printing analysis did not detect the first seven amino acids (M1-L7, grey residues, Fig. 4.5), of the annotated primary sequence of TacA, therefore, we considered the possibility that the true start methionine of TacA was residue M8. This misannotation became relevant during the purification of TacAT complexes,

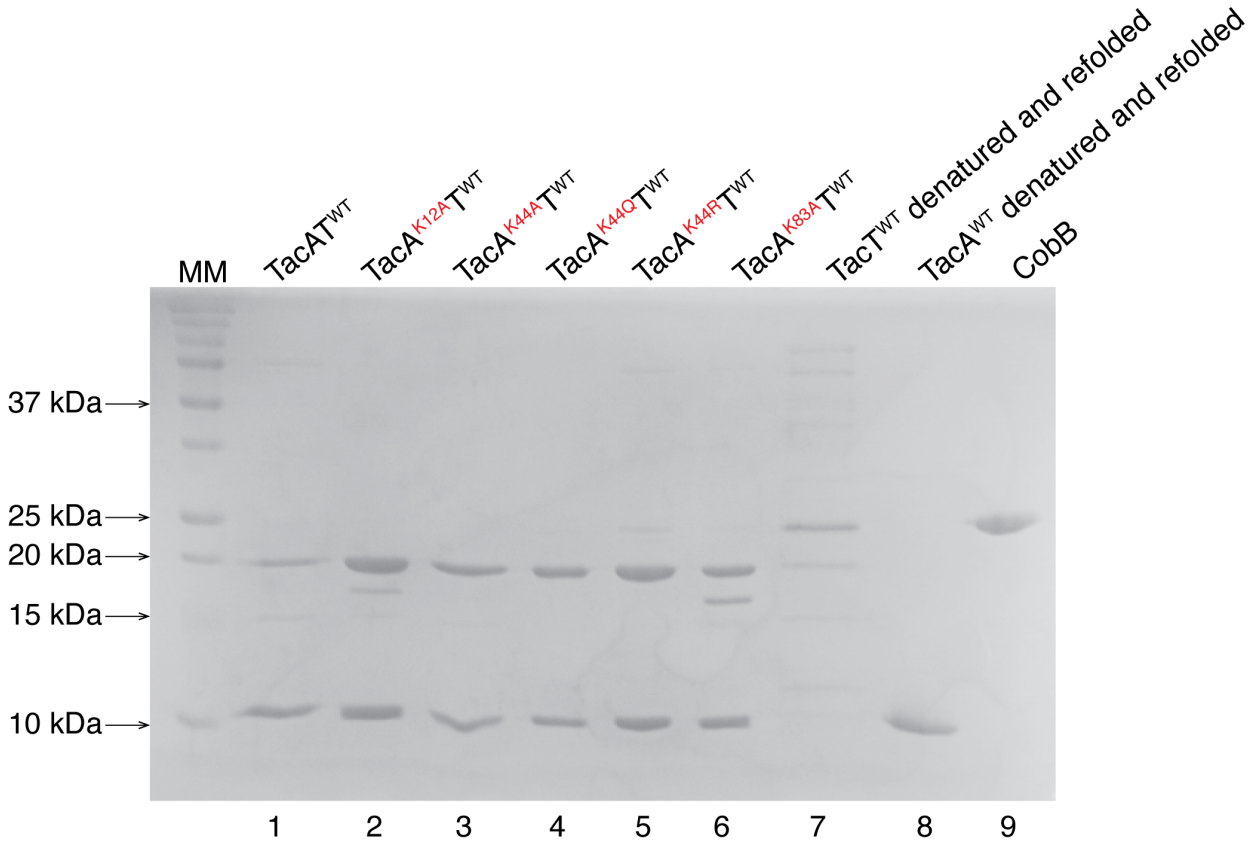


Figure 4.2. Percent purity of proteins used in these studies. A 15% (w/v) SDS-PAGE gel loaded with ~1 μg of each protein that is designated above each well. Percent purities are as follows, TacAT^{WT} : 88%, TacA^{K12A}TacT^{WT} : 97%, TacA^{K44A}TacT^{WT} : 99%, TacA^{K44Q}TacT^{WT} : 99%, TacA^{K44R}TacT^{WT} : 99%, TacA^{K83A}TacT^{WT} : 85%, TacT : 10%, TacA: 99%, CobB: 99%. Percent purity was calculated by diluting proteins (3, 2, 1, 0.5, 0.25 μg) on a separate gel and bands were quantified as percentages using ImageQuant™ TL. Contaminants of TacT denatured and refolded prep include: ClpB (ATP-dependent chaperone), HtpG (molecular chaperone), MdaA (NADPH nitroreductase), elongation factor Tu, peptidyl-prolyl *cis-trans* isomerase, Ferric uptake regulation protein, transposase, and YdaS (uncharacterized protein).

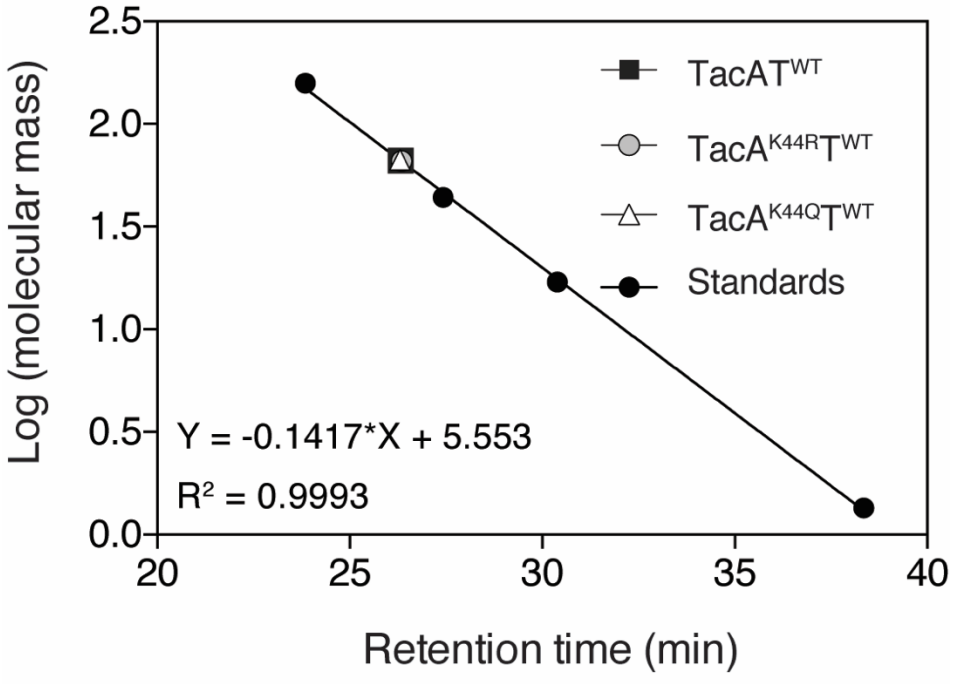


Figure 4.3. Size exclusion chromatography of purified complexes. The molecular masses of $TacAT^{WT}$, $TacA^{K44R}TacT^{WT}$, and $TacA^{K44Q}TacT^{WT}$ were determined by size exclusion chromatography using a Superose 12 10/300 GL column as detailed under Materials and Methods. The molecular mass standards (white circles) were thyroglobulin (bovine; 670,000 Da), gamma globulin (bovine; 158,000 Da), ovalbumin (chicken; 44,000 Da), myoglobin (horse; 17,000 Da), and vitamin B_{12} (1,350 Da).

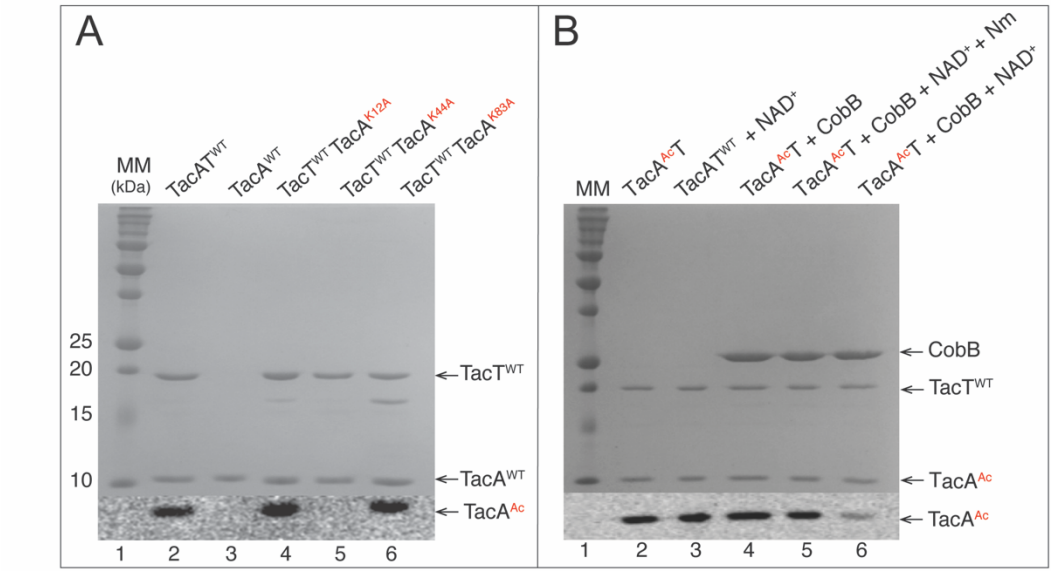


Figure 4.4. A. TacA residue K44 is acetylated by TacT. Wild-type TacA-TacT complex (lane 2) or variant (lanes 4, 5, 6) complexes and TacA (lane 3) were incubated with [1-^{14}C]-acetyl-CoA. Proteins were separated by SDS-PAGE and stained with Coomasie Blue to visualize proteins. Precision Plus Protein™ (Bio-Rad) standard was used as a molecular marker (MM). Acetylation was visualized by phosphor imaging. Extra bands in lanes 4 and 5 were analyzed via mass spectrometry and were identified as TacT, which we presumed was cleaved or lacked the hexahistidine tag. **B. TacA^{Ac} is deacetylated by CobB**. The reaction mixture used as positive control contained wild-type TacAT (TacAT^{WT}) complex and [1-^{14}C]-acetyl-CoA (lane 2); positive control mixture + NAD$^+$ (lane 3); positive control mixture + CobB (lane 4); positive control mixture + CobB + NAD$^+$ + Nm (lane 5); or positive control mixture + CobB + NAD$^+$ (lane 6). Proteins were separated by SDS-PAGE and stained with Coomasie Blue to visualize proteins. Precision Plus Protein™ (Bio-Rad) standard was used as a molecular marker (MM, lane 1). Acetylation was visualized by phosphor imaging. On panels A and B, the thin black line at the bottom of the figures separates the SDS-PAGE gel from the phosphor image. CobB: sirtuin deacetylase, NAD$^+$: nicotinamide adenine dinucleotide, Nm: nicotinamide, TacT: toxin, TacA: antitoxin.

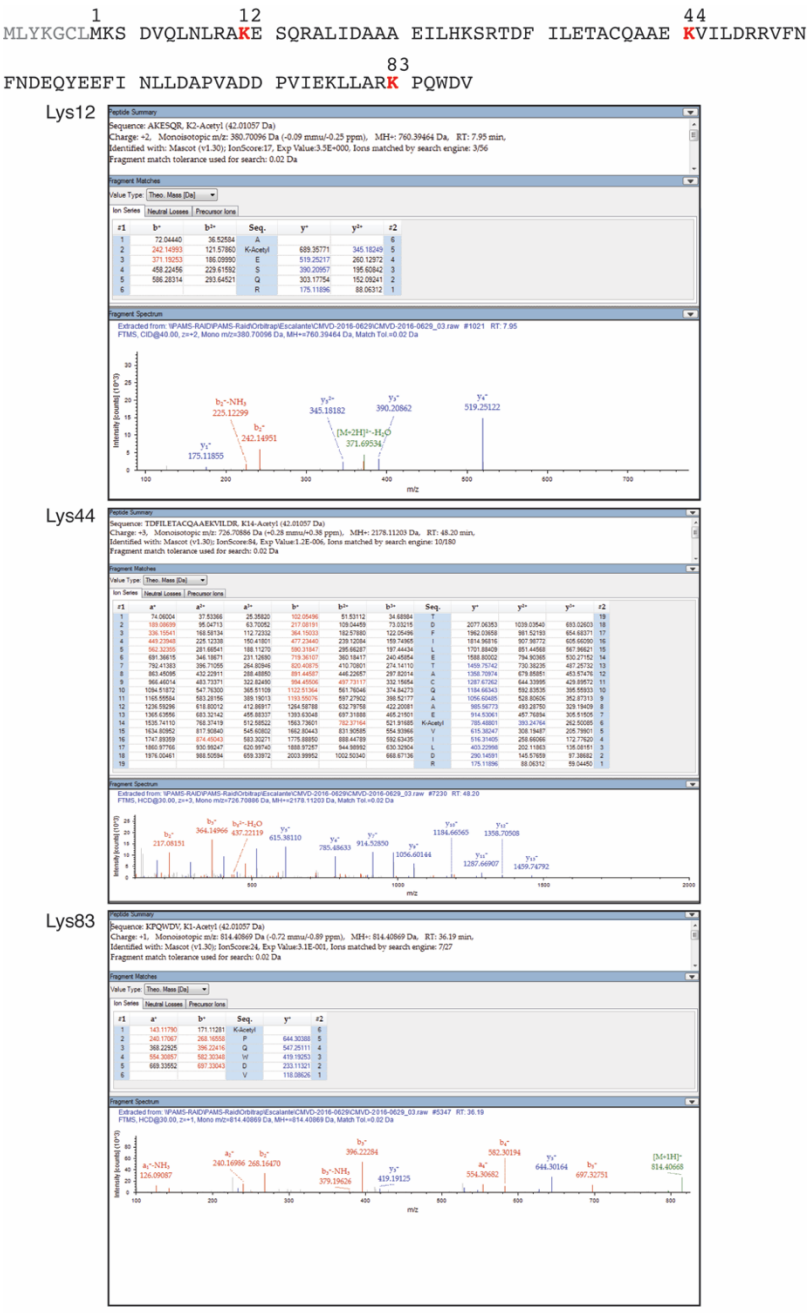


Figure 4.5. Mass spectrometry analysis of acetylated TacA. Sequence above represents annotated TacA protein sequence. Residues in grey represent amino acids not detected by LC/MS MS and therefore the start methionine was repositioned to begin with black sequence. Residues identified to be acetylated in TacAT^{WT} samples containing 1 mM acetyl-CoA are highlighted in red. *b* ions is the series of fragments that extend from the amino terminus; *y* ions are the series of ions that extend from the *C*-terminus. MASCOT software (http://www.matrixscience.com) was the online search engine used to identify peptides on the basis of their masses. Ion series a, b, y are expected ion types. The subscript of a, b or y ions represents the number of residues in the peptide; the superscript represents the peptide charge.

because the nucleotides encoding the first seven amino acids and the *S*-tag fused to TacA were not translated, resulting in co-elution of tag-less, native TacA with H_6-TacT. The availability of tagless TacA became useful for the performance of DNA binding experiments, because the DNA-binding domain of TacA was on its *N*-terminus. For the experiments mentioned above and hereto forth, the translation start codon of TacA was reassigned to the ATG encoding M8. That is, residue M8 became M1, and the numbering for the rest of the residues was modified accordingly. Consequently, residues K19, K51 and K90 became K12, K44, K83, respectively. The adjusted numbering was used throughout these studies.

Residue K44 is a key acetylation site in TacA, a modification reversed by the NAD^+-dependent CobB sirtuin deacetylase.

To validate the putative acetylation sites, TacA variants with substitutions at positions K12, K44 and K83 were isolated. As shown in Figure 4.4A, transfer of the acetyl moiety of [^{14}C]-AcCoA was not observed when TacT was in complex with $TacA^{K44A}$, indicating that either K44 was the only acetylation site in TacA that was modified by TacT, or that K44 acetylation triggered K12 and K83 acetylation by TacT. The latter scenario was not pursued. Instead we focused on the question of whether or not K44 acetylation was reversible. In *Salmonella,* the only known protein lysine deacetylase is the NAD^+-dependent CobB sirtuin. This result raised the question of whether CobB could deacetylate $TacA^{Ac}$. To test this idea $TacA^{[14C]Ac}$ was incubated with CobB, NAD^+, CobB + NAD^+, or CobB + NAD^+ + nicotinamide (Nm). Under the conditions tested, CobB and NAD^+ deacetylated ~70% of $TacA^{Ac}$ (Fig. 4.4B, lane 6 *vs* lanes 2, 3, 4, 5). As expected, CobB activity was inhibited by Nm (Fig. 4.4B, lane 5 *vs* 6), a result consistent with CobB dependent deacetylation.

TacA acetylation enhances TacT activity *in vitro*.

TacA variants were constructed to determine whether or not acetylation of TacA residue K44 had an effect on TacT-dependent arrest of mRNA translation. TacA ariants with substitutions at position K44 that mimicked acetylation (*i.e.,* K44Q) or deacetylation (*i.e.,* K44R) (2) were isolated in complex with $TacT^{WT}$. No effect on the formation or stability of complexes between $TacT^{WT}$ and TacA variants was detected using gel filtration chromatography (Fig. 4.3). TacT was purified from homogeneous TacAT complex by denaturation and re-folding, as described in the Materials and Methods.

In vitro protein synthesis experiments were performed with $TacAT^{WT}$, $TacA^{K44Q}TacT^{WT}$ and $TacA^{K44R}TacT^{WT}$ complexes or $TacT^{WT}$ alone. Proteins were incubated with or without AcCoA in a cell-free protein synthesis system that contained all machinery necessary for transcription and translation from a synthesized DNA product. DNA that is added to the reactions codes for dihydrofolate reductase (DHFR) and its synthesis is used as a reporter of mRNA translation; DHFR synthesis was monitored using SDS-PAGE (32) (~18 kDa, arrow, Fig. 4.6). $TacT^{WT}$ toxin was expected to acetylate aminoacyl-tRNAs resulting in DHFR protein synthesis inhibition (31). To quantify differences in DHFR synthesis in reaction mixtures containing or lacking AcCoA, intensities of three bands in each lane (asterisks in Fig. 4.6), were expressed as percentages of the band intensity of DHFR. For example, the level of DHFR synthesis was compared in reaction mixtures containing $TacAT^{WT}$ with or without AcCoA (Fig. 4.6, lanes 4, 5). Control experiments for these studies included conditions for maximal DHFR synthesis (Fig. 4.6, lane 2), with the baseline for the production of DHFR being established by a reaction mixture devoid of DNA. As expected, no DHFR was synthesized (<1%) in the absence of added DHFR encoding DNA (Fig. 4.6, lanes 2 *vs* 3). Two reaction mixtures lacking $TacA^{WT}$, with or without

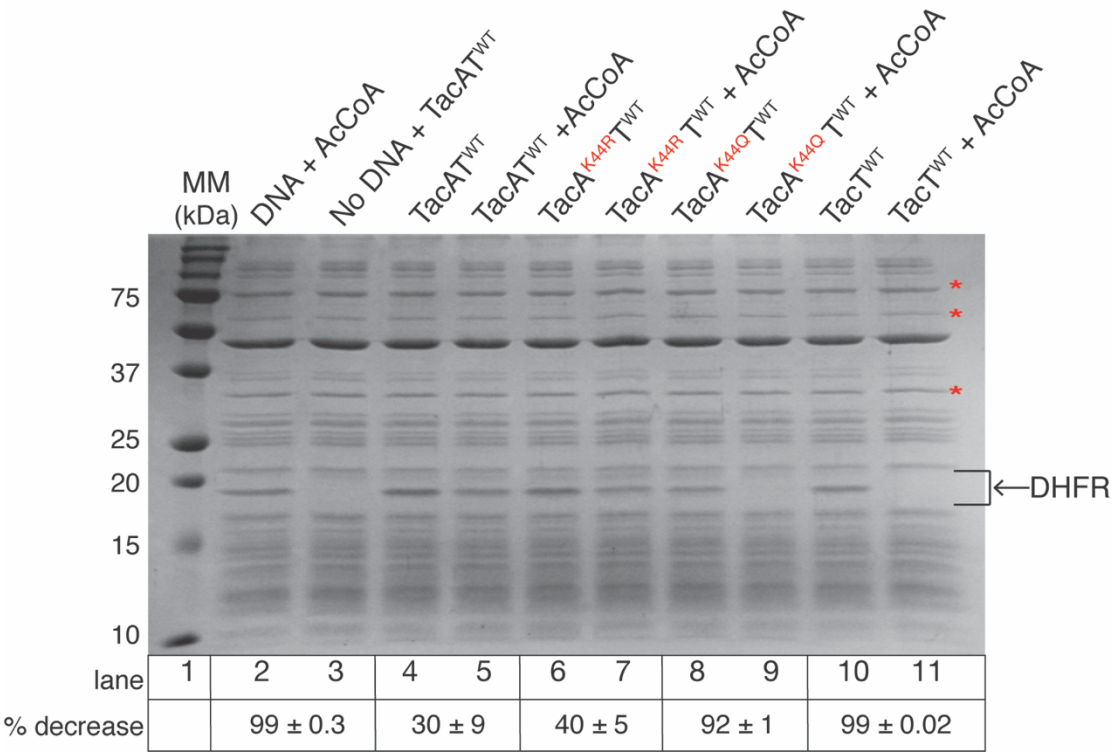


Figure 4.6. *In vitro* DHFR protein synthesis used to quantify aminoacyl-tRNA acetylation by TacT shows TacA acetylation variants increase TacT-mediated aminoacyl-tRNA aceylation. SDS-PAGE gel showing *in vitro* synthesis of dihydrofolate reductase (DHFR, indicated by arrow), in reactions mixtures containing the additions indicated above each lane. Precision Plus Protein™ (Bio-Rad) standard was used as a molecular marker (MM, lane 1). Lane 2 represents a positive control for DHFR production, where the reaction mixture contained DHFR DNA (100 ng) and acetyl-CoA to ensure acetyl-CoA stocks did not interfere with reaction components. The reaction mixture in lane 3 lacked DNA and contained TacAT^{WT} complex to control for any added bands that might interfere with interpretations. Lanes 4-9, reaction mixtures that contained TacAT^{WT} or variant complexes (indicated by labels above each lane) with or without acetyl-CoA and DNA was added to lanes 4-11. Lanes 10 and 11, positive control for aminoacyl-tRNA acetylation. These reaction mixtures contained TacT with or without acetyl-CoA. All reactions were incubated for 2 h at 37°C. Details of the procedure can be found under in materials and methods. Asterisks represent bands used to normalize the intensity of the DHFR protein within each lane (ImageQuant™ v5.2 software). Comparison of DHFR band intensity to bands with asterisks was averaged and the standard deviation of these three numbers was calculated. The percent decrease with SD for each sample is indicated below the corresponding lanes.

AcCoA added, showed $TacT^{WT}$-, AcCoA-dependent arrest of DHFR synthesis (Fig. 4.6, lanes 10 *vs* 11), indicating that the $TacT^{WT}$ protein was active.

DHFR synthesis was reduced by ~30% in reaction mixtures containing $TacAT^{WT}$ in the presence of AcCoA (Fig. 4.6, lanes 4 *vs* 5). This result was unexpected because, to our knowledge, this is the first report of a toxin suggested to be active in complex with its cognate antitoxin. Reaction mixtures containing $TacA^{K44R}T^{WT}$ (deacetylation mimic variant) and acetyl-CoA showed a 40% reduction in DHFR synthesis (Fig. 4.6, lanes 6 *vs* 7). Although it appears there are differences in aminoacyl-tRNA acetylation by the $TacAT^{WT}$ and $TacA^{K44R}T^{WT}$ proteins, these differences are not statistically significant (see percentages as standard deviations, Fig. 4.6). In sharp contrast, DHFR synthesis was reduced by 92% in reaction mixtures containing the acetylation mimic variant, $TacA^{K44Q}TacT^{WT}$ complex, and AcCoA (Fig. 4.6, lanes 8 *vs* 9), suggesting that the $TacA^{K44Q}$ variant enhanced the acetylation of aminoacyl-tRNA by $TacT^{WT}$ (Fig. 4.6, lanes 8 *vs* 9). These results strongly suggested that $TacA^{WT}$ acetylation by $TacT^{WT}$ up-regulated $TacT^{WT}$-dependent acetylation of aminoacyl-tRNAs. A summary of the above results is presented as percentages underneath SDS-PAGE gel (Fig. 4.6).

TacA acetylation enhances TacT activity *in vivo*.

To validate *in vitro* results *in vivo*, *tacA* alleles encoding TacA variants with single-amino acid substitutions at position K44 were introduced by site-directed mutagenesis into a plasmid carrying $tacAT^{+}$ (pTacAT-1). One TacA variant within the operon had a K44Q substitution ($TacA^{K44Q}$) for the purpose of mimicking acetylation (pTacAT-10), another variant had a K44R ($TacA^{K44R}$) to mimic deacetylation (pTacAT-9). The resulting plasmids, which also carried the $tacT^{+}$ allele, were individually introduced into *tacAT*::cat^{+} strains. As shown in Figure 4.7, *tacAT*::cat^{+} strains that synthesized $TacA^{K44Q}TacT^{WT}$ complex had a striking growth phenotype

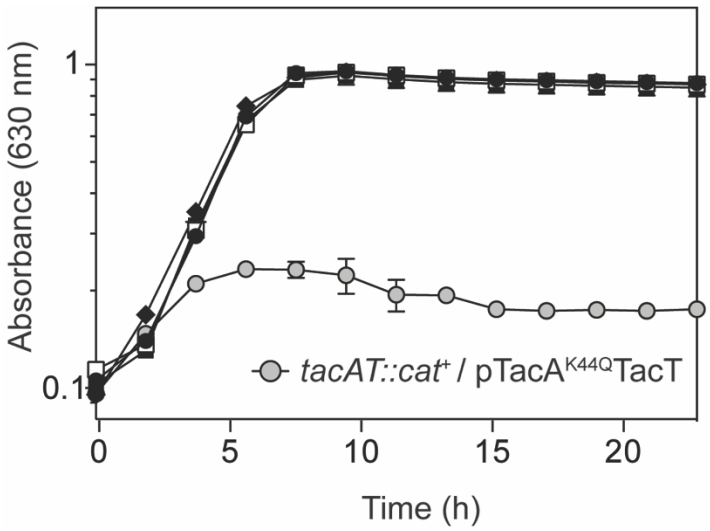


Figure 4.7. TacA variant mimicking acetylation in complex with TacT in *S. enterica* *tacAT*::*cat*$^{+}$ strains causes growth defect. Genes coding for TacT^{WT} (pTacT), TacAT^{WT} (pTacAT^{WT}), TacA^{K44R}TacT^{WT} (pTacA^{K44R}TacT^{WT}), TacA^{K544Q}TacT^{WT} (pTacA^{K44Q}TacT^{WT}), were cloned into a complementation vector pCV1) and introduced into strain backgrounds indicated by the following symbols: *tacAT*::*cat*$^{+}$ / pCV1 (empty cloning vector control), solid circles; *tacAT*::*cat*$^{+}$ / pTacT^{WT}, open squares; *tacAT*::*cat*$^{+}$ / pTacAT^{WT}, solid squares; *tacAT*::*cat*$^{+}$ / pTacA^{K44R}T^{WT}, diamonds; *tacAT*::*cat*$^{+}$ / pTacA^{K44Q}T^{WT}, gray circles. Overnight cultures were grown in NB rich medium. Cells were sub-cultured (1% v/v) and grown in NCE minimal medium supplemented with glycerol (22 mM), ampicillin. Expression of plasmid-borne genes was induced with 25 μM L(+)-arabinose. Growth curves were obtained in technical triplicates of biological triplicates and was repeated three times using a microplate reader (BioTek instruments). Error bars represent standard deviation of technical triplicates.

compared to strains making TacA^{K44R}TacT^{WT} (Fig. 4.7, grey circles, diamonds). This result was consistent with an increase in aminoacyl-tRNA acetylation by TacT^{WT} that slowed down mRNA translation causing the observed phenotype. We note that the growth arrest of strains making TacA^{K44Q}TacT^{WT} was more severe than the growth delay of strains expressing *tacT*$^{+}$ alone and is most likely due to lower inductions of plasmids coding for TacA^{K44Q}TacT^{WT} causing this phenotype (*i.e.* overexpression of *tacT* alone at 25 μM arabinose is not a high enough induction to cause a growth delay). This is most likely due to stability of the toxin when it is in or not in complex with its antitoxin and these results are further discussed below.

Acetylation of residue K44 of TacA alters the DNA-binding of the protein.

Electrophoretic mobility shift assays (EMSAs) were performed to determine whether modifications of K44 would affect the DNA-binding activity of TacA. A 157-bp 6-FAM 5'-labeled probe (-163 to -6, probe 1) upstream of the corrected ATG transcription initiation codon for *tacAT* was chosen for probe design. TacAT^{WT} complex bound to this probe and this binding was specific because TacAT did not bind to the promoter for *argS*, a negative control (Fig. 4.8). To narrow down the minimal sequence necessary for TacAT.binding, two probes within the 157-bp region upstream the putative ATG codon for *tacAT* were designed. One probe was 71-bp long (-163 to -92, probe 2), and the other was 75-bp (-81 to -6, probe 3) long (Fig. 4.9, probe 2 and 3). There was no shift seen with probe 2, but the electrophoretic mobility of the 75-bp probe (probe 3) changed in the presence of the TacAT^{WT} complex (Fig. 4.9).

TacA^{WT} or TacT^{WT} alone did not bind to probe 3 (Fig. 4.10A). This result suggested that TacA needed TacT to bind to DNA. An EMSA was conducted in which TacA and TacT protein were present to test whether or not they could reform a complex capable of binding to probe 1. No binding was detected at 5-fold excess TacA and 50-fold excess TacT, nor was DNA binding

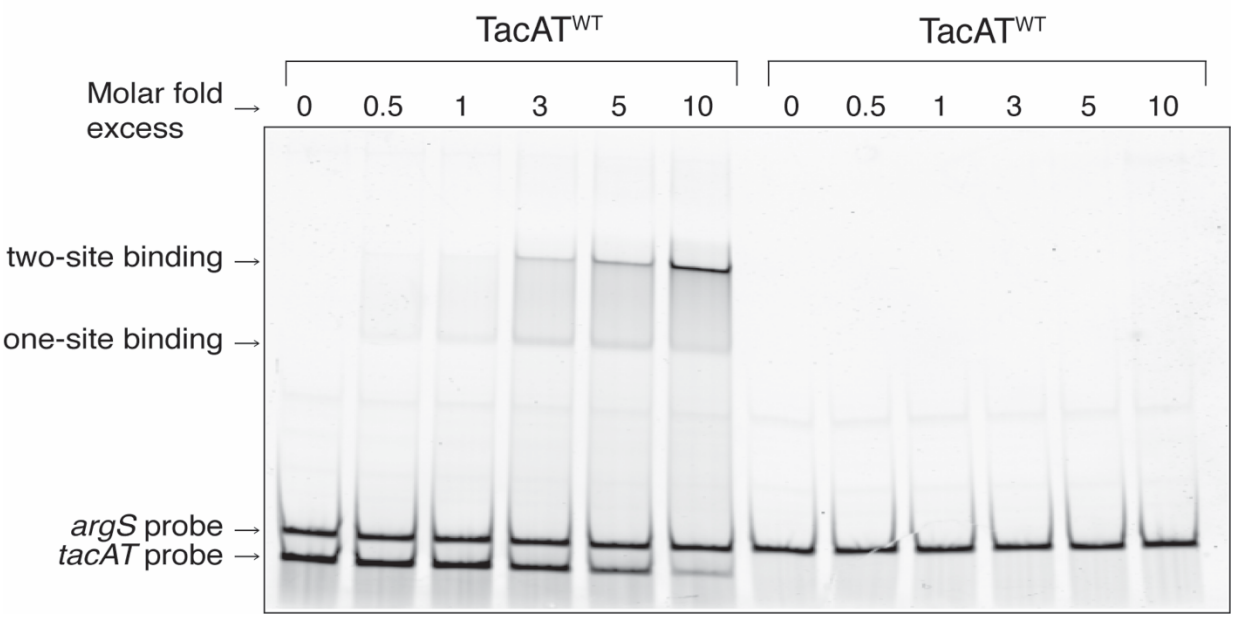


Figure 4.8. TacAT^{WT} complexes bind specifically to the *tacAT* promoter. The specificity of binding of TacAT^{WT} complexes to the *tacAT* promoter was analyzed by electrophoretic mobility shift assays using 6-FAM 5'-labeled probes. Left half: Probe 1 (157 bp, 0.48 pmol, position -163 to -6) and competitor DNA (*argS* promoter DNA, 0.38 pmol, 196 bp) were incubated together with increasing concentrations of TacAT^{WT} complex (0.24, 0.48, 1.45, 2.4, and 4.8 pmol protein to DNA). Right half: Promoter *argS* (0.38pmol) incubated with TacAT^{WT} at increasing molar fold excess (0.54, 1.08, 3.25, 5.41, and 10.8 pmol).

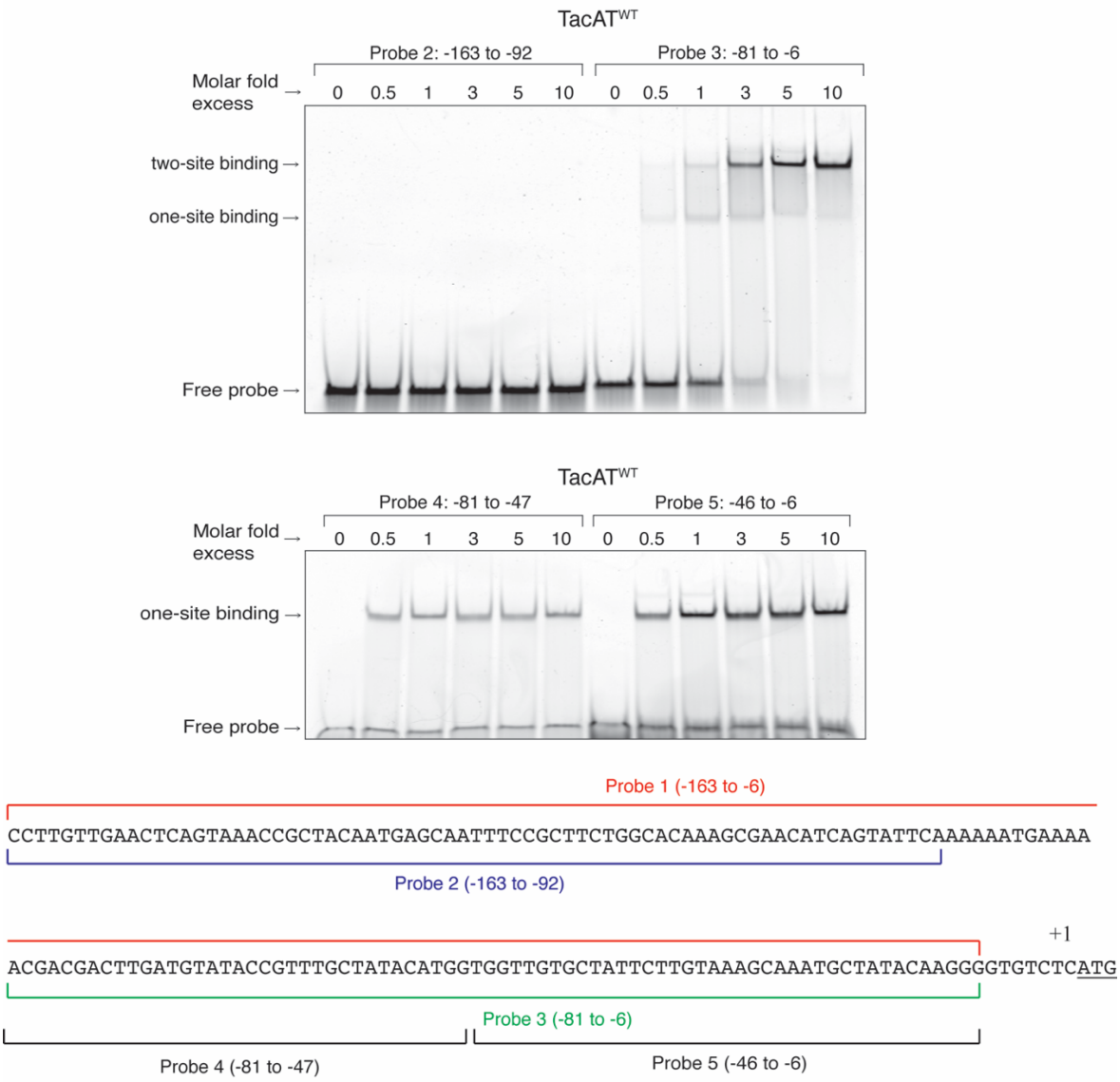


Figure 4.9. TacA-TacT complexes bind to a specific region within the *tacAT* promoter. An attempt to narrow down the binding region of the TacAT^{WT} complex within the *tacAT* promoter was conducted using electrophoretic mobility shift assays with 6-FAM 5'-labeled probes. Top figure, left: A fragment of DNA consisting of the upstream 71 base pairs of Probe 1 called Probe 2 (1.06 pmol, 71 bp) was incubated with increasing concentrations of TacAT^{WT} complex (Protein added to probe 2: 0.53, 1.06, 3.2, 5.3, 10.6 pmol). Top figure, right: TacA^{WT}TacT^{WT} complex (0.5, 1.01, 3.03, 5.05, 10.1 pmol) was incubated with the downstream 75 base pairs of Probe 1 called Probe 3 (1.01pmol). Bottom figure: To determine if two binding sites exist within Probe 3, the sequence of Probe 3 was then divided in half to form Probe 4 (upstream 35 bp region, 2.16 pmol) and Probe 5 (downstream 40 bp region, 1.89 pmol). TacAT^{WT} complex was incubated with either Probe 4 or Probe 5 to show binding to two different sites at the molar fold excess of 1.08, 2.16, 6.49, 10.8, and 21.6 pmol Protein to Probe 4 and 0.95, 1.89, 5.68, 9.5, and 18.9 pmol Protein to Probe 5. DNA sequence represents -163 bp to ATG of *tacA*. ATG start codon matches correct reading frame as mentioned in Results section.

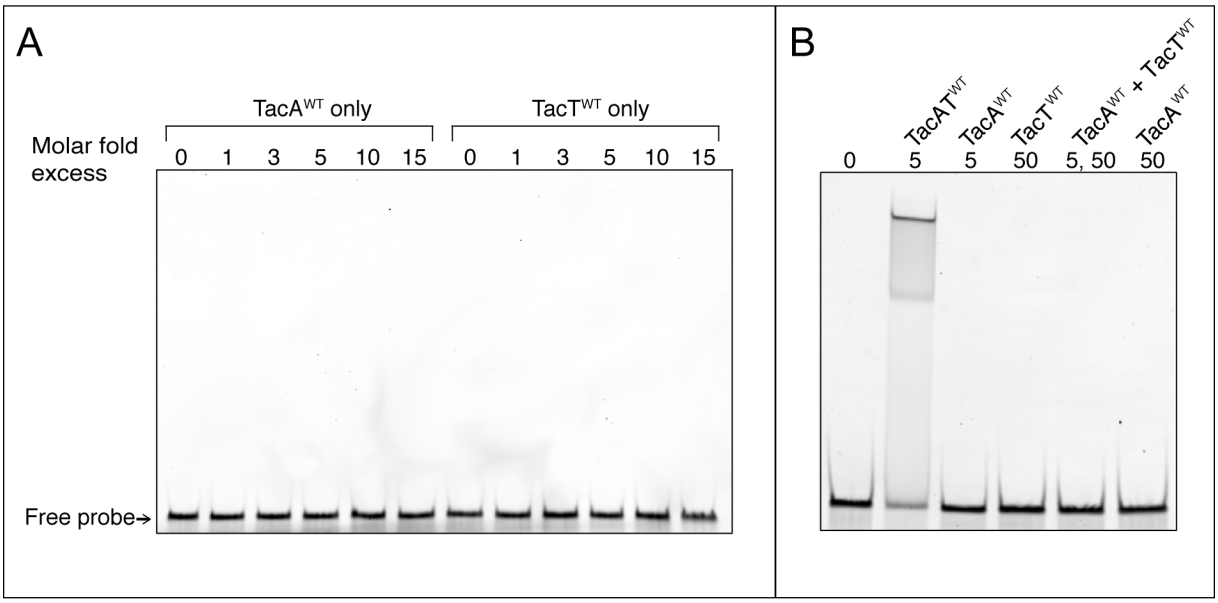


Figure 4.10. TacA and TacT alone do not bind to the *tacAT* promoter. (A) To test whether or not TacA^{WT} could bind to its own promoter in the absence of TacT^{WT}, probe 3 and TacA^{WT} protein were incubated together at 1.01, 3.03, 5.05, 10.1, and 15.15 molar fold excess protein to probe (1.01pmol). As a control to ensure that TacT^{WT} alone had no DNA binding activity, TacT^{WT} protein was incubated with probe 3 at the same molar fold excess as stated above. (B) To test whether TacA and TacT^{WT} protein could re-form a complex and bind to DNA, probe 1 (0.24 pmol) was incubated with either 5 molar-fold excess TacAT^{WT} complex, 5-fold excess TacA^{WT}, 50-fold excess TacT^{WT}, or mixed 5-fold excess TacA^{WT} and 50-fold excess TacT^{WT} together. Fifty-fold excess TacT^{WT} was used because TacT^{WT} was 10% pure (Fig. 5.2).

observed when TacA protein was incubated with probe 1 at 50-fold excess (Fig. 4.10B). We concluded that refolded TacT was active based on the results presented in Figure 4.6. However, we cannot rule out the possibility that TacA did not refold correctly, thus preventing DNA binding. It is also possible that the complex could not reform after TacA and TacT were separated.

We observed two different probe shifts (Fig. 4.11A, labeled with one- and two-site binding), which were not due to a mixed population of complex (*i.e.* acetylated *vs* non-acetylated) since a $TacA^{K44R}TacT^{WT}$ complex (*i.e.,* not acetylatable variant, but positively charged substitution of K44) produced the same two shifts (Fig. 4.11B). To ascertain whether the observed two bands reflected the presence of two binding sites within probe 3 the probe was split into two pieces, probe 4 and 5 (Fig. 4.9, sequence representation). Probe 4 (-81 to -46) or probe 5 (-46 to -6) were mixed with $TacAT^{WT}$ complex, and interactions assessed by EMSAs (Fig. 4.9). The two bands observed when probe 1 and probe 3 were used (Fig. 4.11, Fig. 4.9) were not observed when probes 4 and 5 were used (Fig. 4.9), indicating that only one binding site was present in each probe. Work on the MsqRA TA system in *E. coli* showed that two band shifts occur when two binding sites were present (Brown 2013). Such a scenario is consistent with our observations (Fig. 4.11A) in that the lower band assigned to TA complex bound to either site, and the upper band reflecting interactions of complexes bound to two sites (Fig. 4.11A, cartoon representation).

In contrast to the $TacA^{K44R}TacT^{WT}$ complex, the $TacA^{K44Q}TacT^{WT}$ complex (*i.e.,* mimicking acetylated K44) did not display the second band, in support of the presence of two sites (Fig. 4.11). These data suggested that residue K44 did not directly affect the DNA-binding activity of TacA, but that acetylation altered TacA recognition of or binding to the *tacAT* promoter. It was further validated that K44 was not directly involved in DNA binding because the $TacA^{K12A}TacT^{WT}$

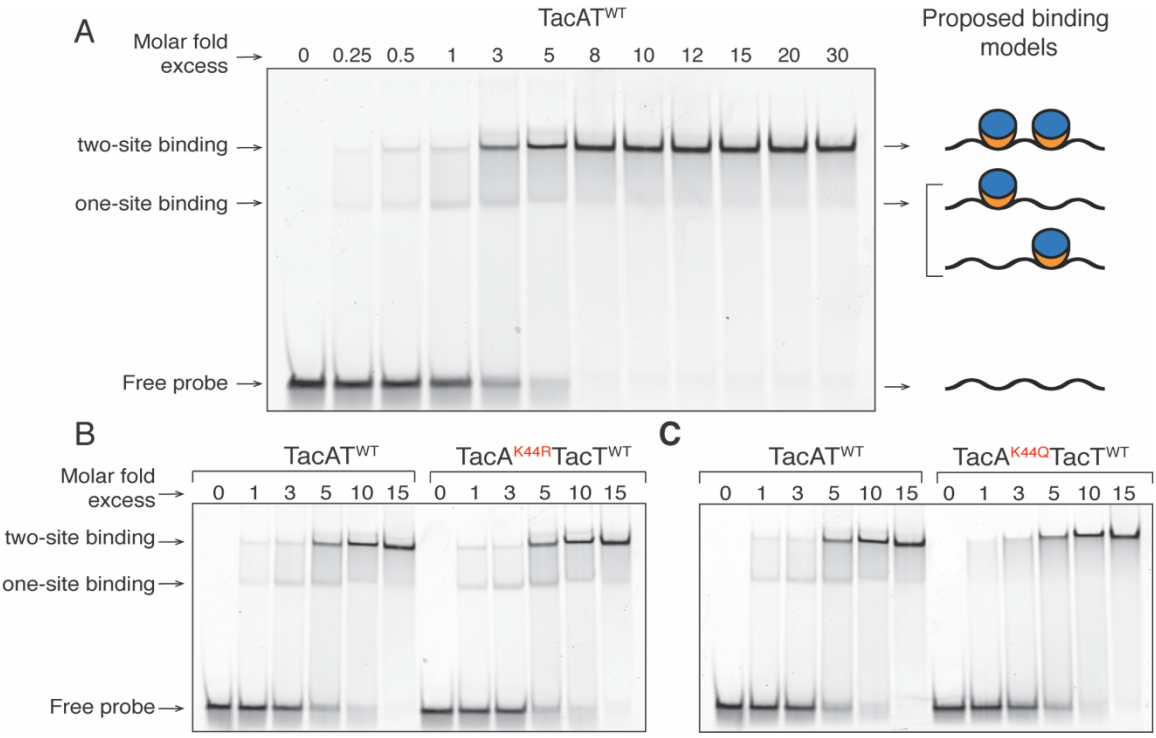


Figure 4.11. Binding of TacA^{WT}TacT^{WT} to the *tacAT* promoter differs from binding of TacA^{K44Q}TacT^{WT}. Binding of increasing concentrations of TacAT complexes to the *tacAT* promoter was assessed by electrophoretic mobility shift assays using 6-FAM 5'-labeled probes. A. Probe 3 (75 bp, 1.01pmol, position -81 to -6) was incubated with increasing concentrations of TacAT^{WT} protein (0.25, 0.5, 1.01, 3.03, 5.05, 8.08, 10.1, 12.12, 15.15, 20.2, and 30.3 pmol). The diagram on the right side of the figure depicts a model that could explain the occurrence of the two shifted bands seen (TacT, blue ovals; TacA, orange crescents). B. Probe 3 (1.06 pmol) incubated with increasing concentrations of either TacAT^{WT} or TacA^{K44R}TacT^{WT} (molar excesses of both complexes: 0.53, 1.06, 3.2, 5.3, and 10.6 pmol). C. Probe 3 (1.06 pmol) incubated with TacAT^{WT} or TacA^{K44Q}TacT^{WT} at same molar fold-excess as Figure 4.11B.

complex had DNA binding activity that was abrogated (Fig. 4.12), suggesting the *N*-terminus included the DNA-binding domain of TacA. In contrast to this finding, binding of $TacA^{K83A}TacT^{WT}$ and $TacAT^{WT}$ complexes to probe 3 was very similar, suggesting that residue K83 and therefore the *C*-terminus was not involved in TacA binding to DNA. A predicted structural representation of these residues can be seen in Figure 4.13.

To determine whether the altered binding to DNA by TacA variants in complex with $TacT^{WT}$ had an effect on repression of *tacAT* transcription, chromosomal mutations coding for K44R or K44Q TacA variants were constructed as described in materials and methods. Cells coding for $TacA^{WT}TacT^{WT}$, $TacA^{K44R}TacT^{WT}$, $TacA^{K44Q}TacT^{WT}$, or cells lacking TacT were grown to mid-log on minimal medium supplemented with 22mM glycerol and total RNA was extracted as described in Materials and Methods. We performed RT-qPCR with total RNA to measure the differences in level of mRNA transcript of the *tacA* gene in the mutants as compared to the $tacAT^{+}$ strain. We observed a large increase in *tacA* transcript (28.6-fold) in the strain in which the *tacT* gene was deleted and only *tacA* remained (Fig. 4.14). This suggested that the native TacAT complex repressed the operon under this condition and that the presence of TacT was needed for TacA-mediated repression. When measuring *tacA* transcript in the chromosomal variants, we did not detect significant transcript differences between the $TacA^{K44R}TacT^{WT}$ or $TacA^{K44Q}TacT^{WT}$ strains as compared to the parent strain. These results suggested that the altered DNA binding activity of $TacA^{K44Q}T$ variants seen *in vitro* might either be an artifact of the assay or that an additional factor for *tacAT* expression might be needed, and may only be induced under different conditions.

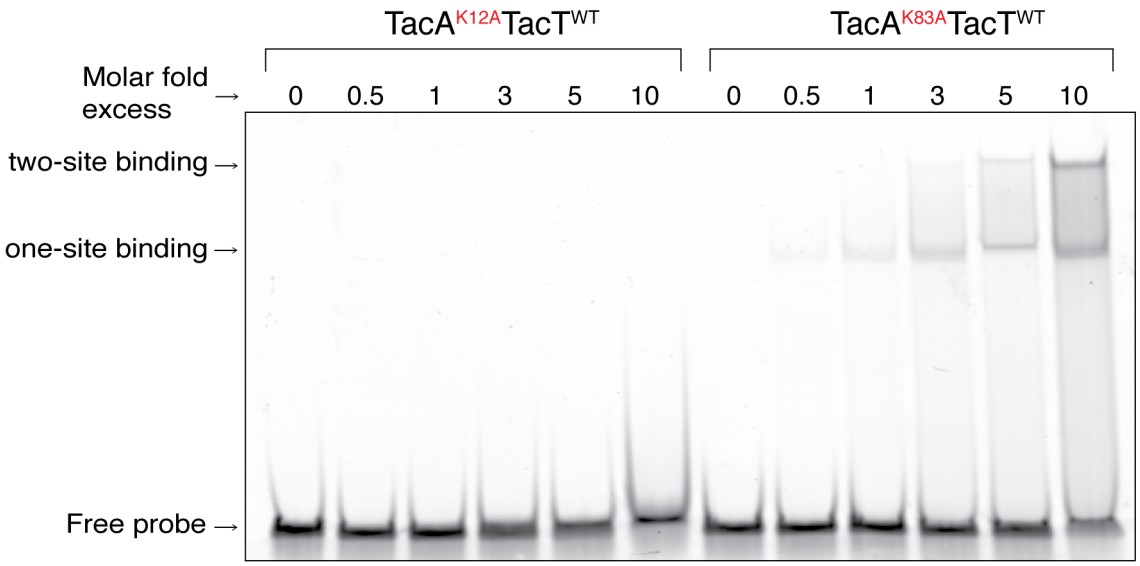


Figure 4.12. Residue K12 of TacA is needed to bind DNA. The ability of lysine variants $TacA^{K12A}T^{WT}$ and $TacA^{83}T^{WT}$ to bind to DNA was analyzed by electrophoretic mobility shift assay. Probe 3 (1.01 pmol) was incubated with increasing molar excess of either $TacA^{K12A}T^{WT}$ or $TacA^{K83A}T^{WT}$ (0.5, 1.01, 3.03, 5.05, 10.1 pmol protein).

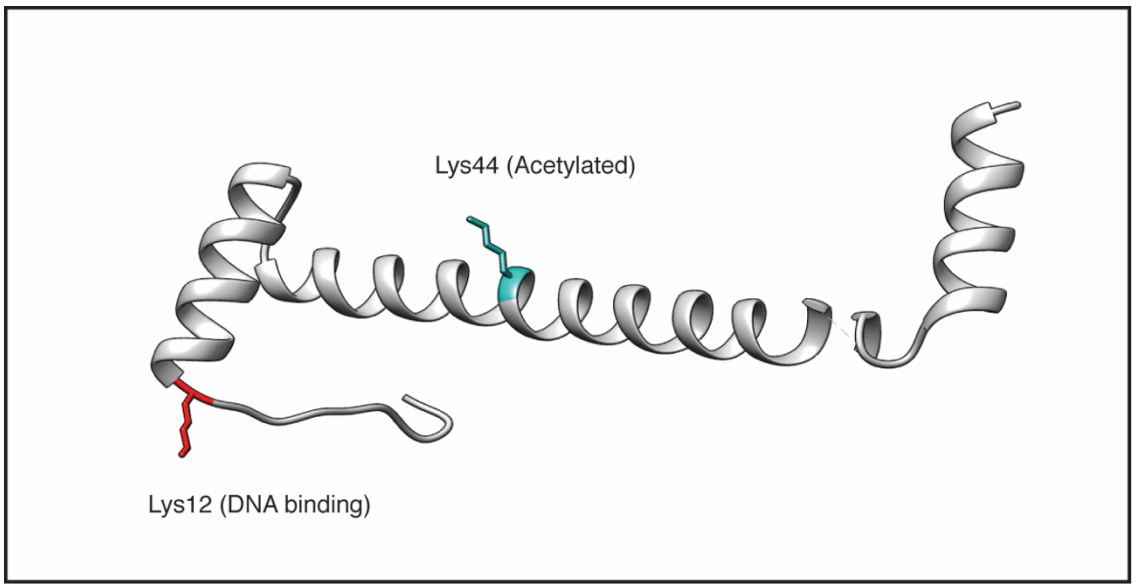


Figure 4.13. PHYRE model representation of antitoxin. Protein sequence of TacA was uploaded to PHYRE protein folding recognition server to analyze predicted structure of the antitoxin. Structure was colored and analyzed using Chimera software. Residue K12, which was required for DNA binding is highlighted in red. Residue K44, which was acetylated by the toxin is highlighted in cyan.

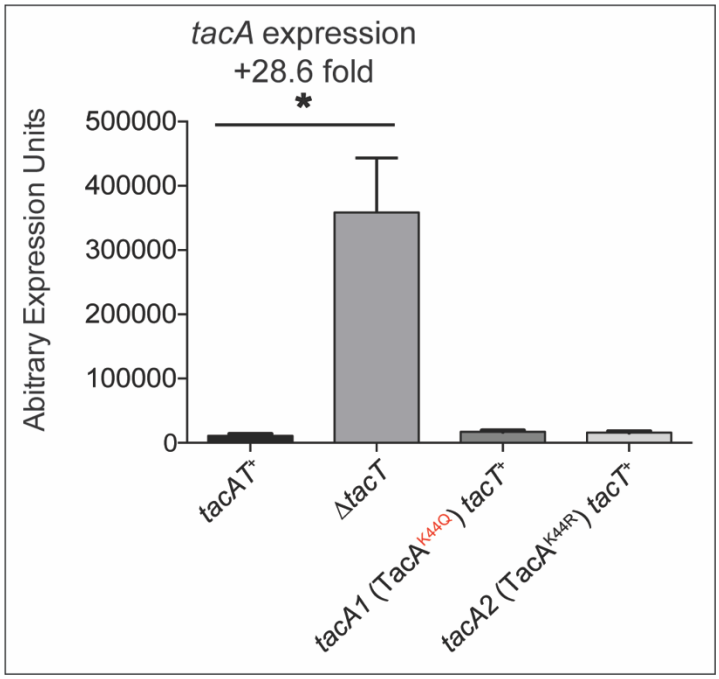


Figure 4.14. Deletion of *tacT* relieves repression of *tacA*. RT-qPCR was used to evaluate the expression of the *tacA* gene in the backgrounds of *tacAT*$^{+}$, Δ*tacT, tacA1 tacT*$^{+}$ (TacA^{K44Q}), and *tacA2 tacT*$^{+}$ (TacA^{K44R}) under the condition of NCE minimal medium supplemented with glycerol (22mM). The gene expression of *tacA* in the toxin deletion strain Δ*tacT* was 28.6-fold higher as compared to the *tacAT*$^{+}$ strain. Interestingly, the two lysine variant strains yielded no significant difference in expression of *tacA* as compared to the *tacAT*$^{+}$ strain. Error bars represent SEM. Asterisk indicates the Welch's *t* test gave a p value of 0.01. Experiment was conducted in technical triplicates of biological triplicates and was repeated three times.

DISCUSSION

Here we report insights into the functionality of one of the three type II toxin-antitoxin systems encoded by the genome of *Salmonella enterica*, a human pathogen. The *S. enterica* TacAT system (encoded by *tacA* and *tacT*, formerly STM3651 and STM3652, respectively) is the focus of the studies reported herein. The TacAT system is the first example of a toxin-antitoxin system whose function is not regulated by dynamic association/dissociation of its components and the first example of a system where the toxin is a GNAT that recognizes a protein and non-protein substrate. In the case of TacAT, TacT acetylates the α-amino group of aminoacyl-tRNAs (29), and also acetylates a lysine residue of TacA (this work). Furthermore, we have shown that sirtuin-dependent reversible lysine acetylation (sRLA) deacetylates $TacA^{Ac}$ and may play a role in controlling the activities of TacA and TacT *in vivo*.

A new role for sRLA in *Salmonella* pathogenesis.

We hypothesize that the *S. enterica* TacAT system is different than other type II toxin-antitoxin systems in that the activity of TacT is not up regulated as a result of the cleavage or cellular degradation of the cognate antitoxin (*i.e.,* TacA) (Aakre 2013; Maisonneuve 2013; Germain 2015). Our data support the conclusion that TacAT activities are modulated post-translationally through sRLA (Figs. 4.6, 5.7). The involvement of sRLA in the modulation of the entrance or egress into or out of a persister state suggests a metabolic link between the latter and the carbon (AcCoA) and energy (NAD^+) statuses of the cell (see below).

Although it is clear that acetylation of TacA alters binding of TacA to the two sites present inside the *tacAT* promoter *in vitro* (Fig. 4.11), more work is needed to understand why and how $TacA^{Ac}$ recognition of two sites is favored over recognition of either one of two sites, and whether or not there are additional factors involved in *tacAT* expression.

Other unique features of the *S. enterica* TacAT system.

As stated above, a unique feature of TacT is that it can acetylate non-protein and protein substrates. Also unique is the fact that TacT activity is enhanced as a result of the acetylation of TacA by TacT while the proteins are in complex. At present, there is no indication that TacA and TacT ever dissociate from each other. In fact, the aminoacyl-tRNA acetylating activity of TacT is enhanced when TacT is in a complex with $TacA^{K44Q}$ (an acetylation mimic) (Fig. 4.6). In contrast, TacT activity decreases in complexes containing TacA or $TacA^{K44R}$ (a deacetylated mimic).

***In vivo* evidence supports the idea that TacA acetylation enhances TacT activity.**

Results from *in vivo* experiments are consistent with *in vitro* results. For example, the phenotype associated with the synthesis of $TacA^{K44Q}TacT^{WT}$ occurred even at low levels of induction (25 μM arabinose), whereas the phenotype generated by high level of expression of TacT alone was erratic, and when it was observed, it occurred only at high levels of induction (150-1000 μM arabinose). These differences could be due to increased turnover rate or to poor solubility of TacT in the absence of TacA. Therefore, when TacT is in complex with $TacA^{K44Q}$, TacT is not only stabilized, but its activity is increased leading to growth arrest (Fig. 4.7). This result suggests that $TacA^{Ac}T^{WT}$ blocks protein synthesis very efficiently. We posit that interaction of *Salmonella* with the macrophage leads to increased AcCoA levels as a result of the inhibition of AcCoA-consuming processes. Such an increase in AcCoA could trigger TacA acetylation by TacT, increased stability of the complex and increased TacT activity, ultimately resulting in the arrest of protein synthesis. These ideas await further experimentation.

Additional perspective.

One central, unanswered question is how *S. enterica* monitors and modulates the levels of $TacA^{WT}$:$TacT^{WT}$ and $TacA^{Ac}$:$TacT^{WT}$ complexes. Our data suggest that the NAD^+-dependent

CobB protein deacetylase prevents the accumulation of $TacA^{Ac}$:$TacT^{WT}$ complex that would drive the cell into a prolonged persister state. Given that CobB can deactylate $TacA^{Ac}$, and that NAD^+ is needed for CobB activity, a plausible answer to the above question may lie on the intracellular NAD^+ level. A decrease in the AcCoA level could be due to its consumption by anabolic processes, which would require a robust energy charge and would correlate with increased NAD^+ levels in the cell. NAD^+ would then be used by CobB to reduce TacT acitivity as a result of $TacA^{Ac}$ deacetylation.

IMPORTANCE.

The TacAT toxin-antitoxin system of *S. enterica* appears to have evolved to use sirtuin-dependent reversible lysine acetylation (sRLA) as a mechanism to rapidly enter and leave a persister state. There are several methods of other type II TA systems that use transcriptional regulation to either replenish pools of antitoxin to neutralize toxin activity as with HicAB (33) or more commonly use the toxin as a co-repressor until a stoichiometric imbalance occurs to de-repress the operon (known as conditional cooperativity) (34, 35). With the TacAT system, the use of sRLA could quickly and efficiently modulate toxin and antitoxin activities without the need for threshold de-repression of transcription. By maintaining a stable TacAT complex the cell may enter or egress the persister state as a function of AcCoA, NAD^+, and probably other as-yet unknown signals. Additional evidence for the involvement of sirtuins in the control of TacAT activities inside the macrophage is needed to advance our understanding of the persister state and how *S. enterica* gets in and out of it. Regardless, the new knowledge reported here suggests a new use of sirtuin activators to help reduce the probability of *S. enterica* maintaining a persister state inside a host.

MATERIALS AND METHODS

Bacterial strains.

All strains constructed were derivatives of *Salmonella enterica* subsp. *Enterica* sv Typhimurium LT2 (hereafter *S. enterica*) using the Wanner in-frame gene deletion method (36). Bacteria were grown shaking at 37 ºC and medium used for growth are described under *culture media and chemicals*.

Culture media and chemicals.

All of the bacterial strains and plasmids used are listed in Table 4.1. Strains were grown in lysogenic broth (LB, Difco), nutrient broth (NB, Difco), or no-carbon essential (NCE) minimal medium (37). Growth studies where glycerol was the sole carbon and energy source, were performed in NCE medium supplemented with $MgSO_4$ (1 mM), Wolfe's trace minerals (1x) (38), and glycerol (22 mM). When used, antibiotics were added at the following concentrations: Ampicillin, 100 μg/ml; chloramphenicol, 20 μg/ml; and kanamycin, 50 μg/ml. Cultures used as inoculum were grown overnight at 37 ºC in NB, and small samples (1%, v/v) were used to inoculate 198 μL of fresh medium placed in each well of a 96-well microtiter plate. L-(+)-arabinose was used as inducer wherever indicated. Microtiter plates were incubated at 37 ºC inside the temperature-controlled chamber of a Powerwave microtiter plate reader (Bio-Tek Instruments), and plates were continuously shaken using the medium setting of the instrument. Cell density was monitored at 630 nm, and data were analyzed using Prism 6 software package (GraphPad).

Table 4.1. Strains and plasmids used in this study.		
Strain	**Relevant genotype**	**Reference/source[a]**
JE10079	*ara-9*	Laboratory strain
Derivatives of JE10079		
JE21469	*tacAT*$^{+}$ / pCV1	This work
JE21486	*tacAT*$^{+}$ / pTacT-4	This work
JE21493	*tacA::cat*$^{+}$ / pCV1	This work
JE21494	*tacA::cat*$^{+}$ / pTacT-4	This work
JE21489	*tacT::cat*$^{+}$/ pCV1	This work
JE21490	*tacT::cat*$^{+}$/ pTacT-4	This work
JE21500	*tacAtacT::cat*$^{+}$/ pCV1	This work
JE21501	*tacAtacT::cat*$^{+}$/ pTacT-4	This work
JE21503	*tacAtacT::cat*$^{+}$/ pTacAT-1	This work
JE23697	*tacAtacT::cat*$^{+}$/ pTacA^{K44R}T	This work
JE23698	*tacAtacT::cat*$^{+}$/ pTacA^{K44Q}T	This work
JE23754	*tacA1 tacT*$^{+}$	This work
JE23755	*tacA2 tacT*$^{+}$	This work
JE23438	Δ*tacT*	
JE10813	Δ*ara-9* / pKD46	
***E. coli* strains**		
E. coli C41 (λDE3)	*pka12*::*kan*$^{+}$ *ompT hsdS* (r_Bm_B) *gal λ (DE3)*	Laboratory collection
E. coli DH5α	Φ80d*lacZ* ΔM15 *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 (r_k^-, m_k^+) *supE*44 *relA*1 *deoR* Δ(*lacZYA-argF)* U169 *phoA*	NEB

Strain construction.

All strains constructed were *ara-9* derivatives of *S. enterica* using the Wanner in-frame gene deletion method (36). Primers used in this study were synthesized by Integrated DNA Technologies (IDT, Coralville, IA), and are listed in Table 4.2. *tacA*::*cat*$^{+}$ and Δ*tacA* strains were engineered as follows. Using Pfu Ultra II Fusion DNA polymerase (Stratagene), flanking regions of the plasmid pKD3 (36) were amplified with primers designed with 36-39bp of overlapping region at the beginning of either the *tacA* or *tacT* genes, and with 50 bp overlapping region at the end of the *tacA* or *tacT* genes. PCR fragments were analyzed by agarose gel electrophoresis on 1% (w/v) agarose gels post-stained with 0.5 μg/ml ethidium bromide for 15 min. PCR fragments were PCR cleaned up using Wizard SV Gel and PCR Clean-Up System (Promega), and 5-10 μl of product were electroporated into *S. enterica* strain JE10813 (Δ*ara-9*) harboring plasmid pKD46 (36). After electroporation, cells were grown up to an optical density (630 nm) of ~0.6 at 30ºC, followed by three washes with glycerol (10%, v/v). Electroporation was performed in 0.2 cm electroporation cuvettes (MidSci) in a BioRad MicroPulser Electroporator on the Ec2 setting. Cells were then recovered in 0.5 ml of LB for 1 h at 37 ºC, plated on LB + agar + antibiotic, and incubated overnight at 37 ºC. Drug resistant transformants were streaked repeatedly on antibiotic plates at 42ºC to cure the strains of plasmid pKD46. Strains were then reconstructed by P22-mediated transduction of the drug marker into strain JE10079. Strains containing chloramphenicol insertions were transformed with the plasmid pCP20 to resolve out the chloramphenicol insertion and make a scarred deletion.

Plasmid construction for complementation and overexpression.

All plasmids used in this work are listed in Table 4.3. Primers used in this study were synthesized by Integrated DNA Technologies (IDT, Coralville, IA), and are listed in Table 4.2.

Table 4.2. Primers and DNA probes used in this study.	
Primer Name	**Primer Sequence 5' → 3'**
Strain construction	
5' *tacT* Wanner primer	CACCAGAACCTTTGTCCGCTTTTCATCAGGTAGCTGGTG TAGGCTGGAGCTGCTTC
3' *tacT* Wanner primer	CGCTTATAGCGATTTGAACATAACAGTCTTTGCATTCTA TTGAGGGAGCCCATATGAATATCCTCCTTAG
5' *tacAT* Wanner primer	GCTATACATGGTGGTTGTGCTATTCTTGTAAAGCAAATG GTGTAGGCTGGAGCTGCTTC
3' *tacAT* Wanner primer	CGCTTATAGCGATTTGAACATAACAGTCTTTGCATTCTA TTGAGGGAGCCCATATGAATATCCTCCTTAG
5' *tacA* Wanner primer	GCTATACATGGTGGTTGTGCTATTCTTGTAAAGCAAATG GTGTAGGCTGGAGCTGCTTC
3' *tacA* Wanner primer	GGTTTCCTTGCCAGCAGTTTTTCGATAACGGGATCATCT GCGACCGGTGCCATATGAATATCCTCCTTAG
5' SOE PCR fragment 1	AATTAACAAGGTCTGTACAGGGATGCTATCAG
3' SOE PCR fragment 1	CTCCAGCCCTCCAGCCTACACCTATTGAGGGAGCCTAA GGAACAATGT
5' SOE PCR fragment 2	CTCCCTCAATAGGTGTAGGCTGGAGCTGCT
3' SOE PCR fragment 2	CCGGCTGTCAGGGCATATGAATATCCTCCTTAG
5' SOE PCR fragment 3	GATATTCATATGCCCTGACAGCCGGAAAGG
3' SOE PCR fragment 3	CGAAACGCGAAAAACCCACG
Site-Directed Mutagenesis	
5' TacA^{K12A}	GCCCGCTGCGATTCCGCAGCTCTAAGGTTAAGTTGAAC
3' TacA^{K12A}	GTTCAACTTAACCTTAGAGCTGCGGAATCGCAGCGGGC
5' TacA^{K44A}	TGCCAGGCTGCCGAGGCAGTGATCCTTGACCG
3' TacA^{K44A}	CGGTCAAGGATCACTGCCTCGGCAGCCTGGCA
5' TacA^{K83A}	CTGCTGGCAAGGGCGCCTCAGTGGGACGTG
3' TacA^{K83A}	CACGTCCCACTGAGGCGCCCTTGCCAGCAG
5' TacA^{K44R}	GTCAAGGATCACTCTCTCGGCAGCCTGGCAG
3' TacA^{K44R}	CTGCCAGGCTGCCGAGAGAGTGATCCTTGAC
5' TacA^{K44Q}	CGGTCAAGGATCACCTGCTCGGCAGCCTGGC
3' TacA^{K44Q}	GCCAGGCTGCCGAGCAGGTGATCCTTGACCG
Cloning [b]	
5' *tacT* cloned into pCV1	NNGCTCTTCNTTCGTGGGACGTGTAACAGCACC
3' *tacT* cloned into pCV1	NNGCTCTTCNTTACTATTGAGGGAGCCTAAGGA
5' *tacA* cloned into pCV1	NNGCTCTTCNTTCATGCTATACAAGGGGTGTCT
3' *tacA* cloned into pCV1	NNGCTCTTCNTTATTACACGTCCCACTGAGGTT
5' *tacT* cloned into MCS1 pACYCDuet BamHI	NNNNNNGGATCCGTGGGACGTGTAACA
3' *tacT* cloned into MCS1 pACYCDuet EcoRI	NNNNNNGAATTCCTATTGAGGGAGCCTAAGGAACAAT GTTC
5' *tacA* cloned into MCS2 pACYCDuet EcoRV	NNNNNNGATATCATGCTATACAAGGGGTGTCT
3' *tacA* cloned into MCS2 pACYCDuet XhoI	NNNNNNCTCGAGTTACACGTCCCACTGAGGTT
5' Upstream *tacA*	NNGCTCTTCNTTCAATTAACAAGGTCTGTACAGGGATG CTATCAG

Table 4.2 cont. Sequence verification primers	
5’ T7 Foward	TAATACGACTCACTATAGGG
3’ T7 Reverse	GCTAGTTATTGCTCAGCGG
5’ pBAD sequencing primer	CTGTTTCTCCATACCCGTT
3’ pBAD sequencing primer	GGCTGAAAATCTTCTCT
5’ *tacT* sequencing primer	TCAACTTAACCTTAGAGCTAAGGA
3’ *tacT* sequencing primer	GCTCACTTTGTACTGACCCC
5’ *tacA* sequencing primer	CACGGCTTTCATATAGTGGAGCG
3’ *tacA* sequencing primer	GCAAGACGGGCAAGTATAATGACAGGG
EMSA primers	
5’ Probe 1, Probe 2 EMSA primer	CCTTGTTGAACTCAGTAAACCG
3’ Probe 1, Probe 3, Probe 5 EMSA primer	ACCCCTTGTATAGCATTTGC
3’ Probe 2 EMSA primer	TGAATACTGATGTTCGCTT
5’ Probe 3, Probe 4 EMSA primer	ACGACGACTTGATGTATAC
3’ Probe 4 EMSA primer	ACCATGTATAGCAAACGG
5’ Probe 5 EMSA primer	TGGTTGTGCTATTCTTGTA
RT-qPCR primers	
5’ *tacA* RTqPCR	CCTTGACCGCCGTGTATTT
3’ *tacA* RTqPCR	GGTTTCCTTGCCAGCAGTT
5’ *rpoB* RTqPCR	AGTCGACCTGAGCACTTCA
3’ *rpoB* RTqPCR	CAAACACTGGTGTGGCAATC
5’ *gyrB* RTqPCR	CCGTTGGATCACGAGTTTG
3’ *gyrB* RTqPCR	AACGCGTCCTCTTCAATCAG

Table 4.3. Plasmids used in this study[a]			
Plasmid	**Genotype**	**Description**	**Source**
pCV1	*araC*$^+$ *bla*$^+$	pBAD24 with BspQI MCS	(39)
pCV3	*araC*$^+$ *cat*$^+$	pBAD33-SD1 with BspQI MCS	(39)
pACYCDuet-1	*cat*$^+$	Vector for coexpression of two genes	Novagen
pTacT-4	*tacT*$^+$ *bla*$^+$	*tacT*$^+$ cloned into pCV1	This study
pTacAT-1	*tacAT*$^+$ *bla*$^+$	*tacAT*$^+$ cloned into pCV1	This study
pTacAT-2	*tacAT*$^+$ *cat*$^+$	pACYCDuet plasmid with *tacT* cloned into MCS1, *tacA* cloned into MCS2	This study
pTacAT-12	*tacAT*$^+$ *cat*$^+$	pACYCDuet plasmid with *tacT* cloned into MCS1, *tacA* coding for TacA^{K12A} cloned into MCS2	This study
pTacAT-13	*tacAT*$^+$ *cat*$^+$	pACYCDuet plasmid with *tacT* cloned into MCS1, *tacA* coding for TacA^{K44A} cloned into MCS2	This study
pTacAT-14	*tacAT*$^+$ *cat*$^+$	pACYCDuet plasmid with *tacT* cloned into MCS1, *tacA* coding for TacA^{K83A} cloned into MCS2	This study
pTacAT-15	*tacAT*$^+$ *cat*$^+$	pACYCDuet plasmid with *tacT* cloned into MCS1, *tacA* coding for TacA^{K44R} cloned into MCS2	This study
pTacAT-16	*tacAT*$^+$ *cat*$^+$	pACYCDuet plasmid with *tacT* cloned into MCS1, *tacA* coding for TacA^{K44Q} cloned into MCS2	This study
pTacAT-9	*tacAT*$^+$ *bla*$^+$	*tacAT*$^+$ cloned into pCV1, *tacA* coding for TacA^{K44R}	This study
pTacAT-10	*tacAT*$^+$ *bla*$^+$	*tacAT*$^+$ cloned into pCV1, *tacA* coding for TacA^{K44Q}	This study
pTacAT-17	P_{tacAT} *bla*$^+$	Upstream region (-792 bp) of *tacAT*$^+$ with *tacAT*$^+$ cloned into pCV1	This study
pTacAT-22	*tacAT*$^+$ *bla*$^+$	Upstream region (-792 bp) of *tacAT*$^+$ with *tacAT* (coding for TacA^{K44R}) cloned into pCV1	This study
pTacAT-23	P_{tacAT} *bla*$^+$	Upstream region (-792 bp) of *tacAT*$^+$ with *tacAT* (coding for TacA^{K44Q}) cloned into pCV1	This study
pCobB71	*cobB*$^+$ *bla*$^+$	*cobB* ORF in pTEV6	(40)
pKD46	*bla*$^+$	Express the Red system to avoid unwanted recombination	(36).
pKD3	*cat*$^+$	Template plasmid carrying *cat*$^+$ gene	(36).

We used the high-efficiency cloning method described elsewhere (39) to clone *tacA*, *tacT*, *tacAT* and *cobB* genes into pCV1 and pCV3 vectors. Plasmid pCV1 is a modified plasmid of pBAD24 (41) with BspQI sites added, and confers ampicillin resistance, and expression of genes cloned into it can be induced with L(+)-arabinose. Plasmid pCV3 is a modified plasmid of pBAD33 (41) with added BspQI sites and confers chloramphenicol resistance and arabinose induction. Genes *tacA, tacT*, *tacAT*, and *cobB* were amplified from the *S. enterica* chromosome using PfuUltra II Fusion DNA polymerase (Stratagene). PCR fragments were analyzed by agarose gel electrophoresis on 1% (w/v) agarose gels stained with ethidium bromide. PCR fragments were PCR cleaned up using Wizard SV Gel and PCR Clean-Up System (Promega) and digested with the restriction enzyme BspQI (NEB) at 50 ºC for 1 h; products were ligated with T4 DNA ligase (Fisher).

The overexpression vector pACYCDuet (EMD Millipore Biosciences) possesses two multiple cloning sites (MCS) for which two separate genes may be cloned into and overexpressed simultaneously from the same vector. This vector was used for the overexpression of TacT and TacA from the same plasmid. pACYCDuet was first digested with FastDigest (ThermoFisher Scientific) BamHI and EcoRI for 1 h at 37 ºC. *tacA* and *tacT* were amplified from the *S. enterica* chromosome using PfuUltra II Fusion DNA polymerase (Stratagene). PCR fragments were analyzed by agarose gel electrophoresis on 1% (w/v) agarose gels stained with ethidium bromide. PCR fragments were PCR cleaned up using Wizard SV Gel and PCR Clean-Up System (Promega). The *tacT* PCR product was digested with FastDigest (ThermoFisher Scientific) BamHI and EcoRI for 1 h at 37 ºC. The digested PCR product and pACYCduet plasmid were PCR cleaned up with the same clean-up kit stated above, and the digested *tacT* product was ligated into MCS1 of digested pACYCDuet with T4 DNA ligase (Fisher) at room temperature for 30 min. Once

transformants were verified, the *tacA* PCR product was digested with EcoRI and XhoI for 1 h at 37 ºC, PCR cleaned, and ligated as described above into the second MCS of pACYCDuet that had the *tacT* gene cloned into MSC1.

Plasmids were isolated using Wizard Plus SV Miniprep Kit (Promega). To confirm the correct sequences were cloned without mutations, DNA sequencing reactions were analyzed at the Georgia Genomics Facility, UGA. Site-directed mutagenesis (Stratagene) was performed on pTacAT-2 or pTacAT-1 to change mentioned residues K12, K44, K83 to A, Q, or R. Polymerase chain reaction was performed using PfuUltra II DNA polymerase using primers listed in (Table 4.2). Modifications included an anneal time of 60 s, an extension temperature of 68 °C, and an extension time of 2.5 min kb^{-1}. DNA changes were confirmed by sequencing.

Protein purification.

Plasmids encoding H_6-$TacT^{WT}$ and $TacA^{WT}$ (pTacAT-2), H_6-$TacT^{WT}$ and $TacA^{K12A}$ (pTacAT-12), H_6-$TacT^{WT}$ and $TacA^{K44A}$ (pTacAT-13), H_6-$TacT^{WT}$ and $TacA^{K44Q}$ (pTacAT-16), H_6-$TacT^{WT}$ and $TacA^{K44R}$ (pTacAT-15), and His_6-$TacT^{WT}$ and $TacA^{K83A}$ (pTacAT-14) were electroporated into *Escherichia coli* strain C41 (λDE3) (42) Δ*pat* (strain JE9314). Cultures of cells containing plasmids were grown to stationary phase (OD_{650} ~1.3) and sub-cultured (1:100 v/v) into 6 L of LB + chloramphenicol. Cultures were grown shaking at 25 °C to an OD_{650} of 0.5, after which ectopic gene expression was induced with IPTG (0.5 mM). Cultures were grown overnight at 25°C, cells were harvested by centrifugation at 6,000 x *g* for 15 min at 4°C, and cell pellets were stored at -80°C until used.

Cell pellets were re-suspended in 50 mL of buffer A containing 4-(2-hydroxymethyl)-1-piperazineethanesulfonic acid (50 mM HEPES, pH 7.0 @ 4°C), NaCl (500 mM), imidazole (20 mM), glycerol (20 %, v/v), lysozyme (1 mg/mL), DNase (1 μg/ml), and protease inhibitor

phenylmethane sulfonyl fluoride (PMSF, 1 mM). Cells were sonicated for 60 s using a Qsonica sonicator at 60% duty with 2 s pulses. Lysates were centrifuged using a Beckman Coulter Avanti J-251 centrifuge at 40,000 x g for 30 min equipped with JA-25.50 rotor. Clarified lysates were filtered through a 0.45 μm and applied to a 2-mL HisTrap FF (GE Healthcare Sciences) column using an ÄKTA FPLC system (GE Healthcare Sciences). Column was washed with 10 column volumes of bind buffer, 7 CV of 8% elution buffer (50 mM HEPES, pH 7.0 @ 4 °C, 500 mM NaCl, 500 mM imidazole, and 20% glycerol v/v), and a 20-column volumes gradient to 100% elution buffer. When separation of TacT and TacA complexes was necessary, proteins were denatured on the column as described (31). Fractions were run on an SDS-PAGE gel and fractions containing desired protein were combined and dialyzed for 3 h each in storage buffer 1 (50 mM HEPES, pH 7.0 @ 4 °C, 400 mM NaCl, and 20% glycerol v/v), storage buffer 2 (50 mM HEPES, pH 7.0 @ 4 °C, 200 mM NaCl, and 20% glycerol v/v), and storage buffer 3 (50 mM HEPES, pH 7.0 @ 4 °C, 150 mM NaCl, and 20% glycerol v/v). Proteins were flash frozen in liquid N_2 and stored at -80 °C. Proteins were quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific) using the molecular mass (28.67 kDa) and extinction coefficient (18,700 M^{-1} cm^{-1}; ExPASy ProtParam) of the complex (assuming a 1:1 ratio of TacT to TacA, as determined via size exclusion chromatography, see below). Percent purity was calculated using ImageQuant™ v5.2 software. CobB protein was purified as described elsewhere (40). TacT was isolated from TacAT complex using a purification protocol reported elsewhere (31). Briefly, $TacAT^{WT}$ complex was purified as above and was separated into its components by denaturation with guanidine-HCl (5M) followed by overnight dialysis. Denatured proteins were resolved by Ni-affinity chromatography, and refolded by dialysis of the denaturant. As refolding occurred, the bulk of TacT became

insoluble while TacA remained stable in solution. This procedure yielded 6-fold larger amounts of TacA than TacT (*e.g.,* 0.5 mg of TacA *vs* 0.08 mg TacT per liter of culture).

Size exclusion chromatography.

A Superose 12 10/300 GL gel filtration column (GE Healthcare Life Sciences) was equilibrated as per manufacturer's protocol using water and elution buffer (50 mM HEPES, pH 7.0 @ 4°C, 150 mM NaCl, 20% glycerol v/v). Samples were applied to column using a 100-μL superloop. Gel filtration standards (Bio-Rad) were applied to column first and a standard curve was calculated using the MW_{log} of each standard against retention time of each protein. Purified TacAT complexes were eluted from the column in the same manner and retention times were recorded and molecular weights were determined using equation calculated from the standard curve.

***In vitro* acetylation assays.**

Homogeneous TacAT complex (3 μM) was incubated with or without [1-^{14}C]-acetyl-CoA (20 μM) in HEPES (50 mM, pH 7.5) and *tris*(2-carboxyethyl)phosphine (TCEP, 1 mM), for 1 h at 37 ºC in a total volume of 25 μl. Reactions were quenched by the addition of SDS-loading buffer (60% (v/v) glycerol, Tris-HCl pH 6.8 (0.3 M), EDTA (12 mM), 12% SDS, 2-mercaptoethanol (0.87 mM), bromophenol blue (0.05%, w/v) and reaction mixtures were resolved by SDS-PAGE on a 15% (w/v) polyacrylamide gel with Tris-HCl buffer pH 8.8 (resolving gel)/Tris-HCl pH 6.8 (stacking gel). Samples were run at 200V for 45 min. Transfer of the radiolabel onto TacA was visualized using a Typhoon Trio+ variable mode imager (GE Healthcare).

***In vitro* deacetylation assays.**

To determine whether or not TacA^{Ac} was a substrate for CobB sirtuin, 200 μL reaction mixtures containing radiolabeled TacA^{Ac} synthesized as described above was treated NAD^+-

dependent CobB sirtuin deacetylase [1-^{14}C]-acetyl-CoA. Briefly, excess [1-^{14}C]-acetyl-CoA was removed by buffer exchange using Amicon®Ultra – 0.5 ml centrifugal filters (Ultracell®, 10K MWCO) and HEPES buffer (50 mM, pH 7.5). Reaction mixtures were concentrated to 100 μL and served as a 2X stock of TacA^{Ac}. After removal of [1-^{14}C]-acetyl-CoA, TacA^{Ac} was added to reaction mixtures (1X) that contained; i) NAD^+ (1 mM) + CobB (75 pmol; 3 μM final concentration), ii) NAD^+ (1 mM) + CobB (3 μM), or iii) NAD^+ (1 mM) + CobB (3 μM) + nicotinamide (5 mM). Samples were incubated at 37 ºC for 1 h, then resolved by SDS-PAGE. Deacetylation of TacA^{Ac} was monitored with a phosphor imager as described above.

***In vitro* DHFR protein synthesis assay.**

The PureExpress® *In Vitro* Protein Synthesis Kit (New England BioLabs) reactions were set up per manufactures protocol in RNase free tubes with the following modifications. Reactions were adjusted to 15 μL total volume (5 μL of Solution A, 3.75 μL Solution B, 100 ng DHFR DNA) and when noted, supplemented with acetyl-CoA (2 mM), TacAT complex (2 μM), or TacT (2 μM). All samples contained 0.5 μL of SUPERase In™ RNase inhibitor (ThermoFisher Scientific). Reactions were brought up to equal volumes with the addition of RNase free water, per manufactures protocol. Reaction mixtures were incubated in a 37°C sand bath for 2 h, after which tubes were placed on ice and 2.5 μL of each reaction mixture was added to 12 μL of 1x loading dye (60% (v/v) glycerol, Tris-HCl pH 6.8 (0.3 M), EDTA (12 mM), 12% SDS, 2-mercaptoethanol (0.87 mM), bromophenol blue (0.05%, w/v)) and heated at 100 °C for 10 min. 5 μL of each denatured sample was resolved by SDS-PAGE on a 15-comb, 15% (w/v) polyacrylamide gel with Tris-HCl buffer pH 8.8 (resolving gel)/Tris-HCl pH 6.8 (stacking gel). Gels were run for 45 min at 220 V and visualized via Coomassie Blue staining and acetic acid destaining. Gels were imaged and analyzed for DHFR protein production using ImageQuant™ v5.2 software. The intensity of

DHFR from each lane on the polyacrylamide gel was normalized to the asterisk-indicated bands in Figure 4.6 using the ImageQuant™ v5.2 software. These values were used to calculate the percent decrease in DHFR per reaction. The mean percentages were plotted using Prism6 software to obtain the standard deviation of each percent decrease as shown in Figure 4.6.

DNA-binding assays.

Electrophoretic mobility shift DNA-binding assays were performed using DNA probes with 6-carboxyfluorescein (6-FAM) covalently attached to the 5' end of the probe. Primers used in these experiments were manufactured by Integrated DNA Technologies (IDT, Coralville, IA). Probes were generated from PCR amplification of strain JE10079 (*tacAT*$^{+}$) chromosomal DNA. PCR products were sized on a 1% (w/v) agarose gel and were purified using a Wizard SV gel and the PCR cleanup system (Promega). Binding reaction mixtures contained 6-FAM dsDNA probe (50 ng), HEPES-NaOH buffer (50mM, pH 7) containing KCl (50 mM), $MgCl_2$ (10 mM), disodium ethylenetetraacetic acid (Na_2EDTA, 0.5 mM), glycerol (10%, v/v), 25 μg / μl poly dI-Dc (deoxyinosinic-deoxycytidylic acid [(Poly(dI-dC) Sigma-Aldrich]), and when added, TacT-TacA complex protein in molar excess of probe as indicated in figure legends. Reaction mixtures (25 μl) were incubated at 25 ºC for 45 min. Glycerol (27 μmol; 5 μl of a 50% v/v solution) was added reaction mixtures, which were resolved using a non-denaturing Criterion Tris-HCl buffer (375 mM, pH 8.6) 7.5% (w/v) polyacrylamide gel (BioRad) at 120 V. Gels were imaged using a Typhoon Trio+ variable mode imager (GE Healthcare) wavelength 488nm (Blue) and analyzed with ImageQuant v5.2 software.

Construction of a strain carrying a chromosomal *tacA* allele encoding TacA^{K44Q} or TacA^{K44R}.

A region 792-bp upstream of *tacAT* in frame with the coding region of *tacAT* was cloned into pCV1 (pTacAT-17). Site-directed mutagenesis was performed to mutate *tacA* to encode either

TacA^{K44R} to TacA^{K44Q}, as described above (pTacAT-22, pTacAT-23). This DNA fragment (792 bp upstream *tacAT* in frame with *tacAT* coding for variants) was designated as fragment 1 (Fig. 4.15). Gene splicing by overlap extension quantitative polymerase chain reaction (SOE-qPCR) was utilized to fuse this fragment to a *cat*$^{+}$ gene (fragment 2) and the downstream region of *tacAT* (fragment 3) The *cat*$^{+}$ gene from pKD3 (36) was used as a template for fragment 2. Fragments 1, 2, and 3 were amplified with Pfu Ultra II Fusion DNA polymerase (Stratagene) with primers listed in Table 4.2 and an annealing temperature of 61.8°C. PCR products were sized on a 1% (w/v) agarose gel and were purified using a Wizard SV PCR cleanup system (Promega). Fragments 1 and 2 (50 ng each per 50 μL PCR reaction) were annealed using SOE PCR primers (5' SOE PCR fragment 1 and 3' SOE PCR fragment 2, Table 4.2) with an annealing temperature of 61.8 °C and an extension time of 30 sec/kb PCR products were sized on a 1% (w/v) agarose gel and were purified using a Wizard SV PCR cleanup system (Promega). Fragment 1-2 was annealed to fragment 3 using the same protocol. The linear PCR fragment 1-3 was transformed into a Δ*tacAT* strain harboring the helper plasmid pKD46 using protocol described above under *Strain construction* (36). The linear PCR fragment recombined with the upstream and downstream region of *tacAT,* which inserted the *tacAT* operon coding for K44Q or K44R in place of its absence. Cells were plated on LB agar + chloramphenicol (10 μg/mL) and individual colonies were screened for acquisition of *tacAT* compared to Δ*tacAT*. Chromosomal mutations were confirmed by sequencing.

RNA isolation.

Strains JE10079 (*tacAT*$^{+}$), JE23754 (*tacA1* encoding TacA^{K44Q} *tacT*$^{+}$), JE23755 (*tacA2* encoding TacA^{K44R} *tacT*$^{+}$), JE23438 (Δ*tacT*) were grown overnight in triplicate in nutrient broth (2 ml; NB, Difco) with shaking at 37ºC. After incubation, strains were diluted 1:100 into 5 ml of

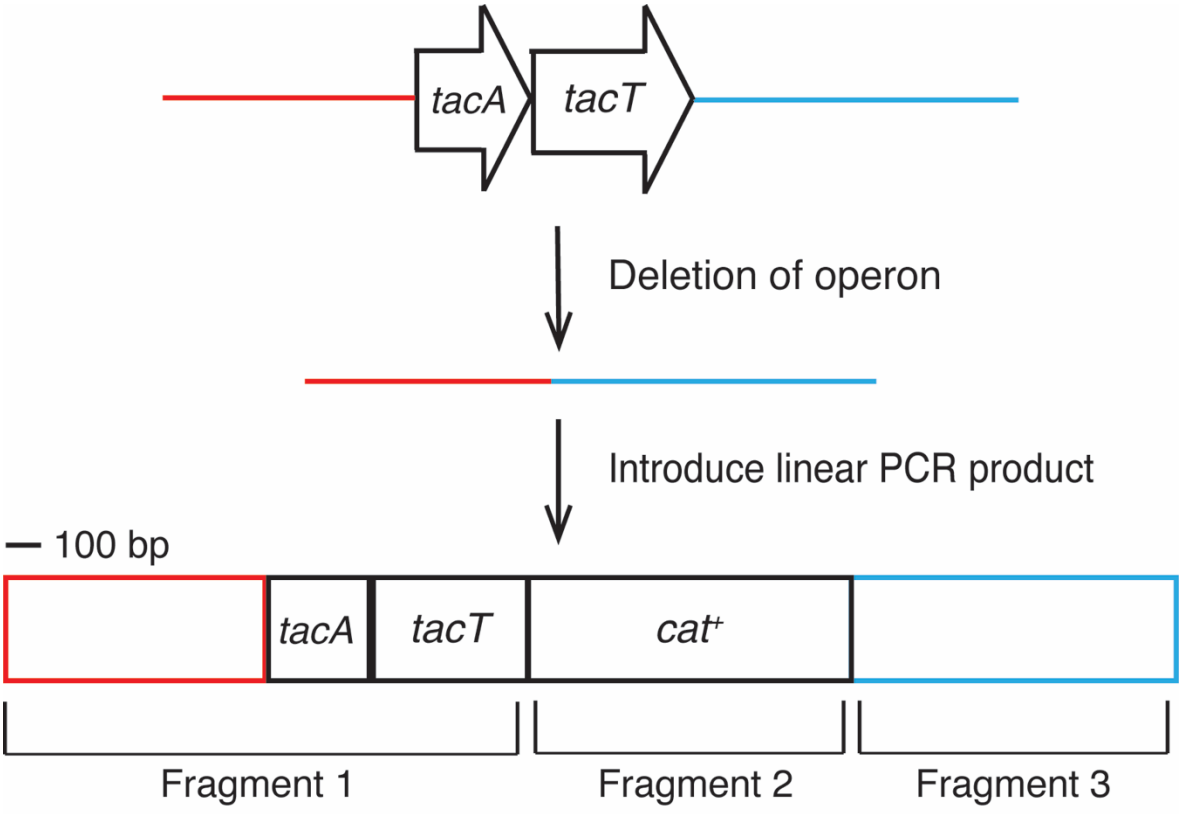


Figure 4.15. Schematic of SOE-qPCR fragment assembly as described in *Materials and methods*. Upstream region in red comprised of 792 bp upstream of *tacAT*. Downstream region ~1 kB of *tacAT*. Fragments were fused together using SOE-qPCR using primers listed in Table S2. DNA sizes to scale.

fresh NCE minimal medium supplemented with $MgSO_4$ (1 mM), Wolfe's trace minerals (1x), and glycerol (22 mM). Cultures were grown shaking at 37 ºC to an optical density (600 nm) of 0.5, then 5 ml of each sample were quickly centrifuged in 1.5 mL Eppendorf tubes at 6000 x g, supernatant was removed, and pellets were flash-frozen in liquid nitrogen and kept on dry ice. RNA was isolated following the RNAsnap™ protocol (43). Pellets were re-suspended in 150 µl of boil solution (ethylenediaminetretraacetic acid (EDTA, 18 mM), SDS (0.025%, w/v) formamide (95%, v/v; RNA grade), 2-mercaptoethanol (1% v/v) in RNase-free water) and were mixed vigorously to break up the cell pellet. Pellets were incubated at 95 ºC for 7 min and centrifuged at 16,000 x g for 5 min at room temperature; 100 µl of supernatant was transferred to a fresh tube. A sodium acetate/ethanol RNA precipitation was then conducted by the addition of 300 µl of RNase-free water, 40 µl of sodium acetate (3 M, pH 5.2; final concentration of 0.3 M), and finally 400 µl of ice-cold absolute ethanol (100%), with mixing briefly before the addition of the next reagent. The mixture was incubated on ice for 15 min, centrifuged at 16,000 x g for 15 min at 4ºC, and ethanol was decanted off. Ethanol (400 µl of cold 70% v/v) was added and pellets were centrifuged at 12,000 x g for 10 min at 4ºC in an Eppendorf 5415D centrifuge. Ethanol was removed and pellets were allowed to dry. RNA pellets were re-suspended in RNase-free water at 4ºC on ice overnight. Subsequent RNase-free DNase I treatment was conducted using the Ambion Turbo DNA-free kit according to manufacturer's instructions (ThermoFisher Scientific). After DNA cleavage, a final sodium acetate/ethanol precipitation was performed as described above. RNA was allowed to re-suspend at 4ºC on ice for 4 h, then were flash frozen in liquid nitrogen and stored at -80ºC until used. A small aliquot of each sample was sent for quality control analysis using the RNA 600 nano kit of the Agilent 2100 bioanalyzer through the Georgia Genomics Facility. Primers

for qPCR were designed using Primer 3 software and were evaluated for specificity and melting curve prior to running the qPCR.

cDNA synthesis and quantitative real time polymerase chain reaction.

Total RNA (620 ng) from each sample was used for the synthesis of cDNA using the iScript^{TM} cDNA synthesis Kit from BioRad Laboratories according to manufacturer's protocol. Each cDNA reaction was then dilute to 7.5 ng/µl and used as template for PCR. For real-time PCR, 20 µl reactions were prepared with 10 µl of 2X FastSYBR Green master mix (Applied Biosystems), 500 nM of each gene-specific primer (1µl of 10 µM primer stock), and 15 ng of cDNA (2 µl of 7.5 ng/µl cDNA). The real-time PCR reaction was performed using a 7500 Fast real-time PCR system (Applied Biosystems). The threshold cycle values of *rpoB* and *gyrB* were checked first to ensure that both genes were optimal for use as reference genes for these strains under the conditions chosen for RT-qPCR. Cycle threshold (C_T) data were normalized to the *rpoB* gene (44). These normalized values (ΔC_T) were transformed using $2(e\text{-}\Delta Ct)/10^{-6}$ (45), and were reported as arbitrary gene expression units (EU), or the gene expression ratio of the mutant strains/the parent strain (JE10079 *tacAT*$^{+}$). Mean EU values were used to calculate the standard error of the mean (SEM) using Prism6 from three biological replicates that were each tested in technical triplicate. Differences in EU between mutant strains and JE10079, and between JE23754 (*tacA1* encoding TacA^{K44Q} *tacT*$^{+}$), JE23755 (*tacA2* encoding TacA^{K44R} *tacT*$^{+}$), were compared using Welch's t-test with the Graphpad Prism6 software as shown in Figure 4.14.

AUTHOR CONTRIBUTIONS.

CMV, ARP designed and performed the reported experiments, analyzed the data and wrote the paper; they contributed equally to the performance of the work. JCES conceived the project, helped design experiments, analyzed data and wrote the paper.

FUNDING INFORMATION.

This work was supported by USPHS grant from the National Institutes of Health R01 GM062203 to J.C.E.-S. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

ACKNOWLEDGMENTS.

The authors have no conflict of interest to declare. We gratefully acknowledge valuable discussions with Diana Downs, Janet Westpheling, Joseph Groom, and Andrew Borchert.

REFERENCES

1. Hentchel KL, Escalante-Semerena JC. 2015. Acylation of biomolecules in prokaryotes: a widespread strategy for the control of biological function and metabolic stress. Microbiol Mol Biol Rev 79:321-346.
2. Bheda P, Jing H, Wolberger C, Lin H. 2016. The Substrate Specificity of Sirtuins. Annu Rev Biochem 85:405-429.
3. Rocker A, Meinhart A. 2016. Type II toxin: antitoxin systems. More than small selfish entities? Curr Genet 62:287-290.
4. Chan WT, Espinosa M, Yeo CC. 2016. Keeping the wolves at bay: Antitoxins of prokaryotic type II toxin-antitoxin systems. Front Mol Biosci 3:9.
5. Gerdes K, Rasmussen PB, Molin S. 1986. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. Proc Natl Acad Sci U S A 83:3116-3120.

6. Roberts RC, Strom AR, Helinski DR. 1994. The parDE operon of the broad-host-range plasmid RK2 specifies growth inhibition associated with plasmid loss. J Mol Biol 237:35-51.

7. Ogura T, Hiraga S. 1983. Partition mechanism of F plasmid: two plasmid gene-encoded products and a cis-acting region are involved in partition. Cell 32:351-360.

8. Bernard P, Kezdy KE, Van Melderen L, Steyaert J, Wyns L, Pato ML, Higgins PN, Couturier M. 1993. The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. J Mol Biol 234:534-541.

9. Lehnherr H, Maguin E, Jafri S, Yarmolinsky MB. 1993. Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. J Mol Biol 233:414-428.

10. Buts L, Lah J, Dao-Thi MH, Wyns L, Loris R. 2005. Toxin-antitoxin modules as bacterial metabolic stress managers. Trends Biochem Sci 30:672-679.

11. Page R, Peti W. 2016. Toxin-antitoxin systems in bacterial growth arrest and persistence. Nat Chem Biol 12:208-214.

12. Christensen SK, Mikkelsen M, Pedersen K, Gerdes K. 2001. RelE, a global inhibitor of translation, is activated during nutritional stress. Proc Natl Acad Sci U S A 98:14328-14333.

13. Moyed HS, Bertrand KP. 1983. hipA, a newly recognized gene of Escherichia coli K-12 that affects frequency of persistence after inhibition of murein synthesis. J Bacteriol 155:768-775.

14. Germain E, Castro-Roa D, Zenkin N, Gerdes K. 2013. Molecular mechanism of bacterial persistence by HipA. Mol Cell 52:248-254.

15. Christensen SK, Pedersen K, Hansen FG, Gerdes K. 2003. Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. J Mol Biol 332:809-819.

16. Maisonneuve E, Shakespeare LJ, Jorgensen MG, Gerdes K. 2011. Bacterial persistence by RNA endonucleases. Proc Natl Acad Sci U S A 108:13206-13211.

17. Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. 2004. Bacterial persistence as a phenotypic switch. Science 305:1622-1625.

18. Keren I, Shah D, Spoering A, Kaldalu N, Lewis K. 2004. Specialized persister cells and the mechanism of multidrug tolerance in Escherichia coli. J Bacteriol 186:8172-8180.

19. Lewis K. 2007. Persister cells, dormancy and infectious disease. Nat Rev Microbiol 5:48-56.

20. Hayes F. 2003. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. Science 301:1496-1499.

21. Van Melderen L, Bernard P, Couturier M. 1994. Lon-dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. Mol Microbiol 11:1151-1157.

22. Maisonneuve E, Castro-Camargo M, Gerdes K. 2013. (p)ppGpp controls bacterial persistence by stochastic induction of toxin-antitoxin activity. Cell 154:1140-1150.

23. Aakre CD, Phung TN, Huang D, Laub MT. 2013. A bacterial toxin inhibits DNA replication elongation through a direct interaction with the beta sliding clamp. Mol Cell 52:617-628.

24. Germain E, Roghanian M, Gerdes K, Maisonneuve E. 2015. Stochastic induction of persister cells by HipA through (p)ppGpp-mediated activation of mRNA endonucleases. Proc Natl Acad Sci U S A 112:5171-5176.

25. Cataudella I, Trusina A, Sneppen K, Gerdes K, Mitarai N. 2012. Conditional cooperativity in toxin-antitoxin regulation prevents random toxin activation and promotes fast translational recovery. Nucleic Acids Res 40:6424-6434.

26. Gelens L, Hill L, Vandervelde A, Danckaert J, Loris R. 2013. A general model for toxin-antitoxin module dynamics can explain persister cell formation in E. coli. PLoS Comput Biol 9:e1003190.

27. Bertram R, Schuster CF. 2014. Post-transcriptional regulation of gene expression in bacterial pathogens by toxin-antitoxin systems. Front Cell Infect Microbiol 4:6.

28. Hayes F, Kedzierska B. 2014. Regulating toxin-antitoxin expression: controlled detonation of intracellular molecular timebombs. Toxins (Basel) 6:337-358.

29. Helaine S, Cheverton AM, Watson KG, Faure LM, Matthews SA, Holden DW. 2014. Internalization of *Salmonella* by macrophages induces formation of nonreplicating persisters. Science 343:204-208.

30. Lobato-Marquez D, Moreno-Cordoba I, Figueroa V, Diaz-Orejas R, Garcia-del Portillo F. 2015. Distinct type I and type II toxin-antitoxin modules control Salmonella lifestyle inside eukaryotic cells. Sci Rep 5:9374.

31. Cheverton AM, Gollan B, Przydacz M, Wong CT, Mylona A, Hare SA, Helaine S. 2016. A *Salmonella* toxin promotes persister formation through acetylation of tRNA. Mol Cell 63:86-96.

32. Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.

33. Turnbull KJ, Gerdes K. 2017. HicA Toxin of Escherichia coli derepresses hicAB transcription to selectively produce HicB antitoxin. Mol Microbiol doi:10.1111/mmi.13662.

34. Turnbull AL, Surette MG. 2010. Cysteine biosynthesis, oxidative stress and antibiotic resistance in *Salmonella typhimurium*. Res Microbiol 161:643-650.

35. Winther KS, Gerdes K. 2012. Regulation of enteric vapBC transcription: induction by VapC toxin dimer-breaking. Nucleic Acids Res 40:4347-4357.

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

37. Berkowitz D, Hushon JM, Whitfield HJ, Jr., Roth J, Ames BN. 1968. Procedure for identifying nonsense mutations. J Bacteriol 96:215-220.

38. Balch WE, Wolfe RS. 1976. New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. Appl Environ Microbiol 32:781-791.

39. VanDrisse CM, Escalante-Semerena JC. 2016. New high-cloning-efficiency vectors for complementation studies and recombinant protein overproduction in Escherichia coli and Salmonella enterica. Plasmid 86:1-6.

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

41. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177:4121-30.

42. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289-298.

43. Stead MB, Agrawal A, Bowden KE, Nasir R, Mohanty BK, Meagher RB, Kushner SR. 2012. RNAsnap: a rapid, quantitative and inexpensive, method for isolating total RNA from bacteria. Nucleic Acids Res 40:e156.

44. Rocha DJ, Santos CS, Pacheco LG. 2015. Bacterial reference genes for gene expression studies by RT-qPCR: survey and analysis. Antonie Van Leeuwenhoek 108:685-693.

45. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402-408.

CHAPTER 5

IN *STREPTOMYCES LIVIDANS*, ACETYL-COA SYNTHETASE ACTIVITY IS CONTROLLED BY *O*-SERINE AND N^{ε}-LYSINE ACETYLATION [1]

[1]VanDrisse C.M and Escalante-Semerena J.C. 2018. *Mol. Microbiol.* 107:577-594.

ABSTRACT

Protein acetylation is a rapid mechanism for control of protein function. Acetyl-CoA synthetase (AMP-forming, Acs) is the paradigm for the control of metabolic enzymes by lysine acetylation. In many bacteria, type I or II protein acetyltransferases acetylate Acs, however, in actinomycetes type III protein acetyltransferases control the activity of Acs. We measured changes in the activity of the *Streptomyces lividans* Acs (*Sl*Acs) enzyme upon acetylation by PatB using *in vitro* and *in vivo* analyses. In addition to the acetylation of residue K610, residue S608 within the acetylation motif of *Sl*Acs was also acetylated (PKTR**S**G**K**^{610}). S608 acetylation rendered *Sl*Acs inactive and non-acetylatable by PatB. It is unclear whether acetylation of S608 is enzymatic, but it was clear that this modification occurred *in vivo* in *Streptomyces.* In *S. lividans,* an NAD^+-dependent sirtuin deacetylase from *Streptomyces*, SrtA (a homologue of the human SIRT4 protein) was needed to maintain *Sl*Acs function *in vivo.* We have characterized a sirtuin-dependent reversible lysine acetylation system in *Streptomyces lividans* that targets and controls the Acs enzyme of this bacterium. These studies raise questions about acetyltransferase specificity, and describe the first Acs enzyme in any organism whose activity is modulated by *O*-Ser and N^{ε}-Lys acetylation.

INTRODUCTION

Control of protein function by chemical modifications is widespread in biology. Such modifications are diverse, including phosphorylation, hydroxylation, automodification, methylation, glycosylation, sumoylation, ubiquitination, prenylation, ADP-ribosylation, acylation, etc. (1). Protein acylation itself is also diverse since acyl chains can vary in length, can have one

or more carboxylate functional groups, the carbon skeleton may or may not be unsaturated, and the moieties can be used to generate *O*-acylated or *N*-acylated derivatives (2-5).

Here we investigated the function of GCN5-related *N*-acetyltransferases (GNATs) (PF00583) from actinomycetes. Some GNATs control the activity of acetyl-CoA (AMP-forming) synthetases (Acs), which are the paradigm of metabolic enzymes whose activities are controlled by sirtuin-dependent, reversible lysine acetylation (sRLA) (6-9). Briefly, AMP-forming Acs enzymes (EC 6.2.1.1) of all domains of life catalyze the reaction shown in reactions 1 and 2. The reactions are reversible as long as pyrophosphate is not hydrolyzed to *o*-phosphate by pyrophosphatase (reaction 3).

Acetate + ATP → acetyl-AMP + PP_i ($\Delta G'^{o}$ = 10.04 kJ / mol) ……(1)

(10)

Acetyl-AMP + CoA → acetyl-CoA + AMP ($\Delta G'^{o}$ = -55.65 kJ / mol)……(2)

(10)

$PP_i + H_2O \rightarrow 2P_i$ ($\Delta G'^{o}$ = - 21.92 kJ / mol)……(3)

(11)

In a cell lacking sirtuin deacetylase, Acs is acetylated by a protein acetyltransferase (Pat) enzyme effectively abolishing reaction (1) (7). Acs proteins across all domains of life are very similar in length (650-700 residues) and contain 10 conserved regions named A1-A10 (12, 13). Region A10 contains the active-site lysine that is modified by GNATs (6).

Given the relevance of Acs to central metabolism, it is not surprising that the activity of Acs orthologues in cells of all domains of life is under sRLA control (14-17). There are exceptions, however. For example, the Acs enzyme of *Streptomyces lividans* (*Sl*Acs) is very weakly acetylated by the type II protein acetyltransferase A enzyme (*Sl*PatA) of this bacterium (18). This observation

is not due to *Sl*PatA being a poor enzyme, because *Sl*PatA can effectively acetylate the *Salmonella enterica* Acs (*Se*Acs) enzyme (19, 20). These results are intriguing, because *Sl*Acs and *Se*Acs share 52% end-to-end identity and an identical acetylation motif (PKTRSGK).

Recent work with the actinomycete *Micromonospora aurantiaca* provided insights into why *Sl*PatA may not acetylate *Sl*Acs as efficiently as expected. In the alluded work, Xu and co-workers identified a GNAT that acetylated the Acs enzyme of *M. aurantiaca* (21). This *M. aurantiaca* GNAT has two domains, a catalytic domain at the *C*-terminus, and an allosteric ACT-domain (ACT standing for aspartate kinase, chorismate mutase, and TyrA, named after the proteins that contain the domain) at the *N*-terminus. Members of this type of GNATs are referred to as type III or ACT-Pats and they have only been found in actinomycetes (22). The regulatory domain of type III protein acetyltransferases are substantially smaller (~400 residues) than those in type I and II protein acetyltransferases (~700 residues), which are found in actinomycetes and other microorganisms (Fig. 5.1) (22). Type I protein acetyltransferases have the large regulatory domain on the *N*-terminus, with type II acetyltransferases having the large regulatory domain on the *C*-terminus (Fig. 5.1). Currently, the regulatory function of this domain is unknown but shares homology to NDP-forming CoA ligases (7).

Here, we show that the *Ma*ACT-Pat enzyme (hereafter referred to as *Ma*PatB) acetylates and deactivates *Sl*Acs, and that the *Ma*PatB homologue from *Streptomyces lividans* (*Sl*PatB) can also acetylate *Sl*Acs *in vivo*. Unexpectedly, we found that in *S. lividans, Sl*Acs was acetylated at two sites within the acetylation motif. One modification was on residue S608 located at the -2 position relative to residue K610 of the acetylation motif, which was also acetylated. *Sl*Acs was acetylated at S608 only when purified from *S. lividans*, but the enzyme had poor activity *in vitro* and could

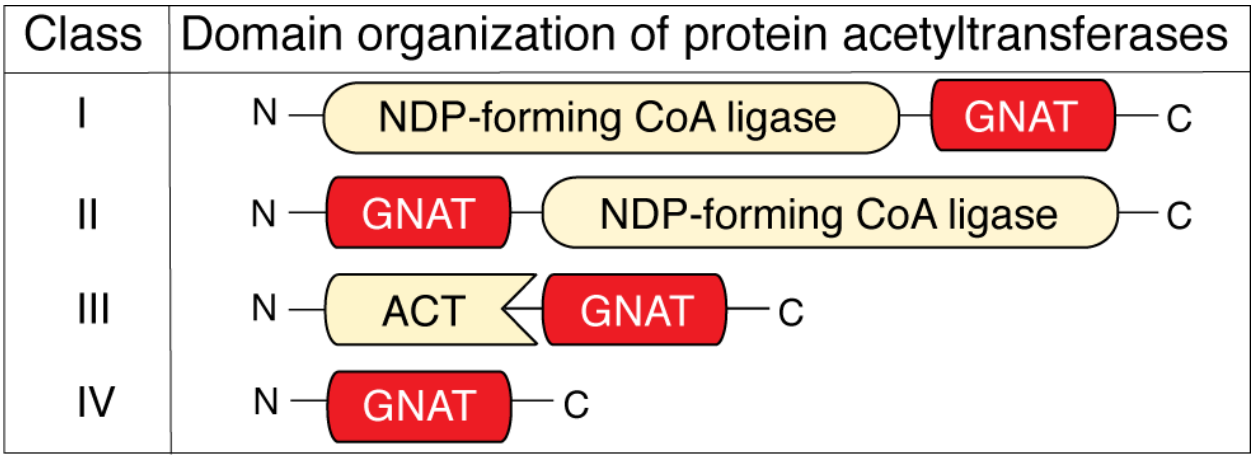


Figure 5.1. Classification of prokaryotic protein acetyltransferases based on domain organization.

not be acetylated by *Ma*PatB. These effects were reversed when a variant of *Sl*Acs (*Sl*Acs^{S608T}) was purified from *S. lividans.* That is, *Sl*Acs^{S608T} isolated from *S. lividans* was active and was acetylated by *Ma*PatB.

*Sl*Acs isolated from *E. coli* was not acetylated at either Ser or Lys, suggesting that the *E. coli* PatZ enzyme (14) could not acetylate K610 of *Sl*Acs, and that *E. coli* lacked the enzyme that modifies S608. At present, it is unclear whether or not Ser acetylation is enzymatic.

Lastly, quantification of *Sl*Acs activity in *S. lividans* strains lacking either of its putative deacetylases (*i.e.,* LdaA, SrtA, CobB) suggested that the SrtA sirtuin of this bacterium was likely responsible for maintaining *Sl*Acs active *in vivo* via lysine deacetylation. To our knowledge, this is the first example of an acyl-CoA (AMP-forming) synthetase whose function is regulated by *O*- and *N*-acetylation.

RESULTS

The *Micromonospora aurantiaca Ma*PatB enzyme acetylates *Streptomyces lividans* acetyl-CoA synthetase (*Sl*Acs) *in vitro.* Previous work from our laboratory showed that *Sl*PatA (a type II Pat, Fig. 5.1) did not recognize *Sl*Acs as substrate, but it did acetylate the acetoacetyl-CoA synthetase (*Sl*AacS) from the same organism (18). Recent work with *Micromonospora aurantiaca* identified a type III Pat enzyme containing an ACT-domain (PF01842, *Ma*PatB, Fig. 5.1) that acetylated its cognate Acs (Micau_0428, hereafter *Ma*Acs) (21). Analysis of the genome of *S. lividans* identified locus EFD70633 as one encoding a homologue of *Ma*PatB (41% identity, Fig. 5.2; hereafter *Sl*PatB). Results in Figure 5.3A show that *Sl*PatA did not recognize *Ma*Acs or *Sl*Acs as substrates, but it did recognize *Sl*AacS, a *bona fide* substrate of *Sl*PatA (18), indicating that *Sl*PatA was active. In contrast, *Ma*PatB acetylated *Sl*Acs at residue K610 (Fig. 5.3B, lanes 2, 3, 4).

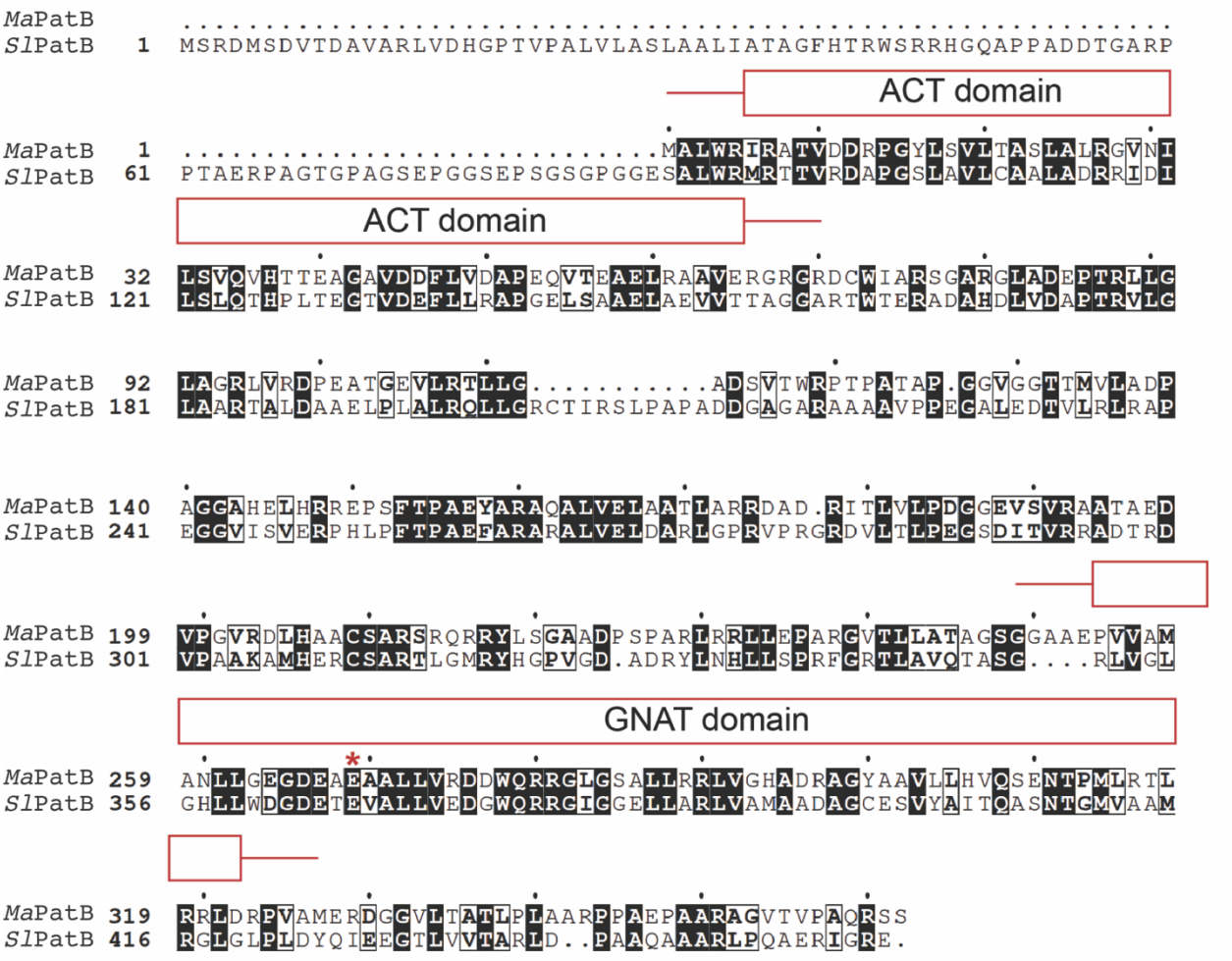


Figure 5.2. Protein alignment of *Ma*PatB and *Sl*PatB. Protein sequences of type III Pat enzymes containing ACT-domains (PF01842), *Ma*PatB (Micau_0428) and *Sl*PatB (EFD70633), were aligned using ESPrit 3.0 (23) with default parameters. *Ma*PatB and *Sl*PatB share 41% identity (BLASTp) and the active site glutamate of these proteins is indicated by the red asterisk.

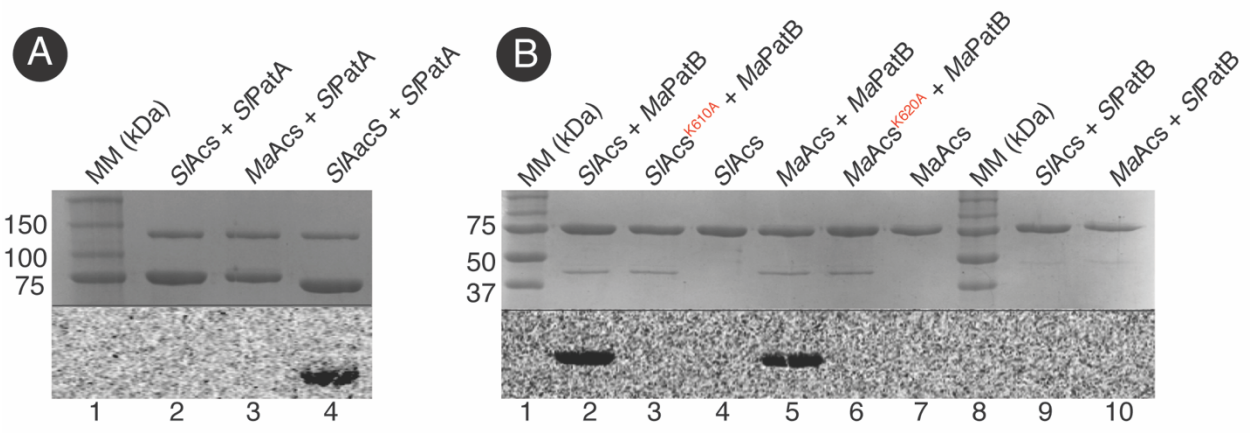


Figure 5.3. *Sl*Acs is acetylated *in vitro* by varying degrees with various protein acetyltransferases. Target proteins of interest (Acs) were incubated with acetyltransferase and [1-^{14}C]-acetyl-CoA. Proteins were separated by SDS-PAGE and stained with Coomassie Brilliant Blue R to visualize proteins. Precision Plus Protein^{TM} (BioRad) standard was used as a molecular marker (MM) with relevant molecular masses indicated in kDa. Acetylation was visualized by phosphor imaging (separated by horizontal black line). A. *Sl*Acs and *Ma*Acs incubated with [1-^{14}C]-acetyl-CoA and ± *Sl*PatA. *Sl*AacS and *Sl*PatA used as a positive control. *Sl*Acs and *Ma*Acs are bands corresponding to 71 kDa, *Sl*AacS at 72 kDa, and *Sl*PatA at 107 kDa. B. *Sl*Acs^{WT}, *Sl*Acs^{K610A}, *Ma*Acs^{WT}, or *Ma*Acs^{K620A}, were incubated with [1-^{14}C]-acetyl-CoA and ± *Ma*PatB or *Sl*PatB (~35 kDa). *Ma*Acs incubated with *Ma*PatB was used as a positive control.

These results were similar to those obtained with control experiments where *Ma*PatB was incubated with its known substrate *Ma*Acs (Fig. 5.3B, lanes 5, 6, 7). Despite the level of identity between *Ma*PatB and *Sl*PatB (Fig. 5.2), *Sl*PatB failed to acetylate *Sl*Acs or *Ma*Acs under the conditions tested (Fig. 5.3B, lanes 9, 10).

Effects of acetylation on *Sl*Acs activity.

Since *Sl*PatB did not acetylate *Sl*Acs *in vitro* but *Ma*PatB did (Fig. 5.3B), we used *Ma*PatB to assess the effect of acetylation on *Sl*Acs activity. The *Sl*Acs enzyme used in the experiments described below was isolated from *E. coli*. Under the conditions tested, *Sl*Acs activity decreased proportionally to the amount of *Ma*PatB used in the experiment (0-3 μM), but leveled off after an ~50-60% decrease even in the presence of excess *Ma*PatB (Fig. 5.4). In these experiments, a specific activity of ~30 μmol AMP min^{-1} mg^{-1} was considered to be 100%. The observed decrease in activity was consistent with *Ma*PatB acetylation on *Ma*Acs (21). The limited reduction in *Sl*Acs activity was unusual in that homologues from other bacteria are >80% deactivated upon acetylation (7, 18, 24-28). We investigated possible reasons for the limited deactivation of *Sl*Acs.

The acetylation state of *Sl*Acs purified from *S. enterica.*

To determine whether the loss of *Sl*Acs activity was due to the acetylation of residue K610, we analyzed the acetylation state of *Sl*Acs proteins synthesized by *S. enterica* using rabbit polyclonal anti-acetyllysine (α-AcK) antibodies. Results of control experiments with non-acetylated *Sl*Acs and *in-vitro* acetylated *Sl*Acs (acetylated by *Ma*PatB) showed that AcK was readily detected under the conditions tested (Fig. 5.5, lanes 17, 18). A signal for AcK was detected in *Sl*Acs purified from *S. enterica* cells that synthesized *Sl*PatB or *Ma*PatB but not from cells harboring vectors coding the inactive variants of both PatB enzymes

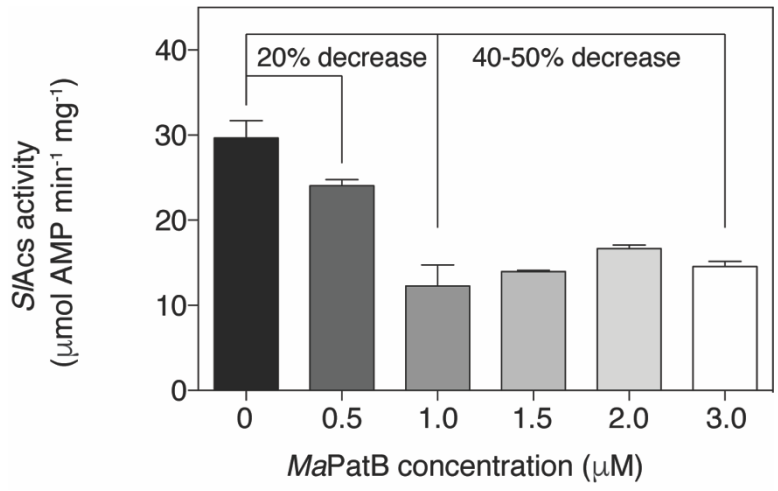


Figure 5.4. *Sl*Acs activity decreases after incubation with acetyl-CoA and *Ma*PatB. *Sl*Acs (2 μM) isolated from *E. coli* (*i.e., Sl*Acs in which residue S608 was not acetylated) was incubated with *Ma*PatB (concentrations indicated on *x*-axis) and acetyl-CoA and specific activity (μmol AMP min^{-1} mg^{-1}) was calculated using a continuous spectrophotometry assay as described in *Experimental procedures*. The activity of *Sl*Acs measured after incubation in the absence of *Ma*PatB was set as a value of 100% activity and *Sl*Acs activity was compared with increasing amounts of *Ma*PatB as indicated by the *y*-axis. Error bars represent standard deviation. Experiment was repeated three times in technical triplicate.

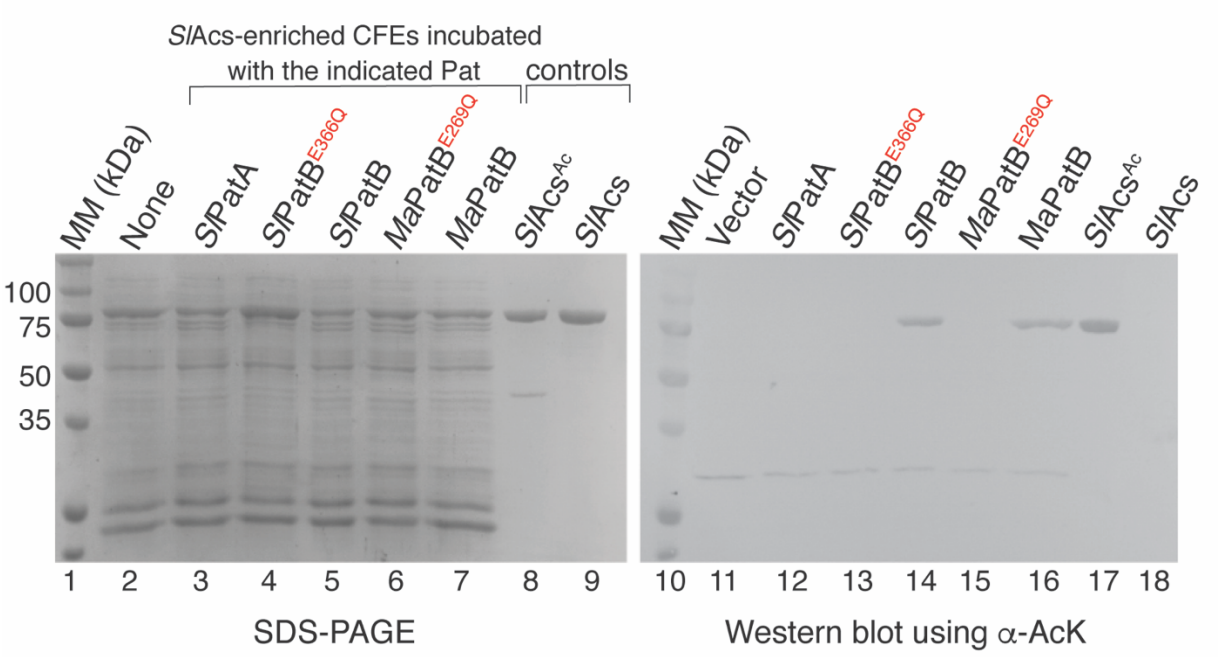


Figure 5.5. *Sl*Acs is acetylated in *S. enterica* strains that synthesize active PatB acetyltransferases. Cell lysates were obtained from Δ*acs* / p*Sl*Acs strains expressing the acetyltransferase as indicated above each lane. p*Sl*PatB^{E366Q} and p*Ma*PatB^{E269Q} refer to plasmids carrying alleles that encode inactive forms of *Sl*PatB or *Ma*PatB, respectively. Acetylation state of *Sl*Acs from lysates (left panel, SDS-PAGE) was analyzed by western blot analysis using anti-acetyllysine (right panel, α-K^{Ac}). *In vitro* acetylated and non-acetylated *Sl*Acs were used as positive and negative controls respectively. *Sl*Acs is the band at ~72 kDa. CFEs, cell-free extracts.

(Fig. 5.5, lanes 13-16). These results raised two points. First, that *Sl*PatB and *Ma*PatB recognized *Sl*Acs as substrate, and second, that residues E366 (*Sl*PatB) and E269 (*Ma*PatB) were important for enzyme activity.

Discovery of *O*-Ser acetylation in *Sl*Acs.

To determine acetylation state of *Sl*Acs when it is synthesized in *S. lividans*, H_6-*Sl*Acs^{WT} protein was purified from *S. lividans*. Unexpectedly, the specific activity of *Sl*Acs^{WT} enzyme isolated from *S. lividans* (~2 μmol AMP min^{-1} mg^{-1}) was 20- to 30-fold lower than that of *Sl*Acs^{WT} purified from *E. coli* (Fig. 5.6A). In fact, the activity of *Sl*Acs purified from *S. lividans* was similar to that of an inactive *Sl*Acs variant (*Sl*Acs^{K610A}; Fig. 5.6A). Results from size-exclusion chromatography experiments (column volume 25 mL with a flow rate of 0.5 ml min^{-1}) showed that the retention time and peak distribution of *Sl*Acs purified from *S. lividans* (24 min) was unchanged from that of *Sl*Acs purified from *E. coli* (24 min, Fig. 5.7), suggesting that the proteins were probably not misfolded.

We note surprisingly, that the *Sl*Acs^{WT} protein purified from *S. lividans* strains was not acetylated by *Ma*PatB *in vitro* compared to the positive control (Fig. 5.6B, lane 2 vs 3.). To determine whether this observation was due to unknown posttranslational modifications on *Sl*Acs, LC/MS/MS peptide fingerprinting was performed on *Sl*Acs purified from different sources, namely i) *Sl*Acs purified from *E. coli* and incubated with *Ma*PatB + AcCoA, ii) *Sl*Acs purified from *E. coli* and incubated with AcCoA but no *Ma*PatB, and iii) *Sl*Acs purified from *S. lividans*. *Sl*Acs^{WT} protein purified from *S. lividans* was acetylated at residue serine (S608) and residue K610 (Fig. 5.8A). In contrast, residue S608 was not acetylated in *Sl*Acs^{WT} protein purified from *E. coli*, and this protein could be acetylated by *Ma*PatB at residue K610 as shown by mass spectrometry (Fig. 5.8B). The SGKIMR peptide of untreated *Sl*Acs protein purified from *E. coli* was not detected

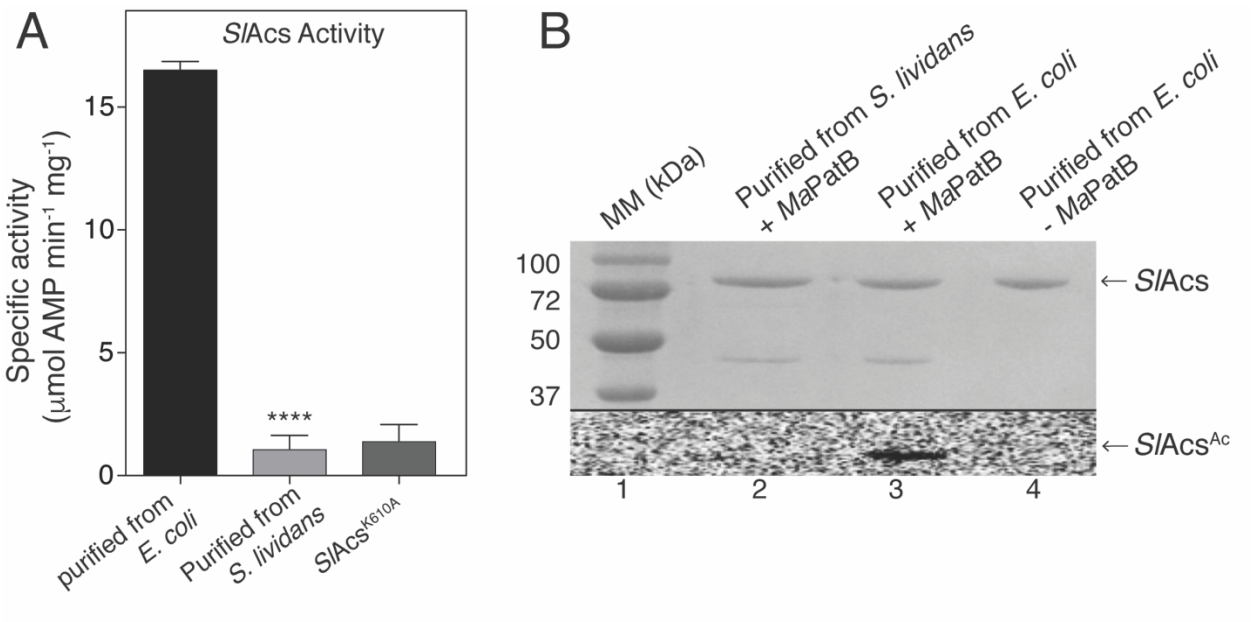


Figure 5.6. *Sl*Acs^{WT} purified from *Streptomyces* strains does not have activity *in vitro* and is not acetylated by *Ma*PatB. A. His$_6$-*Sl*Acs^{WT} was purified from *E. coli* or *Streptomyces* as indicated on *x*-axis, and activity of *Sl*Acs was tested using a spectrophotometry coupled assay. *Sl*Acs purified from *E. coli* was used as a positive control and an inactive variant (*Sl*Acs^{K610A}) was used as a negative control. Specific activity is reported as μmol AMP min^{-1} mg^{-1} of *Sl*Acs. *Sl*Acs^{WT} was purified from *S. lividans* in biological triplicate and each biological replicate was tested in technical triplicate. The technical triplicates of one biological replicate shown as a representative with error bars indicating standard deviation. Asterisks represent significance of *Sl*Acs^{WT} purified from *E. coli* compared to *Sl*Acs^{WT} purified from *S. lividans* strains (indicated on *x*-axis) by calculating P value using an unpaired t-test, **** represents a P value below 0.00005. B. *Sl*Acs^{WT} purified from *Streptomyces* (lane 2) or *E. coli* (lane 3) was incubated with *Ma*PatB and [1-^{14}C]-acetyl-CoA. Last lane (lane 3) was a negative control of *Sl*Acs purified from *E. coli* incubated only with [1-^{14}C]-acetyl-CoA. Proteins were separated by SDS-PAGE and stained with Coomassie Brilliant Blue R to visualize proteins. Precision Plus Protein^{TM} (BioRad) standard was used as a molecular marker (MM) with relevant molecular masses indicated in kDa. Acetylation was visualized by phosphor imaging (separated by horizontal black line). His$_6$-*Sl*Acs at ~72 kDa and *Ma*PatB at ~35 kDa.

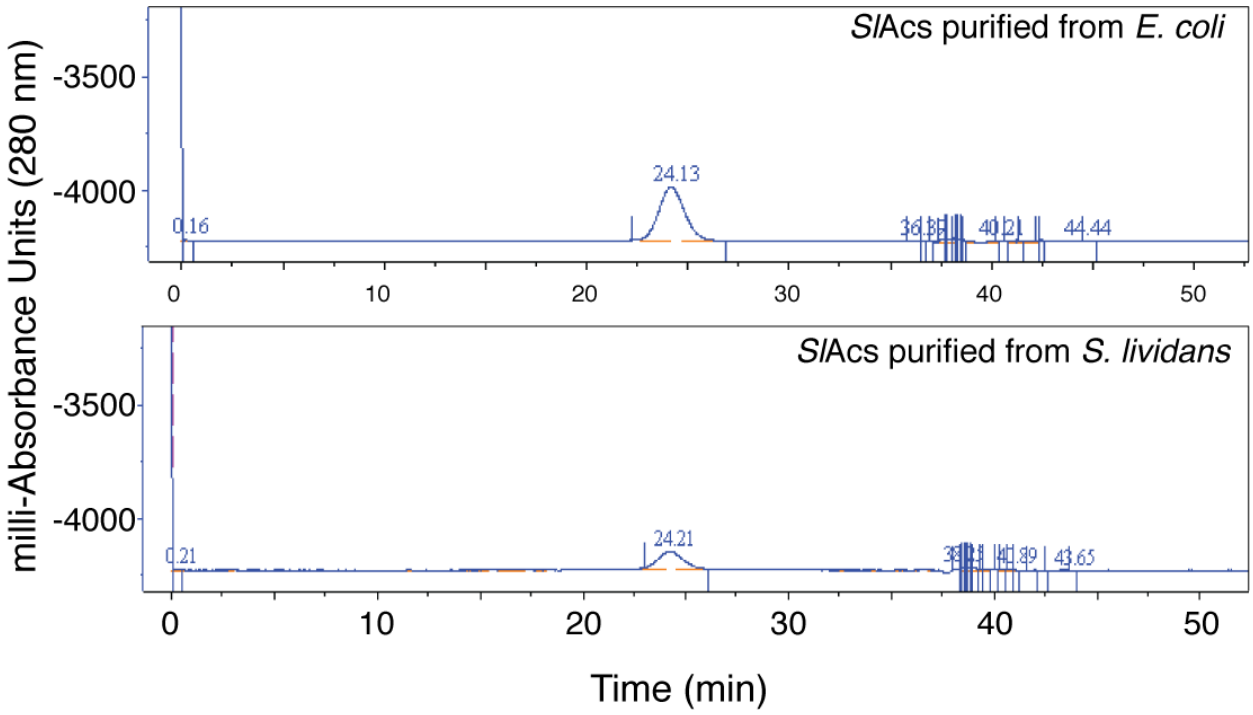
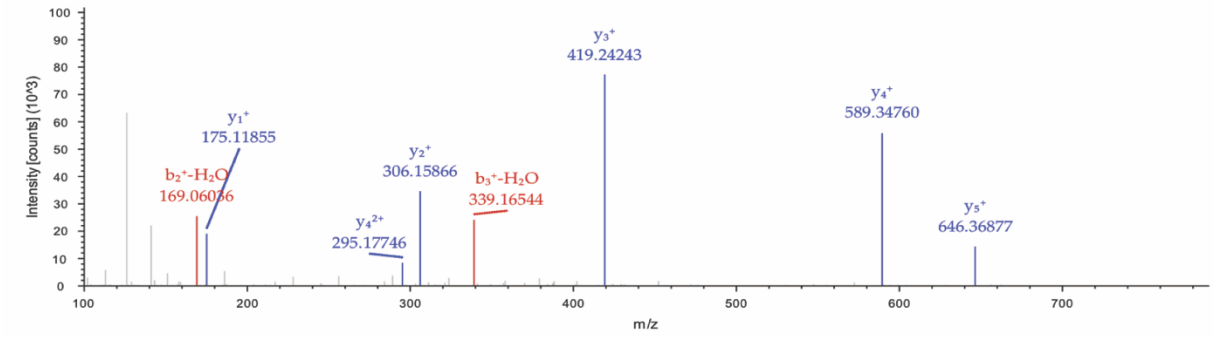


Figure 5.7. *Sl*Acs purified from *E. coli* has the same retention time as *Sl*Acs purified from *S. lividans*. Protein (1 μM, 100 μL) was loaded onto a 25 mL Superose 12 10/300 GL size exclusion column and was run at a flow rate of 0.5 ml min^{-1}. Elution was monitored at 280 nm and retention times were compared to a standard curve plotted from retention times of gel filtration standards (BioRad). Time above peaks corresponds to retention time of *Sl*Acs purified from *E. coli* (top panel) or from *S. lividans* (bottom panel).

A

*Sl*Acs purified from *Streptomyces*

#1	b^+	$b\text{-}H_2O^+$	b^{2+}	Seq.	y^+	y^{2+}	#2
1	130.04987	112.03931	65.52857	S-Acetyl			6
2	187.07134	169.06078	94.03931	G	646.37054	323.68891	5
3	357.17688	339.16631	179.09208	K-Acetyl	589.34907	295.17817	4
4	470.26095	452.25038	235.63411	I	419.24353	210.12540	3
5	601.30145	583.29088	301.15436	M	306.15946	153.58337	2
6				R	175.11896	88.06312	1



B

*Sl*Acs purified from *E. coli* and acetylated *in vitro* with *Ma*PatB

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	88.03931	44.52329	S			6
2	145.06078	73.03403	G	646.37054	323.68891	5
3	315.16631	158.08679	K-Acetyl	589.34907	295.17817	4
4	428.25038	214.62883	I	419.24353	210.12540	3
5	559.29088	280.14908	M	306.15946	153.58337	2
6			R	175.11896	88.06312	1

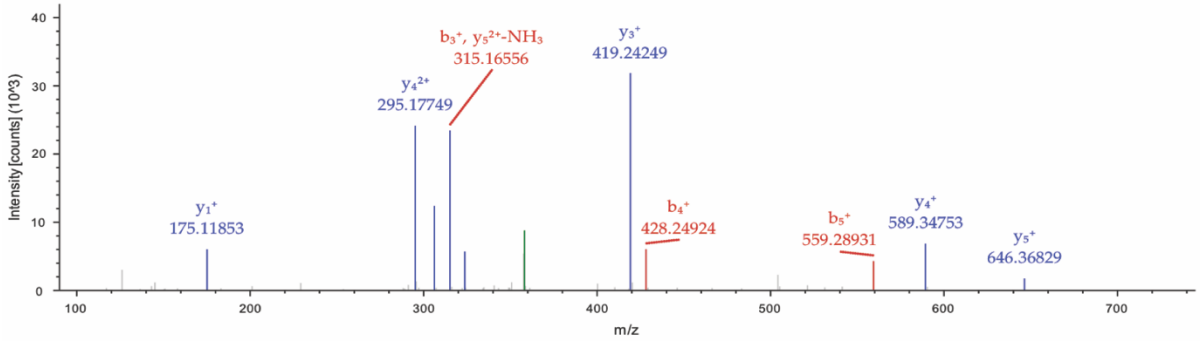


Figure 5.8. *Sl*Acs purified from *Streptomyces* is acetylated at S608 and K610. A. $S^{Ac}GK^{Ac}IMR$ peptide identified by LC/MS/MS analysis of *Sl*Acs purified from *S. lividans*. B. SGKIMR peptide identified by LC/MS/MS of *Sl*Acs acetylated by *Ma*PatB *in vitro*. No peptide was detected for non-acetylated control. *b* ions are the series of fragments that extend from the *N*-terminus; *y* ions are the series of ions that extend from the *C*-terminus. MASCOT software (http://www.matrixscience.com) was the online search engine used to identify peptides on the basis of their masses.

in the LC/MS/MS most likely because trypsin cleaved this peptide at K and R, rendering it too small for detection (data not shown).

In the past, our group reported that single substitutions in the acetylation motif can render an acyl-CoA synthetase non-acetylatable (9). For this reason, the *S. lividans acs*$^{+}$ allele was mutated to encode a S608T variant (*Sl*Acs^{S608T}). Threonine was chosen because Thr is a residue that is conserved in other *S. lividans* CoA synthetases (18). The specific activity of variant H_6-*Sl*Acs^{S608T} purified from *S. lividans* was ~25% lower than that of *Sl*Acs^{WT} purified from *E. coli* (Fig. 5.9A, black *vs* blue bar), but more importantly, the H_6-*Sl*Acs^{S608T} protein purified from *S. lividans* became a substrate for *Ma*PatB *in vitro* (Fig. 5.9B, lanes 2-7). Notably, the level of activity of *Sl*Acs^{WT} isolated from *S. lividans* was not different than the level of the activity of the inactive variant *Sl*Acs^{K610A} also isolated from *E. coli* (Fig. 5.9, red *vs* gray bar). From an enzyme functionality standpoint, this result showed that although Ser and Thr have a hydroxyl group in their side chains that could be modified, the Ser side chain was acetylated, inactivating the enzyme, while the *Sl*Acs^{S608T} protein was not, resulting in active enzyme.

The *Sl*SrtA enzyme (a SIRT4-like sirtuin) deacetylates *Sl*Acs^{Ac} in *S. lividans*.

To determine whether acetylated *Sl*Acs^{S608T} could be enzymatically deacetylated, we looked for genes in the *S. lividans* genome putatively encoding protein deacetylases; we found three such genes (*ldaA, srtA, and cobB*). One of them encoded a homologue of the NAD^{+}-dependent SIRT5 deacetylase, (a homologue of *Se*CobB), another one encoded a homologue of the NAD^{+}-dependent SIRT4 sirtuin, and a third one encoded a homologue of the Zn-dependent, acetate-forming *Rp*LdaA deacetylase (18, 27). H_6-*Sl*Acs^{S608T} protein was purified from five *S. lividans* strains, namely i) *patB*$^{+}$ *ldaA*$^{+}$ *cobB*$^{+}$ *srtA*$^{+}$, ii) Δ*patB ldaA*$^{+}$ *cobB*$^{+}$ *srtA*$^{+}$, iii) *patB*$^{+}$Δ*ldaA cobB*$^{+}$ *srtA*$^{+}$, iv) *patB*$^{+}$ Δ*ldaA* Δ*cobB*

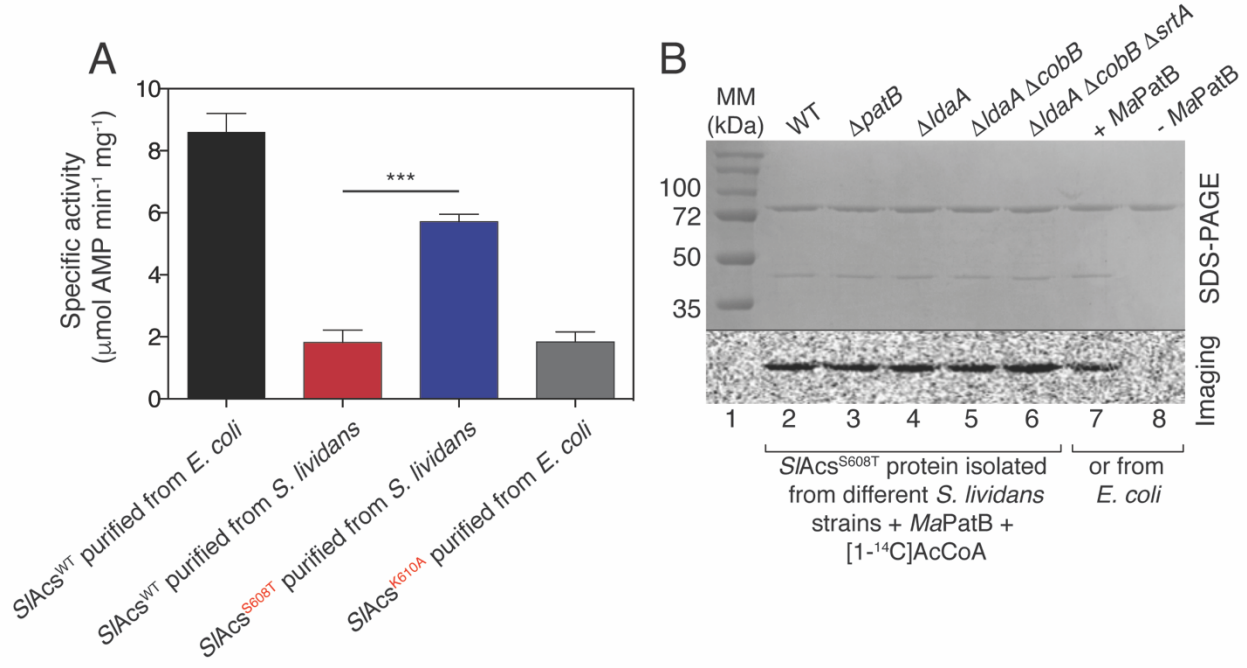


Figure 5.9. Variant *Sl*Acs^{S608T} has higher activity compared to *Sl*Acs^{WT} when purified from *Streptomyces*, and *Sl*Acs^{S608T} is acetylated *in vitro* by *Ma*PatB. Panel A. His$_6$-*Sl*Acs^{S608T} and His$_6$-*Sl*Acs^{WT} was purified from *Streptomyces* and *Sl*Acs activity was calculated using a continuous spectrophotometric assay and reported in μmol AMP min^{-1} mg^{-1}. *Sl*Acs was purified from *S. lividans* in biological triplicate and each biological replicate was tested in technical triplicate. The technical triplicates of one biological replicate shown as a representative with error bars indicating standard deviation. *Sl*Acs purified from *E. coli* was used as positive control and an inactive variant (*Sl*Acs^{K610A}) was used as negative control. Asterisks represent significance of *Sl*Acs^{WT} compared to *Sl*Acs^{S608T} by calculating P value using an unpaired t-test, *** represents a P value below 0.0005. Panel B. His$_6$-*Sl*Acs^{S608T} purified from *Streptomyces* strains (as indicated above lanes) or *E. coli* was incubated with *Ma*PatB and [1-^{14}C]-acetyl-CoA. Last lane was a negative control of His$_6$-*Sl*Acs purified from *E. coli* incubated with only [1-^{14}C]-acetyl-CoA.

srtA^{+}, and v) *patB*$^{+}$ Δ*ldaA* Δ*cobB* Δ*srtA*. The H_6-*Sl*Acs^{S608T} variant was purified from each strain and its activity was quantified. The activity of H_6-*Sl*Acs^{S608T} was not significantly different among proteins purified from *patB*$^{+}$ *ldaA*$^{+}$ *cobB*$^{+}$ *srtA*$^{+}$, *patB*$^{+}$Δ*ldaA cobB*$^{+}$ *srtA*$^{+}$, or *patB*$^{+}$ Δ*ldaA* Δ*cobB srtA*$^{+}$ strains (Fig. 5.10A). However, the activity of *Sl*Acs^{S608T} isolated from the *patB*$^{+}$ Δ*ldaA* Δ*cobB* Δ*srtA* strain was 50% lower than that of the H_6-*Sl*Acs^{S608T} variant purified from the control strain *patB*$^{+}$ *ldaA*$^{+}$ *srtA*$^{+}$ *cobB*$^{+}$ (Fig. 5.10A, black *vs* yellow bar)). This reduction of activity strongly suggested that in *S. lividans, Sl*Acs was under reversible lysine acetylation control exerted by the *Sl*PatB and SrtA enzymes. In *S. coelicolor*, SrtA is equivalent to CobB1, which was shown to deacetylate *Sc*Acs^{Ac} (29). Unfortunately, *S. lividans* SrtA and CobB deacetylases were not active when purified from *E. coli* (18). Based on the *in vivo* data reported in Figure 5.10 we propose that SrtA is likely the deacetylase that used *Sl*Acs^{Ac} as substrate in *S. lividans*.

Acetylation of the *Sl*Acs^{S608T} variant protein by *Ma*PatB decreases its activity.

To assess the effect of acetylation on *Sl*Acs^{S608T} activity, *Sl*Acs^{S608T} and *Sl*Acs^{WT} were purified from *E. coli* and were acetylated (5 μM) *in vitro* by *Ma*PatB (1 μM). When *Sl*Acs^{WT} was acetylated, it lost ~20% activity compared to the non-acetylated control (Fig. 5.10B). These results are comparable to the data in Figure 5.4, where a 20% decrease in *Sl*Acs^{WT} activity was seen with similar ratios of *Sl*Acs to *Ma*PatB (i.e ~1:5) In contrast, when *Sl*Acs^{S608T} was acetylated under the same conditions it lost ~80% activity (Fig. 5.10B). These results raised important questions about the control of CoA synthetases in *S. lividans* and the potential sequential control of CoA synthetases through posttranslational modifications of multiple residues. These data also raised questions about how this protein may be deacetylated, and if the deacetylation of serine or lysine is dependent on deacetylation of the subsequent residue. These results are discussed below.

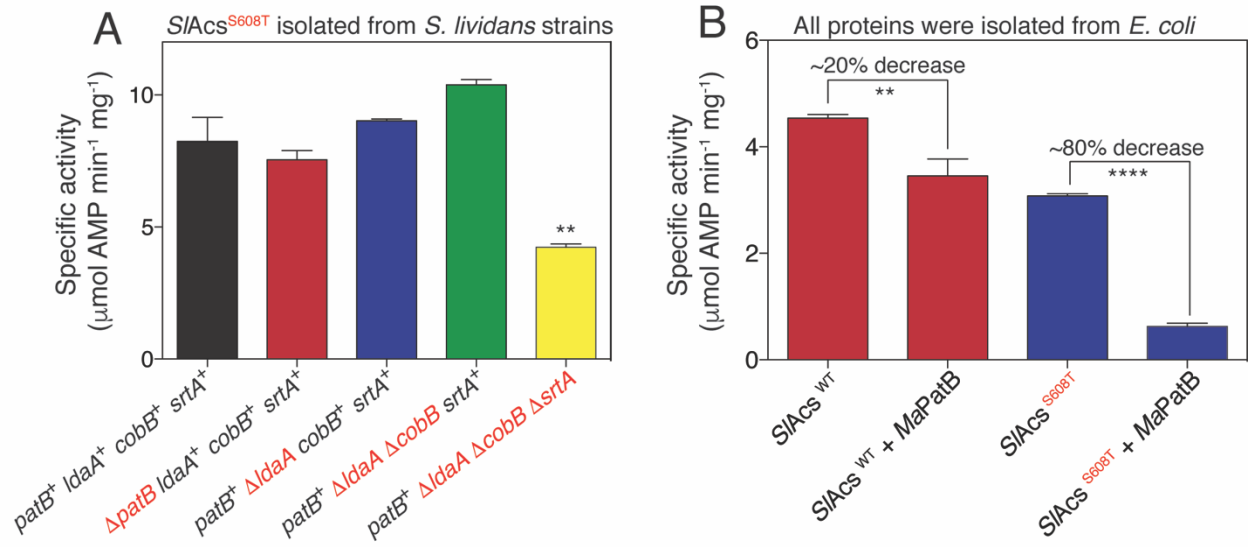


Figure 5.10. *Sl*Acs is deacetylated by SrtA in *S. lividans*. A. *Sl*Acs^{S608T} was purified from *S. lividans* strains indicated on the *x*-axis. Specific activity is reported as µmol AMP min^{-1} mg^{-1} of *Sl*Acs. *Sl*Acs^{S608T} was purified from *S. lividans* in biological triplicate and each biological replicate was tested in technical triplicates. Genes deleted in strains from which *Sl*Acs^{S608T} was isolated are shown in red type. The technical triplicates of one biological replicate shown as a representative with error bars indicating standard deviation. Asterisks represent significance compared to *patB*$^+$ *ldaA*$^+$ *cobB*$^+$ *srtA*$^+$ by calculating P value using an unpaired t-test, ** represents a P value below 0.005. B. *Sl*Acs^{WT} or *Sl*Acs^{S608T} (5 µM) both isolated from *E. coli,* was incubated with Acetyl-CoA with or without *Ma*PatB (1 µM). Activity of *Sl*Acs was monitored using a continuous spectrophotometry assay and activity reports as µmol AMP min^{-1} mg^{-1} of *Sl*Acs. **** represents a P value of < 0.0001. Experiment was repeated in triplicate and error bars represent standard deviation of triplicates from one replicate.

Initial *in vitro* studies of protein substrate specificity among Pat enzymes.

The differences in *Sl*Acs acetylation by *Sl*PatA and *Ma*PatB (Fig. 5.3A, B) raised questions about the specificity of Pat enzymes for Acs substrates. To further examine this specificity, *Sl*Acs was incubated with other Pat enzymes. We found that *Rhodopseudomonas palustris* Pat (*Rp*Pat, a type I enzyme) efficiently acetylated *Sl*Acs (Fig. 5.11A, lanes 2, 3, 4) compared to the positive control of *Rp*Pat with its known substrate *Rp*BadA (Fig. 5.11A, lane 8) (26). In contrast, incubation of *Sl*Acs with *Salmonella enterica* Pat (*Se*Pat) yielded a signal that was only ~16% as strong as the signal obtained with *Se*Acs, a *bona fide* substrate of *Se*Pat (Fig. 5.11, lane 8). Although *Ma*Acs and *Sl*Acs are 76% identical (Fig. 5.12), *Se*Pat acetylated *Ma*Acs more efficiently (~5-fold stronger signal than *Sl*Acs) (Fig. 5.11B, lanes 2 *vs* 5). With the exception of *Se*Pat, all other Pat enzymes tested acetylated *Ma*Acs and *Sl*Acs very similarly (Fig. 5.3).

***In vivo* evidence supports the observed difference in substrate specificity among Pat enzymes.**

We performed *in vivo* experiments to determine whether the differences in protein substrate specificity of Pat enzymes observed *in vitro,* could be replicated under conditions where Acs activity was required for growth. To do this, we grew *S. enterica* strains with low concentrations of acetate (10 mM), a condition where Acs activity is needed (12). In wild-type *S. enterica*, the absence of CobB results in a substantial decrease in cell growth with 10 mM acetate (30). Such a phenotype is due to the acetylation of Acs by Pat (7). Surprisingly, a *S. enterica* Δ*acs* Δ*cobB pat*$^+$ strain carrying a plasmid encoding *Sl*Acs grew better than the *acs*$^+$ *cobB*$^+$ *pat*$^+$ strain with 10 mM acetate (Fig. 5.13, compare gray squares, black circles, white diamonds). These results were consistent with the idea that *Se*Pat, which was present in all strains tested, did not acetylate *Sl*Acs *in vivo* to the point of affecting the ability of the cell to grow with low acetate concentrations. In contrast, when the Δ*acs* Δ*cobB pat*$^+$ strain carrying a plasmid encoding *Sl*Acs also carried the

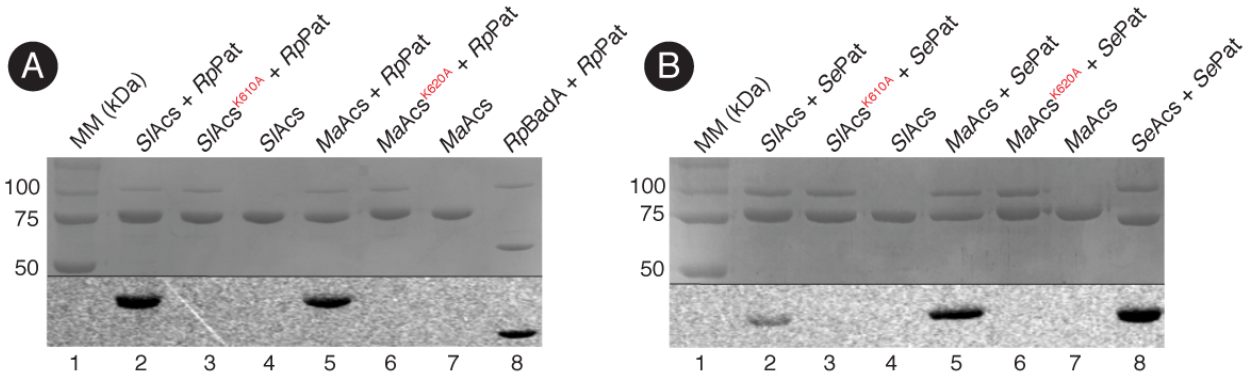


Figure 5.11. *Sl*Acs is acetylated *in vitro* to varying degrees by protein acetyltransferases. Target proteins of interest (Acs) were incubated with acetyltransferase and [1-^{14}C]-acetyl-CoA. Proteins were separated by SDS-PAGE and stained with Coomassie Brilliant Blue R to visualize proteins. Precision Plus Protein^{TM} (BioRad) standard was used as a molecular marker (MM) with relevant molecular masses indicated in kDa. Acetylation was visualized by phosphor imaging (separated by horizontal black line). A. *Sl*Acs^{WT}, *Sl*Acs^{K610A}, *Ma*Acs^{WT}, or *Ma*Acs^{K620A}, were incubated with [1-^{14}C]-acetyl-CoA and ± *Rp*Pat (~90 kDa). *Rp*BadA incubated with *Rp*Pat was used as a positive control. B. *Sl*Acs^{WT}, *Sl*Acs^{K610A}, *Ma*Acs^{WT}, or *Ma*Acs^{K620A}, were incubated with [1-^{14}C]-acetyl-CoA and ± *Se*Pat. *Se*Acs incubated with *Se*Pat (100 kDa) was used as a positive control.

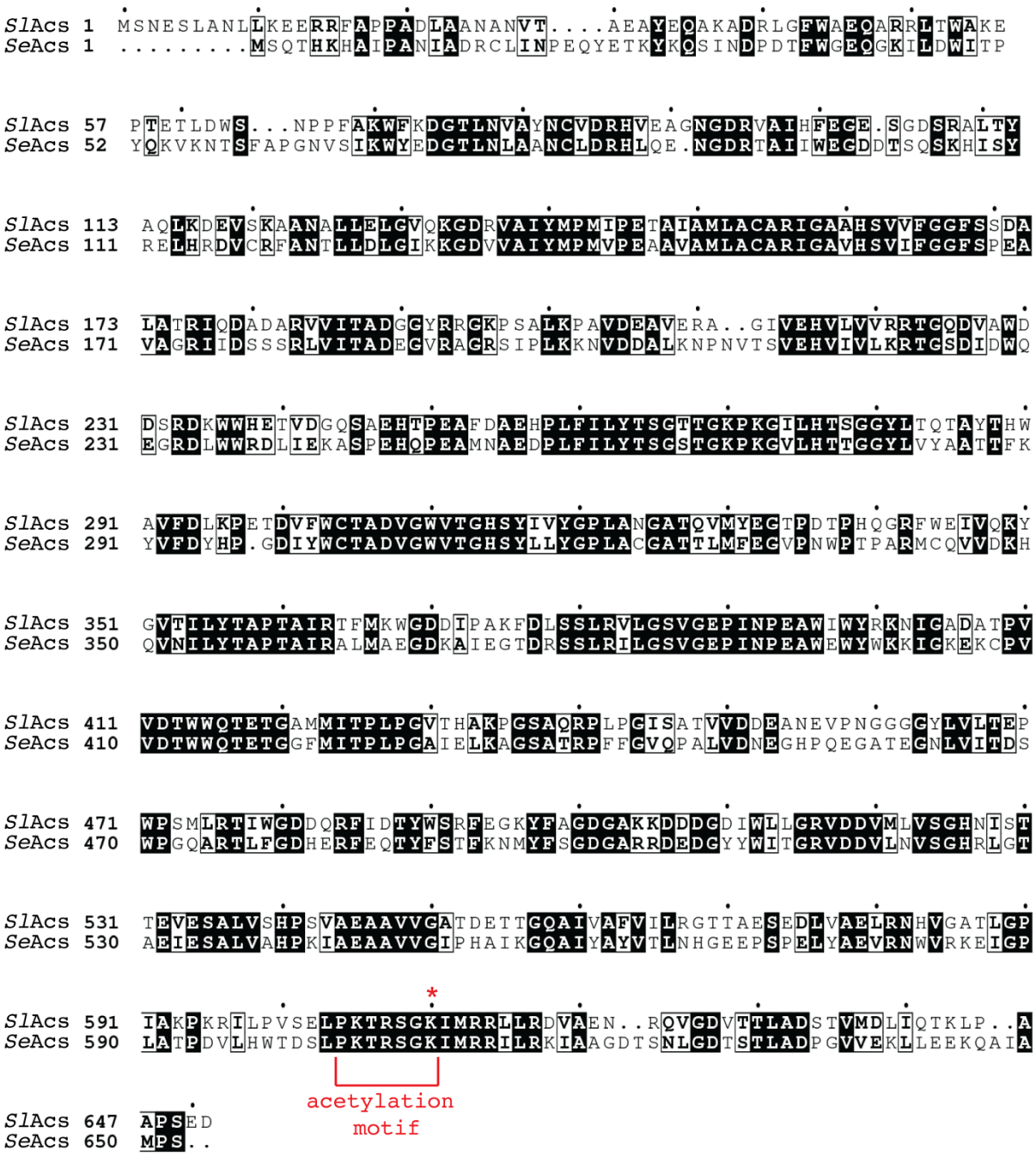


Figure 5.12. Primary sequence alignment of AMP-forming acetyl-CoA synthetases from *Streptomyces lividans* (*Sl*Acs) and *Salmonella enterica* (*Se*Acs). The alignment was generated using ClustalW and ESPript 3.0 software programs. Protein identity 52%; similarity: 67%. Red asterisk designates acetylation sites, which are modified by *Se*Pat (K609 in *Se*Acs) or *Sl*PatB (K610 in *Sl*Acs). The blue asterisk identifies residue S608 of *Sl*Acs, which is modified in *S. lividans* but the modifier has not been identified. The *Se*Acs enzyme is not modified in *S. enterica.*

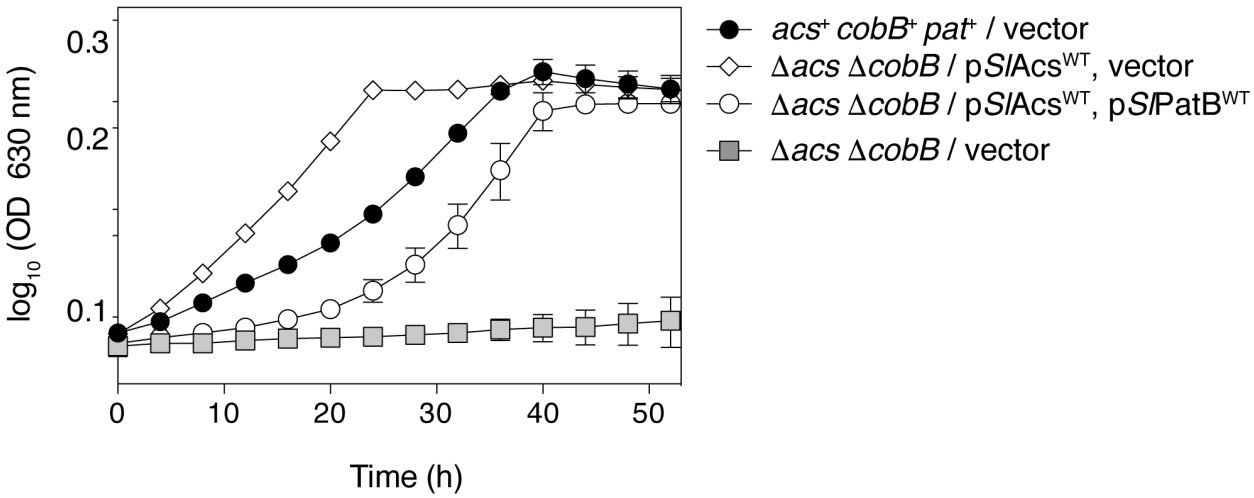


Figure 5.13. *S. enterica Se*Pat and *Se*CobB do not acetylate or deacetylate *Sl*Acs *in vivo*. The gene coding for *Sl*Acs was cloned into a complementation vector, pCV1, and introduced into a Δ*acs* Δ*cobB S. enterica strain.* Genes encoding acetyltransferases *Sl*PatA or *Sl*PatB were cloned into a second complementation vector, pCV3, and introduced into the above-mentioned strain. Strains designated as vector (*i.e., acs*$^+$ *cobB*$^+$ *pat*$^+$ / vector) harbor both pCV1 and pCV3 as controls. Overnight cultures were grown in NB + antibiotic. Cells were sub-cultured (1% v/v) and grown on NCE minimal medium supplemented with acetate (10 mM). Growth curves were obtained in technical triplicate and was repeated three times using a microplate reader (BioTek instruments). Error bars represent standard deviation of technical triplicates.

p*Sl*PatB plasmid growth was delayed for ~20 h, after which the cultures grew at a rate comparable to that of the wild-type strain and reached full density (Fig 5.13, black *vs* white circles).

Exploring the specificity of Pat enzymes for *Sl*Acs using an *in vivo* heterologous system.

To monitor *Sl*Acs activity and its control by different Pat enzymes *in vivo*, we used *Salmonella enterica* strains in which the posttranslational control of Acs and the effects of the absence of control elements (*i.e.*, Pat, CobB) on growth have been extensively characterized (6-8, 30). Several genes were cloned, introduced into a *S. enterica* Δ*acs* strain. The growth behavior with 10 mM acetate as the sole source of carbon and energy was monitored as a function of time. The presence of the *pat*$^+$ allele in all strains was not a concern since *Se*Pat did not acetylate *Sl*Acs extensively *in vitro* (Fig. 5.11B). As discussed below, results from the experiments confirmed this observation.

A *S. lividans acs*$^+$ allele coding for the wild type protein with a hexahistidine *N*-terminal fusion was cloned into an arabinose inducible complementation vector encoding resistance to ampicillin (31). The genes encoding different types of Pat enzymes included *S. lividans patA*$^+$, *S. lividans patB*$^+$, and *M. auriantica patB*$^+$, which were cloned into a compatible arabinose inducible vector encoding chloramphenicol resistance and the resulting plasmids were transformed into strain JE21742 (Δ*acs* / p*Sl*Acs). The resulting strains were grown in minimal medium containing acetate (10 mM). Under those conditions cell growth required Acs function. We predicted that acetylation of Acs would lower its activity reducing the growth rate and final cell density of the cultures.

In Figure 5.14, the behavior of positive- and negative-control strains growing in minimal medium with low acetate (10 mM) as the sole source of carbon and energy is shown by black circles (*acs*$^+$) and grey squares (Δ*acs*), respectively. When the Δ*acs* strain synthesized *Sl*Acs,

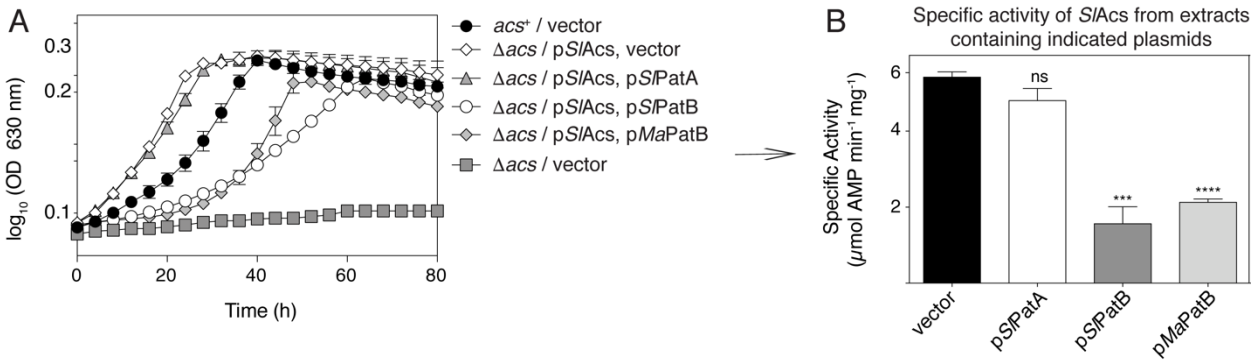


Figure 5.14. Overexpression of *patB* acetyltransferases in *S. enterica* lowers *Sl*Acs activity in cell lysates. A. The gene coding for *Sl*Acs was cloned into pCV1 and introduced into a *S. enterica* Δ*acs* strain carrying genes coding for acetyltransferases of interest cloned into pCV3. Overnight cultures were grown in NB + antibiotic. Cells were sub-cultured (1% v/v) and grown on NCE minimal medium supplemented with acetate (10 mM). Growth curves were obtained in technical triplicate and was repeated three times using a microplate reader (BioTek instruments). Error bars represent standard deviation of technical triplicates. B. Cells grown in 1 L of same medium conditions from A were harvested and lysed and *Sl*Acs specific activity (μmol AMP min^{-1} mg^{-1}) was quantified using a continuous spectrophotometry assay. *Sl*Acs was isolated from strains containing plasmids listed below bars. Error bars represent standard deviation of technical triplicates. Asterisks represent significance compared to vector control by calculating P value using an unpaired t-test, *** represents a P value below 0.0005, **** represents a P value below 0.00005, and ns represents no significance.

growth occurred without delay (Fig. 5.14A, white diamonds), indicating that *Sl*Acs was functional. Notably, growth of the Δ*acs* strain that synthesized *Sl*PatA and *Sl*Acs was not different than that of the Δ*acs* strain that only synthesized *Sl*Acs (Fig. 5.14A, white diamonds *vs* grey triangles). These results suggested that neither *Se*Pat nor *Sl*PatA inactivated *Sl*Acs. In contrast, cultures of Δ*acs* strains that synthesized *Sl*Acs and either *Sl*PatB or *Ma*PatB displayed an ~30-h long lag phase before the onset of exponential growth (Fig. 5.14A, white diamonds *vs* grey diamonds or white circles). We hypothesized that the observed lag phase was associated with an increase in *Sl*Acs acetylation, which would result in lower *Sl*Acs activity. To test this idea, the same strains were grown with acetate, cells were collected at mid-log (OD_{630} ~0.15), and lysed and *Sl*Acs activity was quantified. *Sl*Acs activity decreased by 10% in Δ*acs* / p*Sl*Acs cells harboring a plasmid encoding *Sl*PatA compared to the vector-only control (Fig. 5.14B). In contrast, a ~60-70% decrease in *Sl*Acs activity was measured when the enzyme was purified from strain Δ*acs* / p*Sl*Acs / p*Sl*PatB or from strain Δ*acs* / p*Sl*Acs / p*Ma*PatB. These percent activities were relative to that of *Sl*Acs purified from cells that did not express an acetyltransferase (Fig. 5.14B, black bar). These results were consistent with the idea that that the above-mentioned lag phase was likely due to a decrease in *Sl*Acs activity due to acetylation. These results also suggested that although *Sl*PatB was not active *in vitro*, it did acetylate *Sl*Acs *in vivo*, thereby reducing its activity.

DISCUSSION

The fact that the AMP-forming acetyl-CoA synthetase from *Streptomyces lividans* (*Sl*Acs) was poorly acetylated by *Sl*PatA (18) was the driving force for the work presented in this paper. Here we show that *Sl*Acs activity is controlled by sirtuin-dependent reversible lysine acetylation (sRLA), and that the enzyme responsible for this modification is a type III protein acetyltransferase

thus far found only in actinomycetes. The sRLA control of *Sl*Acs activity is novel because of the unprecedented involvement of an *O*-acetylation modification of the side chain of a serine residue close (position -2) to the lysyl chain that is typically acetylated in AMP-forming acyltransferases. Serine and threonine acetylation in bacteria was only recently reported, with the paradigm acetyltransferase being YopJ from *Yersinia pestis* (32). YopJ acetylates lysine, threonine, and serine residues of host kinases (33). While YopJ is a prokaryotic acetyltransferase that *O*-acetylates a eukaryotic protein, *Sl*Acs is the first example of a prokaryotic AMP-forming acyl-CoA synthetase that is controlled by *O*-Ser and N^{ε}-Lys acetylation. Additionally, a recent acetylome study in *Mycobacterium tuberculosis* identified many acetylated serine residues on a myriad of proteins (34). Although the enzyme responsible for acetylating serine residues has not yet been identified in *M. tuberculosis*, these studies highlight the prevalence of *O*-serine acetylation in prokaryotes, an area that needs further investigation.

Type III protein acetyltransferases acetylate Acs in Actinomycetes.

In several Alpha- and Gamma-proteobacteria, AMP-forming Acs enzymes are regulated by type I protein acetyltransferases [for a review see (22)]. Here we identified the *Sl*PatB enzyme (a type III protein acetyltransferase) of *Streptomyces lividans* as the enzyme that controls the activity of *Sl*Acs at the expense of AcCoA (Fig. 5.3). *Sl*PatB is a homologue of *Ma*PatB from *Micromonospora aurantiaca*, which was shown to be allosterically regulated by cysteine and arginine (21). It is likely that Acs control in *S. lividans* is further affected by physiological conditions that modulate cysteine and arginine metabolism, given that fact that *Ma*PatB activity is allosterically regulated by these two amino acids. At present, it is unclear whether or not *Sl*PatB is allosterically controlled because this enzyme was not active under the *in vitro* conditions tested. Bioinformatics analyses show that only genomes of actinomycetes encode type III protein

acetyltransferases, raising important questions about the possible link between amino acid metabolism and protein acetylation in these bacteria.

_O_-Serine and _N_$^{\varepsilon}$-lysine acetylation controls the activity of _Sl_Acs in _Streptomyces lividans_.

The regulation of Acs in *S. lividans* appears to be more complex than that of Acs homologues found in other bacteria. To our knowledge, *Sl*Acs is the first AMP-forming AcCoA synthetase whose activity is posttranslationally controlled by *O*-Ser and reversible N^{ε}-Lys acetylation.

*Sl*Acs synthesized by *S. lividans* was acetylated at the expected Lys residue within the acetylation motif (site 0), and at residue S608 located two positions upstream the Lys acetylation site (site -2; Fig. 5.8A). At this point, it is not clear what the order of acetylation may be, however, we propose that two different acetyltransferases modify *Sl*Acs in *S. lividans*, since synthesis of *Sl*Acs in *E. coli* did not yield enzyme acetylated at S608, suggesting that *E. coli* lacks the *O*-acetyltransferase that modifies residue S608 (Fig. 5.8B). Although chemical (*i.e.* non-enzymatic) serine acetylation is rare, at present we cannot rule out this possibility. It appears unlikely that this would be the case, since *Sl*Acs isolated from *E. coli* S608 was not acetylated. Further studies on S608 acetylation in *Streptomyces* are needed.

We suggest that in *S. lividans* full downregulation of *Sl*Acs activity only occurs when S608 is acetylated (Fig. 5.10B). We suggest that this is the case because when we attempted to acetylate *Sl*Acs isolated from *E. coli*, *Sl*Acs activity was only decreased by 20-40%, even after extended incubation with the acetyltransferase. Since *E. coli* does not acetylate S608, the acetyltransferase can modify K610. In contrast, *Sl*Acs isolated from *S. lividans* was modified at residue S608 (Fig. 5.8A) and lacked activity (Fig. 5.6). The absence of activity correlated with the presence of acetylated serine, because when serine variant *Sl*Acs^{S608T} was isolated from *S. lividans* strong

*Sl*Acs activity was measured (Fig. 5.9A). The results with *Sl*Acs^{S608T} also suggest that the putative serine acetyltransferase cannot recognize threonine as its substrate.

Acetylation of S608 is of interest given the location of this residue within the acetylation motif found in CoA synthetases [PX_4GK, (9)]. In *Sl*Acs, the acetylation motif is PKTR**S**G**K**^{610}, with the modified serine and lysine residues shown in bold typeface. Based on the data reported here, it appears like there are several layers of control of *Sl*Acs in *S. lividans.* At present is not known whether S608 acetylation is reversible, but if it were, the identification of the S608^{Ac} deacetylase would be an important step towards understanding the integration of two reversible acetylation/deacetylation events that control S608 and K610 modification. Why would such a mechanism of Acs posttranslational control be needed in *Streptomyces*? Is this control mechanism used to modulate the activity of other enzymes in prokaryotes or eukaryotes? These and other interesting questions warrant further investigation in this area of *Streptomyces* physiology.

An opportunity to identify determinants that play a critical role in acyltransferase specificity.

Questions regarding how acyltransferases recognize their protein targets are of interest. We have published structural studies aimed at identifying residues contributing affecting the recognition and specificity of *Sl*PatA for *Se*Acs (20). One limitation of the alluded work was that it was performed with proteins from different sources. The studies reported here open the door for structural work involving enzymes partners from the same organism. A better understanding of the specificity of acyltransferases for their protein targets may help identify such targets.

Deacetylation of *Sl*Acs in *Streptomyces.*

The *S. lividans* genome codes for three deacetylases, two of which are NAD^+-dependent sirtuin deacylases (35). The two sirtuins in this bacterium are different from each other. One of them, *Sl*CobB (formerly CobB2), is homologous to human SIRT5 a known demalonylase,

desuccinylase (36). The other, *Sl*SrtA (formerly CobB1) is homologous to human SIRT4, a known ADP-ribosyltransferase (37). SIRT5-like enzymes are the best characterized family of sirtuins in bacteria (*e.g., Se*CobB, *Rp*SrtN, *Ec*CobB).

Results presented in Figure 5.10A suggest that *Sl*SrtA is responsible for maintaining *Sl*Acs deacetylated (thus active) *in vivo*. Unfortunately, the function of putative deacetylases was not studied *in vitro* due to difficulties in obtaining active enzymes. At present, it is unclear why *Sl*CobB and *Sl*SrtA proteins were inactive in isolation. However, results of *in vivo* experiments show that *Sl*SrtA is is an active participant in the regulation of *Sl*Acs activity, hence offering avenues for future research. Prior to this publication there were no known substrates for *S. lividans* protein deacylases.

MATERIALS AND METHODS

Bacterial strains, culture media, chemicals, and sequencing methods.

All strains used in this study are listed in Table 5.1. *Escherichia coli* C41 (λDE3) (38) and DH5α (New England Biolabs) strains were grown in lysogeny broth (LB, Difco) at 37°C. *Salmonella* strains used for growth analysis were derivatives of *Salmonella enterica* subsp. *enterica* sv Typhimurium strain LT2 (hereafter *S. enterica*) and grown at 37°C in nutrient broth (NB, Difco) with NaCl (85 mM), or no-carbon essential (NCE) minimal medium (39) supplemented with sodium acetate (10 mM), L-methionine (0.5 mM), $MgSO_4$ (1 mM), and trace minerals (40). When in the medium antibiotics were present at the following concentrations: ampicillin, 100 μg mL^{-1}; chloramphenicol, 20 μg mL^{-1}.

Streptomyces strains were all derivatives of *Streptomyces lividans* strain TK24. ISP-2 medium (41) or R2YE medium (42, 43) was used to culture *S. lividans* on solid medium. Liquid

Table 5.1. Bacterial Strains and plasmids used in this study		
Strain	**Relative Genotype**	**Source[1]**
E. coli strains		
E. coli DH5α	Φ80d*lacZ*ΔM15 *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 (r_k-, m_k^+) *supE*44 *relA*1 *deoR* Δ(*lacZYA-argF)* U169 *phoA*	NEB
E. coli C41 (λDE3)	*pka12::kan*$^+$ *ompT hsdS* (r_Bm_B) *gal λ (DE3)*	Laboratory collection
S. enterica strains		
JE6583	*metE205 ara-9*	K. Sanderson via J. Roth
Derivatives of JE6583		
JE22070	pCV1 / pCV3	
JE7758	Δ*acs2*	Laboratory collection
JE21853	Δ*acs2* / pCV1 / pCV3	
JE21760	Δ*acs2* / p*Sl*Acs50 / pCV3	
JE21763	Δ*acs2* / p*Sl*Acs50 / p*Sl*PatA57	
JE21762	Δ*acs2* / p*Sl*Acs50 / p*Sl*PatB4	
JE21761	Δ*acs2* / p*Sl*Acs50 / p*Ma*PatB1	
JE23482	Δ*acs2* Δ*pat3* / p*Sl*Acs50 / pCV3	
JE23483	Δ*acs2* Δ*pat3* / p*Sl*Acs50 / p*Sl*PatA57	
JE23484	Δ*acs2* Δ*pat3* / p*Sl*Acs50 / p*Sl*PatB4	
JE23486	Δ*acs2* Δ*pat3* / p*Sl*Acs50 / p*Sl*PatB7	
JE23485	Δ*acs2* Δ*pat3* / p*Sl*Acs50 / p*Ma*PatB1	
JE23487	Δ*acs2* Δ*pat3* / p*Sl*Acs50 / p*Ma*PatB5	
S. lividans strains		
TK24	Wild type	
Derivatives of *S. lividans* TK24		
JE23392	TK24 / pSlAcs31	
JE23583	ΔEFD68590 / pSlAcs31	
JE23581	ΔEFD68590 ΔEFD71509 / pSlAcs31	
JE23398	ΔEFD68590 ΔEFD71509 ΔEFD65580 / pSlAcs31	
JE23639	ΔEFD70633 / pSlAcs31	
JE23900	TK24 / pSlAcs52	
JE23952	ΔEFD68590 / pSlAcs52	
JE23954	ΔEFD68590 ΔEFD71509 / pSlAcs52	
JE23956	ΔEFD68590 ΔEFD71509 ΔEFD65580 / pSlAc52	
JE23948	ΔEFD70633 / pSlAcs52	

cultures were grown on yeast extract-malt extract (YEME) rich medium in baffled flasks (cultures under 200 mL) or with marine-grade stainless steel springs (NMMP) (cultures over 500 mL) to reduce cell clumping. Strains were grown 2-5 days at 30°C at 220-250 rpm. Antibiotics were used at the following concentrations: thiostrepton, 10 μg mL^{-1}, ampicillin, 100 μg mL^{-1}, chloramphenicol, 20 μg mL^{-1}. All chemicals were purchased from Fisher Chemical Co. unless noted otherwise; chloramphenicol, L(+)-arabinose (Sigma-Aldrich); and isopropyl β-D-1-thiogalactopyranoside (IPTG, IBI Scientific). All restriction enzymes were purchased from Thermo Scientific™ with the exception of BspQI (New England Biolabs).

Molecular techniques.

DNA manipulations were carried out using standard molecular techniques (44). DNA was amplified using Pfu Ultra II Fusion DNA polymerase (Agilent) or Phusion High-Fidelity DNA Polymerase (New England Biolabs). Site-directed mutagenesis was performed using the Quikchange™ Site Directed Mutagenesis protocol (Agilent). Plasmids were purified 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 at the Georgia Genomics Facility. Primers were synthesized from Integrated DNA Technologies and are listed in Table 5.2.

Construction of plasmids for protein production and complementation.

The following genes were amplified from purified genomic DNA using primers listed in Table 5.3: *Streptomyces lividans* TK24 *acs* (EFD68454), *patA* (EFD66247), *patB* (EFD70633), *aacS* (EFD70521), *Micromonospora aurantiaca* ATCC 27029 *acs* (Micau_0428), *patB* (Micau_1670), *Salmonella enterica* LT2 *pat* (STM2651), *acs* (STM4275), *Rhodopseudomonas palustris* CGA009 *badA* (RPA0661), and *pat* (RPA4240). For *Salmonella* complementation

Table 5.2. Plasmids used in this study		
Plasmid	**Genotype**	**Source[1]**
pTEV5	*lacI*$^+$ *bla*$^+$	(45)
pTEV6	*lacI*$^+$ *malE*$^+$ *bla*$^+$	(45)
pTEV18	*lacI*$^+$ *bla*$^+$	(31)
pTYB1	*lacI*$^+$ *bla*$^+$	
pCV1	*araC*$^+$ *bla*$^+$	(31)
pCV3	*araC*$^+$ *cat*$^+$	(31)
pKC1139	Temperature sensitive shuttle vector, *apr*$^+$	(46)
pRK2013	Self-transmissible helper plasmid, *mob*$^+$, *tra*$^+$, *kan*$^+$	(47)
pSE34	PermE constitutive expression vector, *bla*$^+$, *tsr*$^+$	(48)
Overexpression plasmids		
p*Sl*Acs1	*S. lividans acs*$^+$ (EFD68454) in pTEV5, *bla*$^+$	(19)
p*Sl*Acs2	*S. lividans acs*$^+$ (Acs^{K610A}) in pTEV5, *bla*$^+$	
p*Ma*Acs3	*M. aurantiaca acs*$^+$ (Micau_0428) in pTEV18, *bla*$^+$	
pACS67	*S. enterica acs*$^+$ in pTEV18, *bla*$^+$	(31)
pACS28	*S. enterica acs*$^+$ (Acs^{K609A}) in pTYB1 , *bla*$^+$	(8)
pPAT8	*S. enterica pat*$^+$ in pTEV6, *bla*$^+$	(49)
p*Sl*PatA1	*S. lividans patA*$^+$ (EFD66247) in pTEV5, *bla*$^+$	(18)
p*Ma*PatB3	*M. aurantiaca patB*$^+$ (Micau_1670) in pTEV18, *bla*$^+$	
p*Sl*PatB1	*S. lividans patB*$^+$ (EFD70633) in pTEV5, *bla*$^+$	
pCOBB71	*S. enterica cobB$_s$*$^+$ in pTEV6, *bla*$^+$	(50)
p*Rp*Pat13	*R. palustris pat* in pTEV5	(26)
pBadA3	*R. palustris badA* in pTEV5	(26)
Complementation plasmids		
p*Sl*Acs50	*S. lividans acs*$^+$ in pCV1, *bla*$^+$	
p*Ma*PatB1	*M. aurantiaca patB*$^+$ in pCV3, *cat*$^+$	
p*Sl*PatB4	*S. lividans patB*$^+$ in pCV3, *cat*$^+$	
p*Sl*PatA57	*S. lividans patA*$^+$ in pCV3, *cat*$^+$	
p*Sl*PatB7	*S. lividans patB* coding for PatB^{E366Q} in pCV3, *cat*$^+$	
p*Ma*PatB5	*M. aurantiaca patB* coding for PatB^{E269Q} in pCV3, *cat*$^+$	
***Streptomyces* deletion constructs**		
pKC1139EFD68590	Upstream and downstream regions of EFD68590 (*acuC*) in pKC1139	
pKC1139EFD71509	Upstream and downstream regions of EFD71509 (*cobB*) in pKC1139	
pKC1139EFD65580	Upstream and downstream regions of EFD65580 (*srtA*) in pKC1139	
pKC1139EFD70633	Upstream and downstream regions of EFD70633 (*patB*) in pKC1139	
***Streptomyces* complementation**		
p*Sl*Acs31	*S. lividans acs*$^+$ (His$_6$, EFD68454) in pSE34, *tsr*$^+$, *bla*$^+$	
p*Sl*Acs52	*S. lividans acs* coding for His$_6$-Acs^{S608T} (EFD68454) in pSE34, *tsr*$^+$, *bla*$^+$	

Table 5.3. Primers used in this study	
Primer Name	**Primer Sequence[2] 5' → 3'**
Overexpression and complementation primers	
5' *Sl*Acs ATG-His_6 pCV1	NN**GCTCTTC**NTTCATG CATCATCACCATCACCACAGCAAC GAATCCTTGGCCAAC
3' *Sl*Acs pCV1	NN**GCTCTTC**NTTATCAGTCCTCGCTGGGGGCGGCCG
5' *Ma*Acs pCV1 pTEV18	NN**GCTCTTC**NTTCATGAGCGAGGCATTGGCCAACTTGCTG
3' *Ma*Acs pCV1 pTEV18	NN**GCTCTTC**NTTATCAGTCCTCGTCGGACTTCCCGCC
5' *Sl*PatA ATG pCV3	NN**GCTCTTC**NTTCATGTCGTACGCGAGCCGTACTCTG
3' *Sl*PatA pCV3	NN**GCTCTTC**NTTATCAGCCGCCTGGGCGGTGGTGCCCG
5' *Sl*PatB pCV3	NN**GCTCTTC**NTTCATGAGTCGAGACATGTCTGATGTG ACGG
3' *Sl*PatB pCV3	NN**GCTCTTC**NTTATCACTCCCGGCCGATCCGCTCCGCCT
5' *Sl*PatB pTEV5	
3' *Sl*PatB pTEV5	
5' *Ma*PatB pCV3 pTEV18	NN**GCTCTTC**NTTCATGGCGCTCTGGCGGATCCGAGCCA
3' *Ma*PatB pCV3 pTEV18	NN**GCTCTTC**NTTATCAACTCGACCGCTGTGCCGGCACCGT
5' *Sl*Acs pSE34	
3' *Sl*Acs pSE34	
Site directed mutagenesis primers	
5' *Sl*PatB E366Q	CAGCGCGACCTGCGTCTCGTCCC
3' *Sl*PatB E366Q	GGGACGAGACGCAGGTCGCGCTG
5' *Ma*PatB E269Q	CAGCGCCGCCTGGGCCTCGTCGC
3' *Ma*PatB E269Q	GCGACGAGGCCCAGGCGGCGCTG
5' *Sl*Acs S608T	ATGATCTTGCCGGTGCGGGTCTTCGGC
3' *Sl*Acs S608T	ATGATCTTGCCGGTGCGGGTCTTCGGC
S. lividans primers	
EFD70633 5' EcoRV del	CTATGACATGATTACGAATTC**GATATC**GCCCAGCTCGTGCTGCTCGGC
EFD70633 R linker XbaI	GTGACGGATGCCGTGGCGCGC**TCTAGA**GAGCGGATCGGCCGGGAGTGA
EFD70633 F linker XbaI	TCACTCCCGGCCGATCCGCTC**TCTAGA**GCGCGCCACGGCATCCGTCAC
EFD70633 3' HindIII del	TGTAAAACGACGGCCAGTGCC**AAGCTT**GGCCGCACTCCCACTGCTGG
EFD65580 5' EcoRV del	GGC **GAT ATC** CGG CAA CGC CAC GTG CGG CTT C
EFD65580 R linker XbaI	TTC **TCT AGA** CTT GCC GGT CAT GCG GTC GAG
EFD65580 F linker XbaI	**GTC TCT AGA** CTGCTGCGCGGGCTGGGCTGA
EFD65580 3' HindIII del	AAG**AAGCTT**GAACGTCCCTTCGCCGAGACC
EFD68590 5' EcoRV del	GCT**GATATC**TCGGACTTCGACGACGTCTTCC
EFD68590 R linker XbaI	GAC**TCTAGA**CA GGG CGA GCC GGA CCG GGTC
EFD68590 F linker XbaI	TTG**TCTAGA**GTTGCGGGGGTTGTTGGCGTAG
EFD68590 3' HindIII del	CTC**AAGCTT**CTCGATCACGTCGTCCTCGTGC
EFD71509 5' EcoRV del	GCG**GATATC**GCGAAGTACCTGTCCCAGTCGGAG
EFD71509 R linker XbaI	TAC**TCTAGA**CATGCGGCAAGGGTACGGAAC
EFD71509 F linker XbaI	AGG**TCTAGA**GGACGCGGCGACGGCCTGAC
EFD71509 3' HindIII del	GCC**AAGCTT**CCGAGCTTCATCAGTGCTGCTC

[2]Nucleotides shown in bold typeface indicate restriction sites

studies, the start codons of *S. lividans acs* and *patA* genes (GTG, TTG, respectively) were changed to ATG. The DNA fragments and corresponding plasmids were digested with BspQI and cloned using protocols described elsewhere described (31, 51).

The resulting plasmids were: p*Sl*Acs1, p*Sl*PatA1, p*Sl*AacS1, p*Sl*PatB1, p*Ma*Acs3, p*Ma*PatB3, p*Rp*Pat2, p*Rp*BadA3, p*Se*Acs65, and p*Se*Pat8. All plasmids directed the synthesis of proteins fused to an *N*-terminal H_6 tag (with the exception of p*Rp*Pat2, which was fused to an *N*-terminal H_6-MBP tag). All tags were cleavable by recombinant tobacco etch virus (rTEV) protease purified as described (52). Plasmids that directed the synthesis of *Sl*Acs^{K610A} (p*Sl*Acs2), and *Ma*Acs^{K620A} (p*Ma*Acs5) variants were generated from plasmids p*Sl*Acs1 and p*Ma*Acs3. The resulting plasmids (*i.e.,* p*Sl*Acs50, p*Sl*PatB4, p*Sl*PatA57, p*Ma*Acs1 and p*Ma*PatB) were used for complementation in *Salmonella*. For purification from *Streptomyces,* p*Sl*Acs31 or p*Sl*Acs52 directed the synthesis of a protein fused to an *N*-terminal H_6-tag from plasmid pSE34. For identification of plasmid backbones and antibiotic markers for the above-mentioned constructs, refer to Table 5.1.

Construction of gene deletion in *Salmonella enterica*.

An in-frame deletion of *acs* in *Salmonella enterica* was constructed using the phage lambda Red recombinase system as described elsewhere (53). Plasmids were transformed into strains as described (54). In cells that carried two plasmids, pCV1 plasmids were transformed into competent *S. enterica* cells and selected for on ampicillin. Competent cells harboring pCV1 plasmids were then transformed with pCV3 plasmids, whose inheritance was selected by chloramphenicol resistance. Transformants were streaked onto LB + ampicillin and chloramphenicol to select for cells containing both plasmids.

Construction of gene deletions in *Streptomyces lividans*.

In-frame deletions of Δ*ldaA,* Δ*srtA,* and Δ*cobB* were generated as described (55). DNA fragments of 1.5-kb upstream and downstream of the gene of interest were amplified from *S. lividans* TK24 purified genomic DNA. The resulting fragments were cloned into plasmid pKC119 (46) using the In-fusion ND Cloning Kit (Clontech) and transformed into *E. coli* Stellar competent cells (Clontech). Resulting plasmids were transformed into the *E. coli* helper strain HB101 harboring pRK2013 (47). Plasmids were conjugated into *S. lividans* and plated on mannitol soya (MS) agar (42). After 20 h of growth, plates were flooded with apramycin (50 μg ml^{-1} final concentration for 25 ml of plate in 3 mL of soft nutrient agar) to select for the acquisition of deletion construct plasmids. Apramycin-resistant strains were inoculated into 25 ml YEME + apramycin grown in 500 mL baffled flasks at 30°C for 4 days. Cells were pelleted, re-suspended and plated on ISP-2 medium + apramycin at 42°C for 3 d to select for cells that had integrated the plasmids into the chromosome. Apramycin-resistant cells were grown in 25 mL of YEME medium in 500 mL baffled flasks at 30°C for 4 d and plated on ISP-2 at 30°C to promote loss of plasmid. Isolated colonies were screened on ISP-2 and ISP-2 medium containing with apramycin to screen. Apramycin-sensitive strains were screened by PCR for deletion of *ldaA, cobB,* or *srtA.*

Protein purification from *Escherichia coli* or *Salmonella enterica*.

Proteins purified from cells expressing p*Sl*Acs1, p*Sl*Acs2 p*Sl*PatA1, p*Sl*AacS1, p*Rp*BadA3, were obtained from laboratory stocks. Plasmids p*Sl*ACT-Pat1, p*Ma*Acs3, p*Ma*Acs5, p*Ma*PatB3, p*Rp*Pat2, p*Se*Pat8, p*Se*Acs65 were transformed into *E. coli* C41λ(DE3) *pat*::*kan*$^{+}$. The resulting strains were grown overnight in 50-mL cultures in medium containing ampicillin (100 μg ml^{-1}), and were sub-cultured 1:100 into 2 L of lysogeny broth containing ampicillin (100 μg ml^{-1}). Cultures were grown shaking at 30°C to an optical density at 600 nm (OD_{600}) ~0.6, and

protein synthesis was induced by the addition of IPTG (0.5 mM). Upon induction, cultures were grown at 25°C overnight and harvested at 6000 X *g* for 15 min at 4°C in an Avanti J-2 XPI centrifuge with rotor JLA-8.1000 (Beckman Coulter). Cell pellets were stored at -80°C until used. Pellets were thawed and re-suspended in 50 mL of buffer A [(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer [(50 mM, pH 7.0 at 4°C) containing NaCl (500 mM], imidazole (20 mM), lysozyme (1 mg/mL), DNase (1μg/mL), and protease inhibitor phenylmethanesulfonyl fluoride (PMSF, 0.5 mM)]. Cells were lysed on ice by two rounds of sonication (1 min (2s, 50% duty, ea.) using a 550 Sonic Dismembrator (Fischer Scientific) at setting 5. Clarified cell lysates were obtained after centrifugation for 30 min at 4°C at 40,000 X *g* in a Beckman Coulter Avanti J-25I with JA-25.50 rotor followed by filtration of the supernatant through a 0.45 μm filter (Millipore). Samples were applied at 4°C to a pre-equilibrated 1-ml HisPur nickel-nitrilotriacetic acid (Ni-NTA) resin (Thermo Scientific). The column was washed with 10 column volumes (CV) of buffer A, followed by 7 CV of buffer B (HEPES (50 mM, pH 7.0 at 4°C] containing NaCl (500 mM], and imidazole (60 mM). Proteins were eluted with 5 CV of buffer C (HEPES (50 mM, pH 7.0 at 4°C] containing NaCl (500 mM], and imidazole [500 mM]). Proteins were dialyzed at 4°C in decreasing concentrations of NaCl (400 mM, 200 mM, and 150 mM) with a final buffer composition of HEPES (50 mM, pH 7.0 at 4°C), NaCl (150 mM), and glycerol (20% v/v). Proteins concentration was quantified using a NanoDrop with molecular weights and extinction coefficients of each protein as calculated by inputting protein sequences into ExPASy (ProtParam). Proteins were drop-frozen in liquid nitrogen and stored at -80°C.

Purification of *Sl*Acs from *S. enterica* cells was carried out in the same manner, but the column volume of the nickel resin was 100 μL and 2 L of cells were grown in NCE minimal medium supplemented with acetate (10 mM) as carbon and energy source and L-(+)-arabinose

(100 μM) as inducer. Purification of *Sl*Acs from *S. enterica* was performed in biological triplicates and the experiment was carried out twice.

***In vitro* protein acetylation assays.**

Protein acetylation assays were performed as described (7). Briefly, reaction mixtures contained HEPES buffer (50 mM, pH 7.0 at 25°C), *tris*(2-carboxyethyl)phosphine hydrochloride (TCEP; 1 mM), [1-^{14}C]-acetyl-CoA (20 μM), protein substrate (3 μM), and acetyltransferase (1 μM). Reactions (25 μL total volume) were incubated at 37°C for 60 min. Reactions were stopped by the addition of 5 μL of 6x loading dye (60% (v/v) glycerol, Tris-HCl pH 6.8 (0.3 M), EDTA (12 mM), 12% SDS (w/v), 2-mercaptoethanol (0.87 mM), bromophenol blue (0.05%, w/v)) and heated at 100 °C for 10 min. Samples (3 μL) were resolved using SDS-PAGE (56), and visualized by Coomassie Blue R staining (57). Gels were dried, exposed overnight to a MultiPurpose phosphor screen (Packard), and imaged using a Typhoon (GE Healthcare). *In vitro* protein acetylation assays using [1-^{14}C]-acetyl-CoA were performed in triplicate.

***In vitro* acyl-CoA synthetase assays.**

Acyl-CoA synthetases (concentrations indicated in figure legends) were incubated with acetyltransferase (concentration indicated in figure legends) with or without acetyl-CoA (60 μM) for 1 h at 37°C using reaction conditions described above. Activity of acyl-CoA synthetases was quantified using an NADH consumption coupled assay (58). Reactions contained HEPES buffer (50 mM, pH 7.0 at 25 °C), TCEP (1 mM), ATP (2.5 mM), coenzyme A (CoASH, 0.5 mM), $MgCl_2$ (5 mM), phosphoenol pyruvate (3 mM), NADH (0.1 mM), pyruvate kinase (1 U), myokinase (5 U), lactate dehydrogenase (1.5 U), and organic acid substrate (0.2 mM). All reactions were started by addition of 2 μL of above acetylation reactions (final concentration of acyl-CoA synthetase, 60 nM). Changes in NADH absorbance was monitored at 340 nm for 8 min in a 96-well plate format

using a Spectramax Plus UV-visible spectrophotometer (Molecular Devices). Enzyme activities were calculated as described (4). Data presented here were obtained from technical triplicates of one biological replicate, which was repeated in triplicate.

Western blot analysis of His_6-*Sl*Acs purified from *S. enterica*.

Plasmids directing the synthesis of His_6-*Sl*Acs were transformed into a *metE205 ara-9* Δ*acs* Δ*pat* strain (JE11993). Total cell lysates were prepared from 100 ml of *S. enterica* cells grown on minimal medium supplemented with acetate (10 mM), ampicillin, and arabinose (0.1 mM). Cells were pelleted at 6,000 x *g* for 10 min and re-suspended in 10 ml of buffer A (above) with 1 mg ml^{-1} of lysozyme and 1 μg ml^{-1} DNase. Total cell lysates were quantified using a Bradford assay kit (BioRad). From each total cell lysate 7 μg of protein was loaded onto an SDS-PAGE gel. *Sl*Acs purified from *E. coli* was acetylated *in vitro* with *Ma*PatB as described above and loaded on the same SDS-PAGE gel. Three gels were loaded, one for Coomasie Blue R staining and two for technical replicates for transfer. Proteins were resolved by SDS-PAGE at 220 V for 45 min and subsequently transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore) using a Trans-Blot Turbo System (BioRad) using BioRad StandardSD mini gel setting (1.0 A, 25V, 30 min) and transfer buffer (Tris•HCl (25 mM, pH8), glycine (192 mM) and methanol (10% v/v). Membranes were incubated with blocking buffer consisting of condensed milk (12.5% w/v) in 25 ml of PBST (10 mM NaH_2PO_4 (pH 7.2 at 25 °C), NaCl (0.9 % w/v), and Tween-20 (0.1% (v/v). Membranes were probed with polyclonal rabbit α-acetylated lysine antibodies (1:866 dilution into blocking buffer) (Calbiochem) overnight at 4 °C. Membranes were washed thrice with 20 ml of phosphate buffered saline with tween 20 (PBST) and goat α-rabbit immunoglobuilin G conjugated to calf intestinal alkaline phosphatase (Pierce) were diluted (1:5,000) into blocking buffer and applied to membrane for 1 h at 25 °C to detect $Sl\text{Acs}^{Ac}$. Membranes were washed thrice with 20

mL PBST and signal was detected using nitro-blue tetrazolium chloride [~19 mg / 1 ml of 67% dimethylsulfoxide (DMSO)] and 5-bromo-4-chloro-3'-indolylphosphate *p*-toluidine salt (~9 mg / 1 mL of 67% DMSO) (NBT-BCIP).

Protein purification from *Streptomyces lividans.*

Plasmids pSlAcs31 (encoding His_6-*Sl*Acs) or pSlAcs52 (encoding His_6-*Sl*Acs^{S608T}) were introduced into *S. lividans* strains (wild-type, Δ*ldaA,* Δ*ldaA* Δ*cobB, and* Δ*ldaA* Δ*cobB* Δ*srtA*) by polyethyleneglycol (PEG)-assisted protoplast transformation as described (42, 59). Transformed cells were plated on R2YE medium and grown at 30 °C for 20 h. Plates were then flooded with thiostrepton (10 μg ml^{-1}, final concentration for 25 ml plate volume in 3 ml soft agar nutrient broth) and incubated for 1-2 d at 30 °C to select for strains harboring plasmids. Cells containing p*Sl*Acs31 or p*Sl*Acs52 were grown in 25 mL YEME medium + thiostrepton in 500-ml baffled flasks for 4-5 d at 30 °C. Cells were harvested by diluting 1:2 into water and centrifuged at 2000 x *g* for 10 min. Cell pellets were re-suspended in 1-2 ml of YEME medium and 300 μl were plated on multiple MS agar plates. Spores were harvested as described (43) and counted by serial dilutions plated on ISP-2 medium + thiostrepton. Spores (1 x 10^9 / L) were inoculated into 100 mL YEME medium + thiostrepton and grown for 2-3 days at 30 °C. Cells were harvested by diluting cultures 1:2 into water and centrifuging at 2000 x *g* for 15 min. Cells were re-suspended in 30 ml of buffer A (above) with 1 mg ml^{-1} of lysozyme, 1 μg ml^{-1} DNase, and protease cocktail inhibitor (100 μL, Sigma). Cells were lysed by sonication on ice (46 s total with 2 s on, 2 s off) and cell debris was removed by centrifugation at 45,000 x *g* for 30 min. His_6-*Sl*Acs or His_6-*Sl*Acs^{S608T} protein was purified as described above using a 200 μl of a Ni-NTA resin (Pierce). *Sl*Acs-containing fractions were combined and dialyzed at 4 °C for 3 h into HEPES (50 mM, pH 7 at 4°C), NaCl (250 mM), followed by a final dialysis in HEPES (50 mM, pH 7 at 4°C), NaCl (150 mM), glycerol (20%,

v/v). Proteins were drop frozen into liquid-N_2 until use. His_6-*Sl*Acs or His_6-*Sl*Acs^{S608T} activity was assessed using a coupled spectrophotometry assay as described above. The above procedure was performed in triplicate with error bars of activity assays representing technical triplicates of a biological replicate.

ACKNOWLEDGEMENTS

The authors do not have a conflict of interest. This work was supported by HHS NIH grant R01 GM062203 to JCES. We thank Alex Tucker for fruitful discussions and the Proteomics and Mass Spectrometry Core Facility of the University of Georgia for the performance and analysis of LC/MS/MS peptide fingerprinting.

REFERENCES

1. Walsh CT. 2006. Posttranslational Modification of Proteins: Expanding Nature's Inventory. Roberts and Company Publishers, Greenwood Village, CO.
2. Wagner GR, Payne RM. 2013. Widespread and enzyme-independent Nepsilon-acetylation and Nepsilon-succinylation of proteins in the chemical conditions of the mitochondrial matrix. J Biol Chem 288:29036-29045.
3. Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim S, Falck JR, Peng J, Gu W, Zhao Y. 2007. Lysine propionylation and butyrylation are novel post-translational modifications in histones. Mol Cell Proteomics 6:8211-829.
4. Garrity J, Gardner JG, Hawse W, Wolberger C, Escalante-Semerena JC. 2007. *N*-lysine propionylation controls the activity of propionyl-CoA synthetase. J Biol Chem 282:30239-30245.

5. Guan X, Fierke CA. 2011. Understanding protein palmitoylation: Biological significance and enzymology. Sci China Chem 54:1888-1897.

6. Starai VJ, Celic I, Cole RN, Boeke JD, Escalante-Semerena JC. 2002. Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science 298:2390-2392.

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

8. Chan CH, Garrity J, Crosby HA, Escalante-Semerena JC. 2011. In *Salmonella enterica*, the sirtuin-dependent protein acylation/deacylation system (SDPADS) maintains energy homeostasis during growth on low concentrations of acetate. Mol Microbiol 80:168-183.

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

10. Frey PA, Arabshahi A. 1995. Standard free energy change for the hydrolysis of the alpha, beta-phosphoanhydride bridge in ATP. Biochemistry 34:11307-11310.

11. Thauer RK, Jungermann K, Decker K. 1977. Energy conservation in chemotrophic anaerobic bacteria. Bacteriol Rev 41:100-180.

12. Starai VJ, Escalante-Semerena JC. 2004. Acetyl-coenzyme A synthetase (AMP forming). Cell Mol Life Sci 61:2020-2030.

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

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

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

16. You D, Yao LL, Huang D, Escalante-Semerena JC, Ye BC. 2014. Acetyl-CoA synthetase is acetylated on multiple lysine residues by a protein acetyltransferase with single GNAT domain in *Saccharopolyspora erythraea*. J Bacteriol 196:3169-3178.

17. Schwer B, Bunkenborg J, Verdin RO, Andersen JS, Verdin E. 2006. Reversible lysine acetylation controls the activity of the mitochondrial enzyme acetyl-CoA synthetase 2. Proc Natl Acad Sci U S A 103:10224-10229.

18. Tucker AC, Escalante-Semerena JC. 2013. Acetoacetyl-CoA synthetase activity is controlled by a protein acetyltransferase with unique domain organization in *Streptomyces lividans*. Mol Microbiol 87:152-167.

19. Tucker AC, Escalante-Semerena JC. 2014. Determinants within the *C*-terminal domain of *Streptomyces lividans* acetyl-CoA synthetase that block acetylation of its active site lysine i*n vitro* by the protein acetyltransferase (Pat) enzyme. PLoS One 9:e99817.

20. Tucker AC, Taylor KC, Rank KC, Rayment I, Escalante-Semerena JC. 2014. Insights into the specificity of lysine acetyltransferases. J Biol Chem 289:36249-36262.

21. Xu JY, You D, Leng PQ, Ye BC. 2014. Allosteric regulation of a protein acetyltransferase in *Micromonospora aurantiaca* by the amino acids cysteine and arginine. J Biol Chem 289:27034-27045.

22. Hentchel KL, Escalante-Semerena JC. 2015. Acylation of biomolecules in prokaryotes: a widespread strategy for the control of biological function and metabolic stress. Microbiol Mol Biol Rev 79:321-346.

23. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:W320-4.

24. Gardner JG, Escalante-Semerena JC. 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.

25. Gardner JG, Grundy FJ, Henkin TM, Escalante-Semerena JC. 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD^+ involvement in *Bacillus subtilis*. J Bacteriol 188:5460-5468.

26. Crosby HA, Heiniger EK, Harwood CS, Escalante-Semerena JC. 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.

27. Crosby HA, Pelletier DA, Hurst GB, Escalante-Semerena JC. 2012. System-wide studies of *N*-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. J Biol Chem 287:15590-15601.

28. Crosby HA, Rank KC, Rayment I, Escalante-Semerena JC. 2012. Structure-guided expansion of the substrate range of methylmalonyl-CoA synthetase (MatB) of *Rhodopseudomonas palustris*. Appl Environ Microbiol 78:6619-6629.

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

30. Starai VJ, Takahashi H, Boeke JD, Escalante-Semerena JC. 2003. Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in *Salmonella enterica* and *Saccharomyces cerevisiae*. Genetics 163:545-55.

31. VanDrisse CM, Escalante-Semerena JC. 2016. New high-cloning-efficiency vectors for complementation studies and recombinant protein overproduction in Escherichia coli and Salmonella enterica. Plasmid 86:1-6.

32. Mukherjee S, Hao YH, Orth K. 2007. A newly discovered post-translational modification--the acetylation of serine and threonine residues. Trends Biochem Sci 32:210-216.

33. Mittal R, Peak-Chew SY, Sade RS, Vallis Y, McMahon HT. 2010. The acetyltransferase activity of the bacterial toxin YopJ of *Yersinia* is activated by eukaryotic host cell inositol hexakisphosphate. J Biol Chem 285:19927-19934.

34. Birhanu AG, Yimer SA, Holm-Hansen C, Norheim G, Aseffa A, Abebe M, Tonjum T. 2017. *N*epsilon- and *O*-Acetylation in *Mycobacterium tuberculosis* lineage 7 and lineage 4 strains: Proteins involved in bioenergetics, virulence, and antimicrobial resistance are acetylated. J Proteome Res 16:4045-4059.

35. Sauve AA. 2010. Sirtuin chemical mechanisms. Biochim Biophys Acta 1804:1591-1603.

36. Matsushita N, Yonashiro R, Ogata Y, Sugiura A, Nagashima S, Fukuda T, Inatome R, Yanagi S. 2011. Distinct regulation of mitochondrial localization and stability of two human Sirt5 isoforms. Genes Cells 16:190-202.

37. Haigis MC, Mostoslavsky R, Haigis KM, Fahie K, Christodoulou DC, Murphy AJ, Valenzuela DM, Yancopoulos GD, Karow M, Blander G, Wolberger C, Prolla TA, Weindruch R, Alt FW, Guarente L. 2006. SIRT4 inhibits glutamate dehydrogenase and opposes the effects of calorie restriction in pancreatic beta cells. Cell 126:941-954.

38. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289-298.

39. Berkowitz D, Hushon JM, Whitfield HJ, Jr., Roth J, Ames BN. 1968. Procedure for identifying nonsense mutations. J Bacteriol 96:215-220.

40. Balch WE, Wolfe RS. 1976. New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. Appl Environ Microbiol 32:781-791.

41. Shirling EB, Gottlieb D. 1966. Methods for characterization of *Streptomyces* species. Int J Syst Bacteriol 16:313-340.

42. Kieser T, Bibb MJ, Buttner MJ, Chater K, Hopwood DA. 2000. Media, buffers, and suppliers, p 405-420, Practical *Streptomyces* Genetics. John Innes Foundation, Norwich, England.

43. Kieser T, Bibb MJ, Buttner MJ, Chater K, Hopwood DA. 2000. Growth and preservation of *Streptomyces*, p 43-62, Practical *Streptomyces* Genetics. John Innes Foundation, Norwich, England.

44. Elion EA, Marina P, Yu L. 2007. Constructing recombinant DNA molecules by PCR, p Unit 3.17.1-3.17.12. *In* Ausubel FM, R. Brent RE, Kingston DD, Moore JG, Seidman JA,

Smith a, Struhl K (ed), Curr Protoc Mol Biol. Greene Publishing Associates & Wiley Interscience, New York, N.Y.

45. Rocco CJ, Dennison KL, Klenchin VA, Rayment I, Escalante-Semerena JC. 2008. Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*. Plasmid 59:231-237.

46. Bierman M, Logan R, O'Brien K, Seno ET, Rao RN, Schoner BE. 1992. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene 116:43-49.

47. Figurski DH, Helinski DR. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci 76:1648-1652.

48. Yoon YJ, Beck BJ, Kim BS, Kang HY, Reynolds KA, Sherman DH. 2002. Generation of multiple bioactive macrolides by hybrid modular polyketide synthases in *Streptomyces venezuelae*. Chem Biol 9:203-214.

49. Thao S, Escalante-Semerena JC. 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.

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

51. Galloway NR, Toutkoushian H, Nune M, Bose N, Momany C. 2013. Rapid cloning for protein crystallography using Type IIS restriction enzymes. Crystal Growth & Design 13:2833-2839.

52. Blommel PG, Becker KJ, Duvnjak P, Fox BG. 2007. Enhanced bacterial protein expression during auto-induction obtained by alteration of *lac* repressor dosage and medium composition. Biotechnol Prog 23:585-598.

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

54. VanDrisse CM, Parks AR, Escalante-Semerena JC. 2017. A toxin involved in *Salmonella* persistence regulates its activity by acetylating its cognate antitoxin, a modification reversed by CobB sirtuin deacetylase. MBio 8.

55. Martinez A, Kolvek SJ, Yip CL, Hopke J, Brown KA, MacNeil IA, Osburne MS. 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.

56. Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.

57. Sasse J. 1991. Detection of proteins, p 10.6.1-10.6.8. *In* Ausubel FA, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (ed), Curr Protoc Mol Biol, vol 1. Wiley Interscience, New York.

58. Reger AS, Carney JM, Gulick AM. 2007. Biochemical and crystallographic analysis of substrate binding and conformational changes in acetyl-CoA synthetase. Biochemistry 46:6536-6546.

59. Kieser T, Bibb MJ, Buttner MJ, Chater K, Hopwood DA. 2000. Introduction of DNA into *Streptomyces*, p 229-252, Practical *Streptomyces* Genetics. John Innes Foundation, Norwich, England.

CHAPTER 6

ACETAMIDOBENZOATES CONTROL THE DEGRADATION OF BENZOATE AND PHOTOSYNTHESIS IN *RHODOPSEUDOMONAS PALUSTRIS* [1]

[1]VanDrisse C.M. and Escalante-Semerena J.C. 2018.

Submitted to *m*Bio, July 2018

ABSTRACT

The degradation of lignin-derived aromatics such as benzoate has been extensively studied in *Rhodopseudomonas palustris* and the chemistry underpinning the conversion of benzoate to acetyl-CoA is well understood. Here we characterize the last unknown gene*, badL,* of the *bad* (benzoic acid degradation) cluster. BadL function is required for benzoate degradation under photosynthetic conditions (*i.e.*, light plus anoxia). On the basis of bioinformatics, *in vivo* and *in vitro* data we show that BadL, a Gcn5-related *N*-acetyltransferase (GNAT, PF00583), acetylates aminobenzoates to yield acetamidobenzoates. The latter relieved repression of the *badDEFGAB* operon by binding to BadM, triggering the synthesis of enzymes that activate and dearomatize the benzene ring. We uncovered an unexpected connection between acetamidobenzoates and the expression of genes encoding the photosynthetic reaction center light harvesting complexes through a BadM-independent mechanism. The effect of acetamidobenzoates on photosynthesis is new and different than their effect on the catabolism of benzoate.

INTRODUCTION

Lignin is the second most abundant polymer in Nature, second only to cellulose. Unlike cellulose, lignin does not contain carbohydrate monomers, instead, it is comprised of phenyl derivatives (*e.g.,* coumaryl alcohol, syryngyl alcohol, and coniferyl alcohol). Lignin is found in cell walls of plants, and it is estimated to represent approximately 25% of the terrestrial biomass (1). Aromatic compounds released from lignin are rich in energy and carbon, hence it is not surprising that microbes that occupy environments rich in plant materials have evolved metabolic strategies for the degradation of such compounds. The purple non-sulfur photosynthetic α-proteobacterium *Rhodopseudomonas palustris* is an aquatic bacterium that can degrade aromatic compounds into central metabolites. Shown in Figure 6.1A is the pathway used by *R. palustris* for the degradation of aromatics under anoxic conditions (2). Other aromatic compounds such as chorismate, *p*-coumarate, toluene, vanillate, cresol, and phenol (to name a few) feed into the benzoate catabolism pathway via 4-hydroxybenzoate, benzoate, or benzoyl-CoA (2, 3).

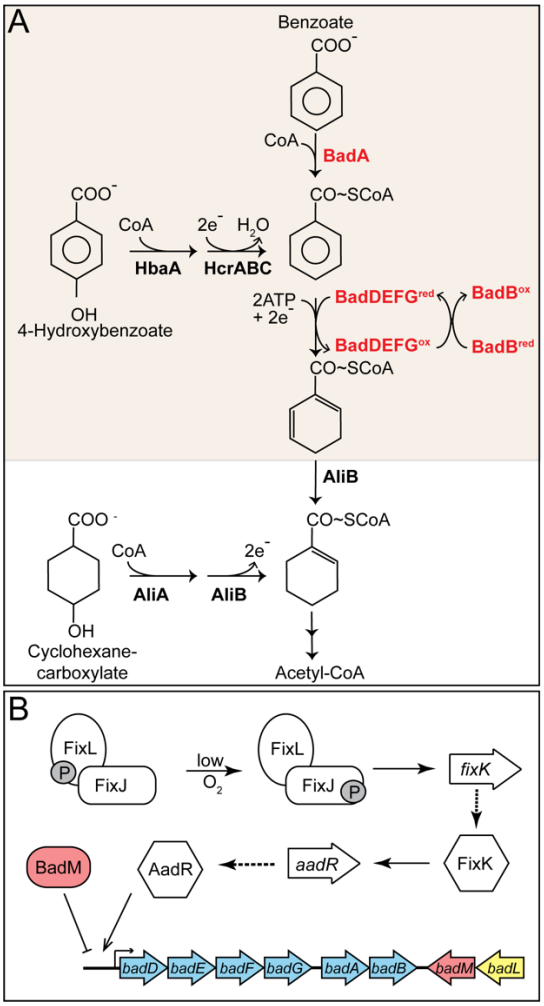


Figure 6.1. Aromatic catabolism and gene regulation in *R. palustris*. A. Benzoate is activated to its CoA thioester by BadA. Hydroxybenzoate can also be converted to benzoyl-CoA in multiple steps by HbaA and HbaBC. The benzene ring of benzoyl-CoA is reduced by BadDEFG yielding cyclohex-1,5-ene-carboxyl-CoA, which is further reduced by BadJ to cyclohex-1-ene-carboxyl-CoA. The latter is the entry point for cyclohexanecarboxylate into the pathway. Cyclohex-1-ene-carboxyl-CoA undergoes a ring reduction, ring cleavage, and several β-oxidations to eventually form acetyl-CoA. Arrows between cyclohex-1-ene-carboxyl-CoA and acetyl-CoA represent a simplified version of enzymatic reactions carried out by BadK, BadH, and BadI. BadA, Benzoyl-CoA synthetase; HbaA, 4-hydroxybenzoyl-CoA synthetase; HcrABC, 4-hydroxybenzoyl-CoA reductase; BadDEFG, benzoyl-CoA reductase; BadB, ferredoxin; AliA, cyclohexancarboxyl-CoA synthetase; AliB, cyclohexancarboxyl-CoA dehydrogenase. Adapted from Harwood 1997 (4). B. *Gene regulation of the* badDEFGAB *operon.* Under low oxygen conditions FixL-P relays its phosphate to FixJ, and FixJ-P activates expression of the global regulator, *fixK*. FixK activates *aadR* transcription and AadR activates expression of the *badDEFGAB* operon (5). BadM is a Rrf2-like regulator that represses the *badDEFGAB* operon (6). Promoter region of BadM has been characterized and lies between *badD* and the upstream gene, *badC* (not shown). FixL, sensor histidine kinase; FixJ, response regulator; FixK, transcriptional regulator; AadR, Crp-like transcriptional activator; BadM, transcription factor.

Benzoate catabolism has been studied in detail in *R. palustris* (3, 7-14). The first step of the pathway is catalyzed by the benzoyl-CoA synthetase (BadA, EC 6.2.1.25) enzyme, which activates benzoate to its CoA thioester, benzoyl-CoA (Fig. 6.1A) (10). The subsequent ring reduction of benzoyl-CoA to cyclohex-1,5-diene-1-carboxyl-CoA is catalyzed by the two $[4Fe\text{-}4S]^{+}$, two $[2Fe\text{-}2S]^{+}$ ATP-dependent reductase BadDEFG enzyme (EC 1.3.7.8, Fig. 1A, (15, 16). The highly oxygen labile iron-sulfur centers of the reductase are re-reduced by the two $[4Fe\text{-}4S]^{+}$ ferredoxin protein, BadB (1). Cyclohex-1,5-diene-1-carboxyl-CoA undergoes a series of reductions, ring cleavage, and β-oxidations, releasing acetyl-CoA and CO_2 (2-4).

Due to the oxygen sensitivity and energetically demanding initial steps of benzoate catabolism (*i.e.* BadDEFG and BadB), the genes encoding these proteins are tightly regulated by activators and repressors (6, 7). A majority of the genes required for benzoate catabolism are clustered within the genome of *R. palustris* (3). Specifically, the genes coding for the benzoyl-CoA synthetase (*badA*), benzoyl-CoA reductase (*badDEFG*), and ferredoxin (*badB*) lie within an operon, *badDEFGAB* (Fig. 6.1B). The *badDEFGAB* operon is activated by AadR (a Crp-type family regulator) and repressed by BadM (a Rrf2-type regulator) (6, 7). The *aadR* gene is additionally regulated via activation by the oxygen sensing two-component Fix (FixL, FixJ, FixK) system (Fig. 6.1B) (5). The Fix-AadR hierarchy mediates the transition from microaerobic to anaerobic growth and further ensures anoxic conditions for the iron-sulfur center proteins of BadDEFG and BadB. To date, it is unknown what signals lead to BadM de-repression of *badDEFGAB.*

Within the aromatic degradation gene cluster, a single gene (*badL*) remains uncharacterized and is annotated as coding for a putative *N*-acetyltransferase. Acetylation occurs in all domains of life and involves the transfer of an acetyl group to α or ε amino groups of proteins (*N*α or *N*ε) and amine groups (*N*α) of small molecules (17). Commonly, protein acetylation occurs on active site lysines of proteins (*N*ε), which in turn modulates their activity (18-25). Small molecule acetylation has been shown to be involved in detoxification (26), translation inhibition (27, 28), and antibiotic neutralization (29, 30).

This study identifies BadL as a small molecule acetyltransferase that modifies aminobenzoates (ABAs). We show that acetylated aminobenzoates (ABA^{Ac}, a.k.a. acetamidobenzoate) bind to BadM and that BadM/ABA^{Ac} complexes can no longer repress *badDEFGAB* operon expression. Our conclusions are supported by results of growth analyses of wild-type and mutant strains, qRT-PCR, and electrophoretic mobility shift assays with BadM and acetamidobenzoates. HPLC and LC/MS/MS data show that BadL acetylates ABAs, and EMSAs show that BadM can bind all three forms of acetamidobenzoates. On the basis of our data we conclude that BadL activity is required for *R. palustris* growth on benzoate and acetylation, and we suggest that BadL works as a sensing mechanism for the presence of ABAs in the environment. Surprisingly, BadL activity affects the synthesis of light harvesting complexes I and II of this bacterium. These results establish BadL as a link between carbon utilization and the generation of a light-driven proton motive force. To our knowledge this is the first example of a Gcn5-type acetyltransferase connecting carbon and energy conservation in a photosynthetic bacterium.

RESULTS

BadL acetyltransferase activity is required for photoheterotrophic growth on benzoate.

Prior to this work, the function of the putative BadL protein was unknown. We began investigating the function of this protein by deleting the *badL* gene and screening for phenotypes related to benzoate utilization. The growth behavior of the *R. palustris badL* strain was assessed under photoheterotrophic conditions. A *R. palustris* Δ*badL* strain failed to grow photosynthetically on benzoate (Fig. 6.2A) or 4-hydroxybenzoate (Fig. 6.2B), compared to $badL^+$ controls. Notably, the Δ*badL* strain grew as well as the $badL^+$ strain on cyclohexanecarboxylate (Fig. 6.2C), suggesting that BadL function was necessary only for substrate activation and ring reduction (Fig. 6.1A, shaded area). The enzymes catalyzing the early steps of the pathway include the benzoyl-CoA synthetase (BadA, EC 6.2.1.25), a benzoyl-CoA reductase (BadDEFG (dearomatizing), EC 1.3.7.8), and ferredoxin (BadB). The above-mentioned genes comprise the *badDEFGAB* operon of *R. palustris*, and are tightly regulated by activation and repression (Fig. 6.1B). On the basis of the results shown in Fig. 6.2A-C we hypothesized that BadL played an as-yet-unidentified role

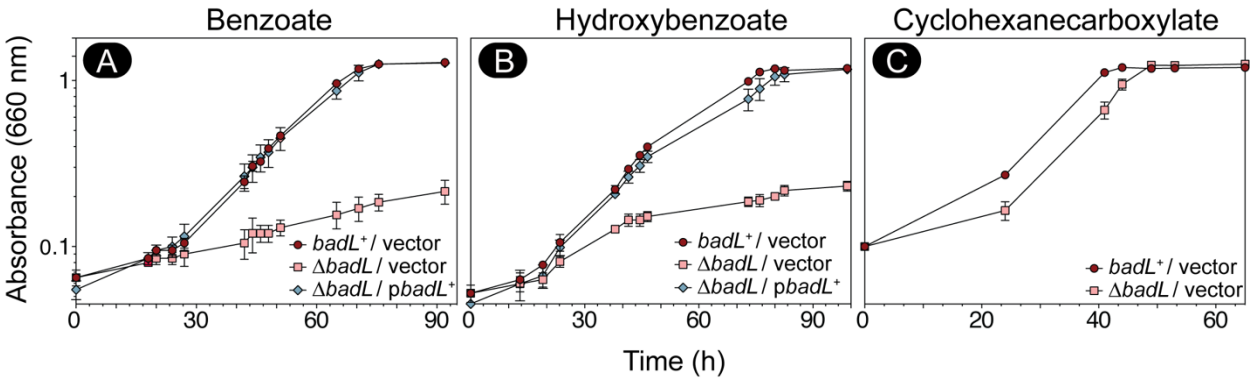


Figure 6.2. ***The BadL acetyltransferase is required for photoheterotrophic growth on benzoate and hydroxybenzoate***. Cells were grown in biological triplicate photoheterotrophically on various carbon sources (3 mM), as listed above each of the panels and growth was monitored at OD_{660}. Cells lacking BadL had lower growth rates compared to *badL*$^+$ controls when grown with benzoate (panel A, red squares *vs* black circles) or hydroxybenzoate (panel B, red squares *vs* black circles), but not when grown with cyclohexancarboxylate (panel C). Growth rates were restored when *badL* was re-introduced on a plasmid (blue diamonds). Each experiment was repeated in triplicate with a representative growth curve shown.

in the regulation of *badDEFGAB* expression, and that its putative function was to either acylate a protein or a small molecule.

Deletion of *badM* restores photoheterotrophic growth of a Δ*badL* strain on benzoate.

To address the possible involvement of BadL in *badDEFGAB* expression, we deleted *badM* in a Δ*badL* strain. In the absence of BadL, growth of *R. palustris* was restored when *badM* was deleted (Fig. 6.3A). A Δ*badL* Δ*badM* strain carrying a plasmid encoding the wildtype allele of *badM* failed to grow photoheterotrophically on benzoate (Fig. 6.3A), suggesting that somehow BadL function affected BadM DNA-binding activity.

Importantly, overexpression of *aadR*, the activator of *badDEFGAB*, did not restore growth of the Δ*badL* strain on benzoate (Fig. 6.3B), leading us to hypothesize that BadL either directly acetylated BadM, or alternatively, BadL acetylated a small molecule effector of BadM. We note, that prior to this work small molecule effectors of BadM DNA-binding activity were not known (6).

We investigated the potential role of BadL in *badDEFGAB* expression. For this purpose, we used qRT-PCR to monitor *badDEFGAB* operon expression using primers and conditions described elsewhere (6). When *R. palustris* was grown photoheterotrophically on succinate supplemented with benzoate, cells lacking *badL* had 30-fold less *badDEFGAB* transcript compared to *badL*$^{+}$ strains (Fig. 6.4A). Additionally, deletion of *badM* in a Δ*badL* strain restored transcription of *badDEFGAB* under photoheterotrophic conditions with succinate + benzoate (Fig. 6.4B). These results supported the idea that BadL was somehow involved in the regulation of *badDEFGAB* expression, but it was unclear how.

BadL homologues acetylate aminobenzoates, and the resulting acetamidobenzoates bind to BadM de-repressing *badDEFGAB* expression.

Attempts to isolate active *R. palustris* BadL protein were unsuccessful among a range of buffers, pHs, ionic strengths, or viscosities. We used SEED viewer version 2.0 (31), to identify BadL homologues that clustered with BadM repressors. We focused our attention on BadL homologues present in *Magnetosprillium magneticum* (*Mm*BadL; locus tag amb3392, 41% identity to *Rp*BadL) and *Geobacter*

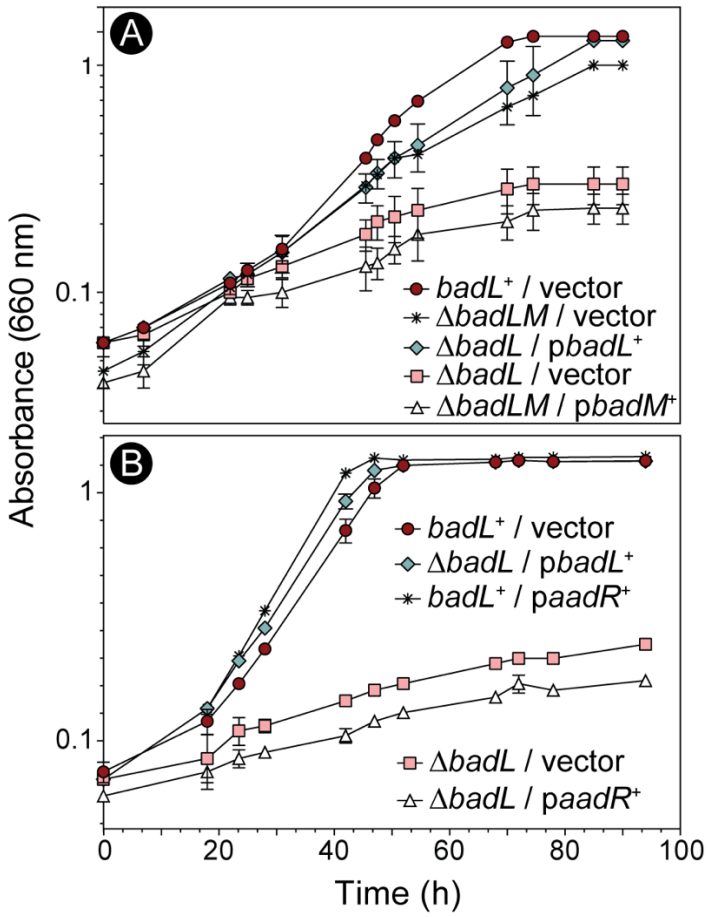


Figure 6.3. BadL dependent phenotypes are restored upon deletion of *badM*, but not overexpression of *aadR*. Cells were grown in biological triplicate photoheterotrophically on benzoate (3 mM) and growth was monitored at OD_{660}. A. A *badM* Δ*badL* strain grew similarily to the *badL*$^+$ strain (black circles *vs* asterisks) and ectopic expression of *badM*$^+$ in the Δ*badM* Δ*badL* strain restored repression, hence lack of growth (triangles). B. Overexpression of *aadR* in a Δ*badL* strain did not recover growth (white triangles). Each experiment was performed in triplicate with a representative growth curve shown.

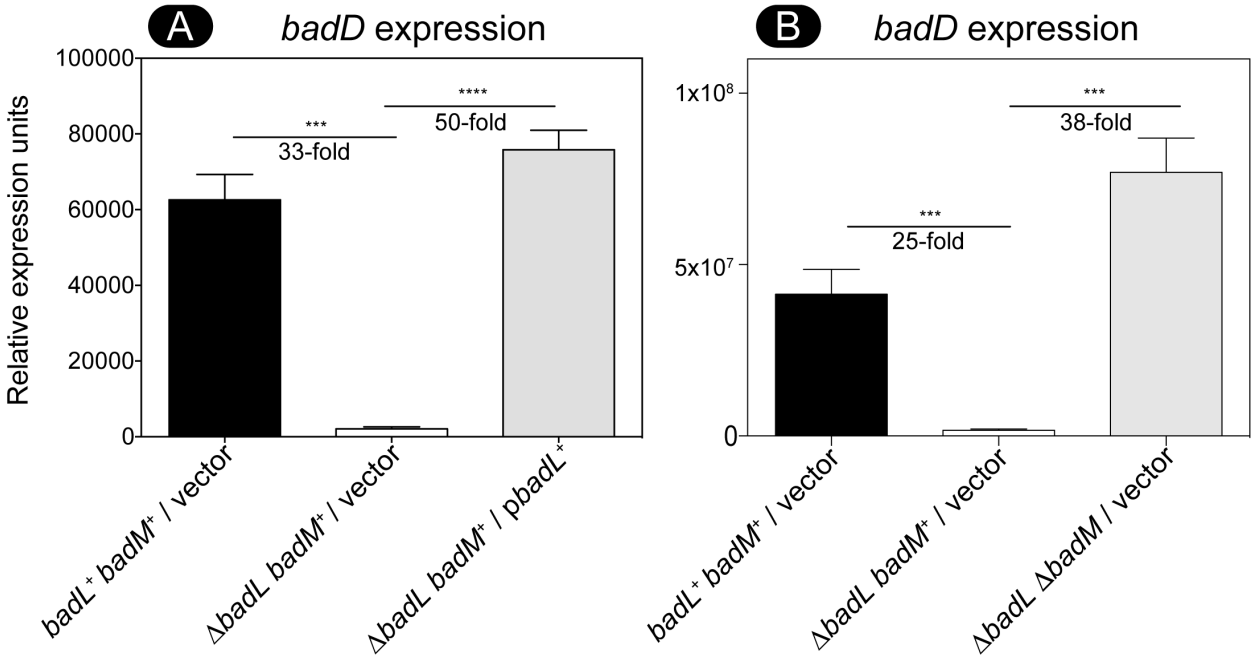


Figure 6.4. In a *badL* strain, *badDEFGAB* operon expression is reduced. Total RNA was obtained from cells grown photosynthetically with succinate and benzoate. A. Expression of the *badDEFGAB* operon was assessed in *badL*$^+$ and Δ*badL* strains using qRT-PCR. Absence of *badL* led to a 30-fold decrease in *badD* expression. When *badL* was reintroduced on a plasmid, transcription increased 50-fold. B. Expression of *badDEFGAB* was compared in *badM*$^+$ or Δ*badM* strains. Cells lacking *badL* had 25-fold lower transcription of *badDEFGAB*. When *badM* was also deleted in cells lacking *badL*, *badDEFGAB* transcription was restored. Transcripts of *badD* were normalized to the housekeeping gene *fixJ*. Cells were grown in biological triplicate and qRT-PCR was performed in technical triplicate of each biological replicate. Error bars represent the standard deviation of the mean. A *p* value of > 0.0005 is represented by *** and *a p* value of less than 0.0001 is represented by ****.

metallireducens (*Gm*BadL; locus tag gmet_2096, 47% identity to *Rp*BadL) (Fig. 6.5). *Mm*BadL and *Gm*BadL were successfully isolated and used in subsequent *in vitro* analyses. Attempts to acetylate BadM with *Mm*BadL or *Gm*BadL did not yield $BadM^{Ac}$ under the conditions tested (data not shown).

To test for small molecule acetylation by BadL, [1-^{14}C]-Ac-CoA was incubated with potential small molecule substrates and either *Mm*BadL or *Gm*BadL. The transfer of radiolabeled acetyl moieties was assessed by thin-layer chromatography using a mobile phase that resolved substrates and acetyl-CoA from products, followed by phosphor imaging. We decided to test small molecules added to *Rhodopseudomonas* growth medium that contained primary amine groups that could be targeted for acetylation. One such molecule, 4-aminobenzoate is added to photosynthetic medium as a precursor for folate synthesis, as the genome of *R. palustris* does not code for enzymes to synthesize its own 4-aminobenzoate. Using this approach, we found that *Mm*BadL and *Gm*BadL acetylated 4-aminobenzoate producing 4-acetamidobenzoate (Fig. 6.6A), We further tested the specificity of BadL by attempting to acetylatate other acetamidobenzoates and found that BadL acetylated 2-, 3-, and 4-aminobenzoate, yielding 2-, 3-, and 4-acetamidobenzoate, albeit to different extents. These results suggested that BadL acetylated the amino group of aminobenzoates. In contrast, *Mm*BadL and *Gm*BadL did not acetylate benzoate, benzoates with hydroxyl groups substituents at the same positions, or acetamidobenzoates (Fig. 6.6B). The proposed reactions and their products are shown in Figure 6.6C. When analyzing unreacted [1-^{14}C]-Ac-CoA at the point of sample application on the TLC plate, it appeared as if *Mm*BadL had a preference for 3-ABA *in vitro* (Fig. 6.6A). We attempted to establish substrate preference by kinetic means, but unfortunately, the level of activity of *Mm*BadL and *Gm*BadL was lower than background limits of the continuous spectrophotometric assay that utilizes Ellman's reagent to quantify sulfhydryl groups (from CoA release), thus we could not evaluate kinetic parameters to determine a preferred substrate.

Mass spectrometry analysis of the product of the *Mm*BadL reaction

Since it appeared that *Mm*BadL acetylated 3-ABA more efficiently and $3\text{-}ABA^{Ac}$ (retention time = 5.0 min) was readily resolved away from 3-ABA (retention time = 2.8 min) by reverse-phase HPLC (Fig. 6.7A), the products of reaction mixtures containing *Mm*BadL and 3-ABA were analyzed by mass

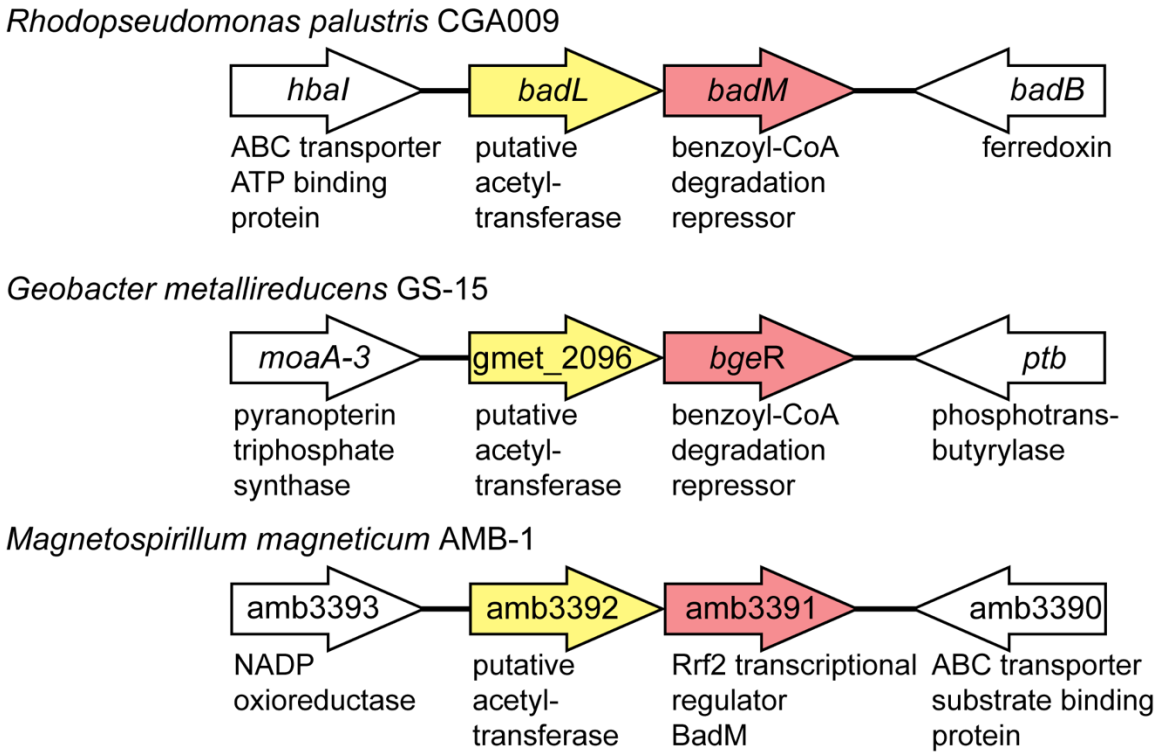


Figure 6.5. Bioinformatics analysis reveal putative *badL badM* gene clustering in other proteobacteria. By analyzing the *R. palustris badL* and *badM* features of The SEED viewer (pubseed.theseed.org/), other organisms were found to have similar clustering of Rrf2-type transcriptional regulators in putative operons with putative acetyltransferases. The putative BadL acetyltransferases from *Geobacter metallreducens* and *Magnetospirillum magneticum* were used for subsequent *in vitro* experiments. Genes surrounding the upstream and downstream *badL badM* regions for each organism are shown and genes are not to scale.

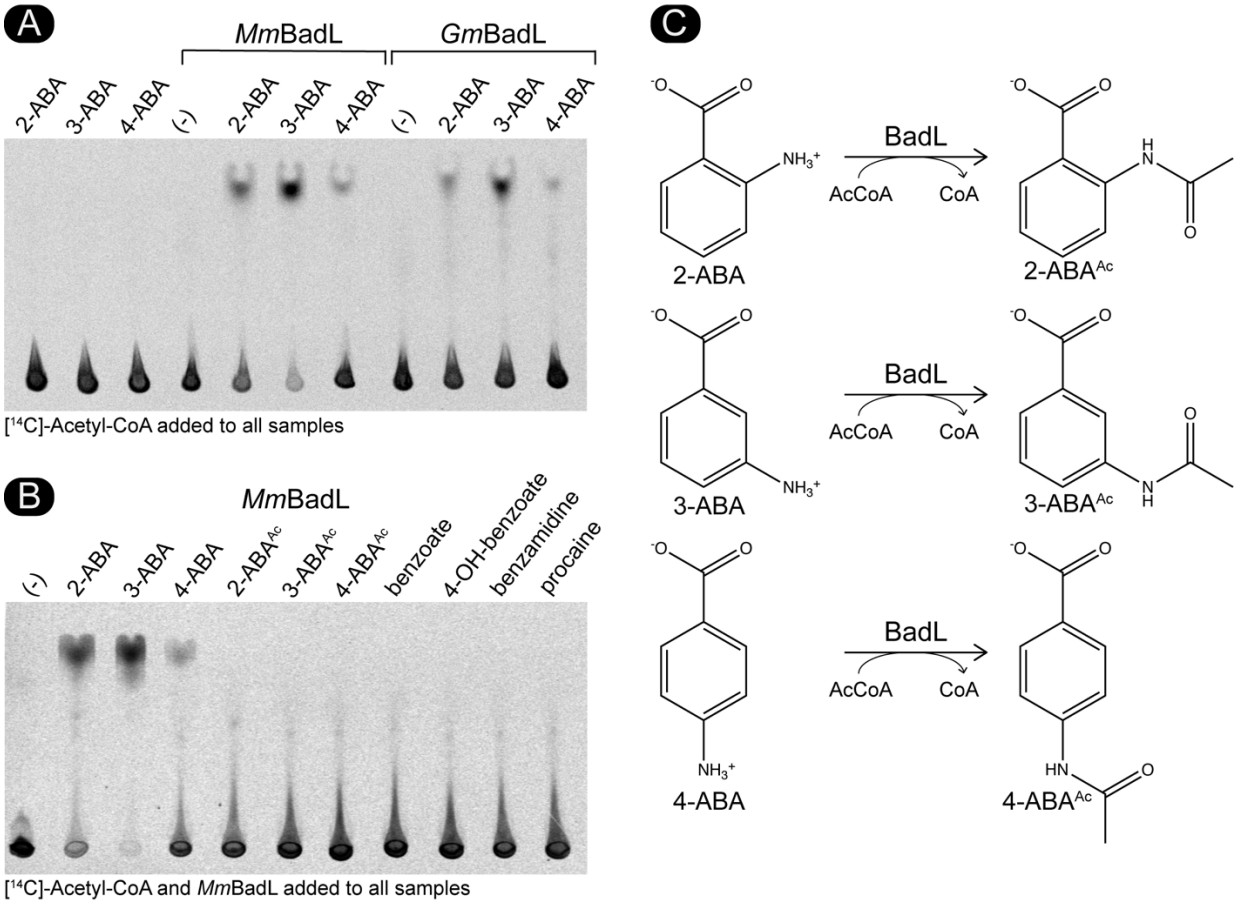


Figure 6.6. BadL homologues acetylate aminobenzoates *in vitro*. A. *Mm*BadL and *Gm*BadL (3 μg) were incubated with [1-^{14}C]-Acetyl-CoA and 2-aminobenzoate (2-ABA), 3-aminobenzoate (3-ABA), or 4-aminobenzoate (4-ABA). As a negative control, 2-ABA, 3-ABA, and 4-ABA were incubated with [1-^{14}C]-Ac-CoA (lanes 1-3). Lanes designated by (-) indicate reaction mixtures that contained BadL, [1-^{14}C]-Ac-CoA, but lacked substrate. Reactions were spotted and acetylated products migrated away from the origin where [1-^{14}C]-Ac-CoA remained. B. *Mm*BadL was incubated with various aromatic substrates as listed above each lane. Lanes labeled as 2-ABA^{Ac}, 3-ABA^{Ac}, and 4-ABA^{Ac} indicate reaction mixtures containing authentic acetylated 2-ABA, 3-ABA, and 4-ABA, as described in materials and methods. Lane designated by (-) identify a reaction mixture where BadL was incubated with [^{14}C-1]-Ac-CoA but no substrate. Reactions were spotted and acetylated products migrated away from the origin, where [1-^{14}C]-Ac-CoA remained. Images were acquired by exposure to a phosphor screen and subsequent imaging. C. Proposed reaction schematic of BadL-mediated aminobenzoate acetylation.

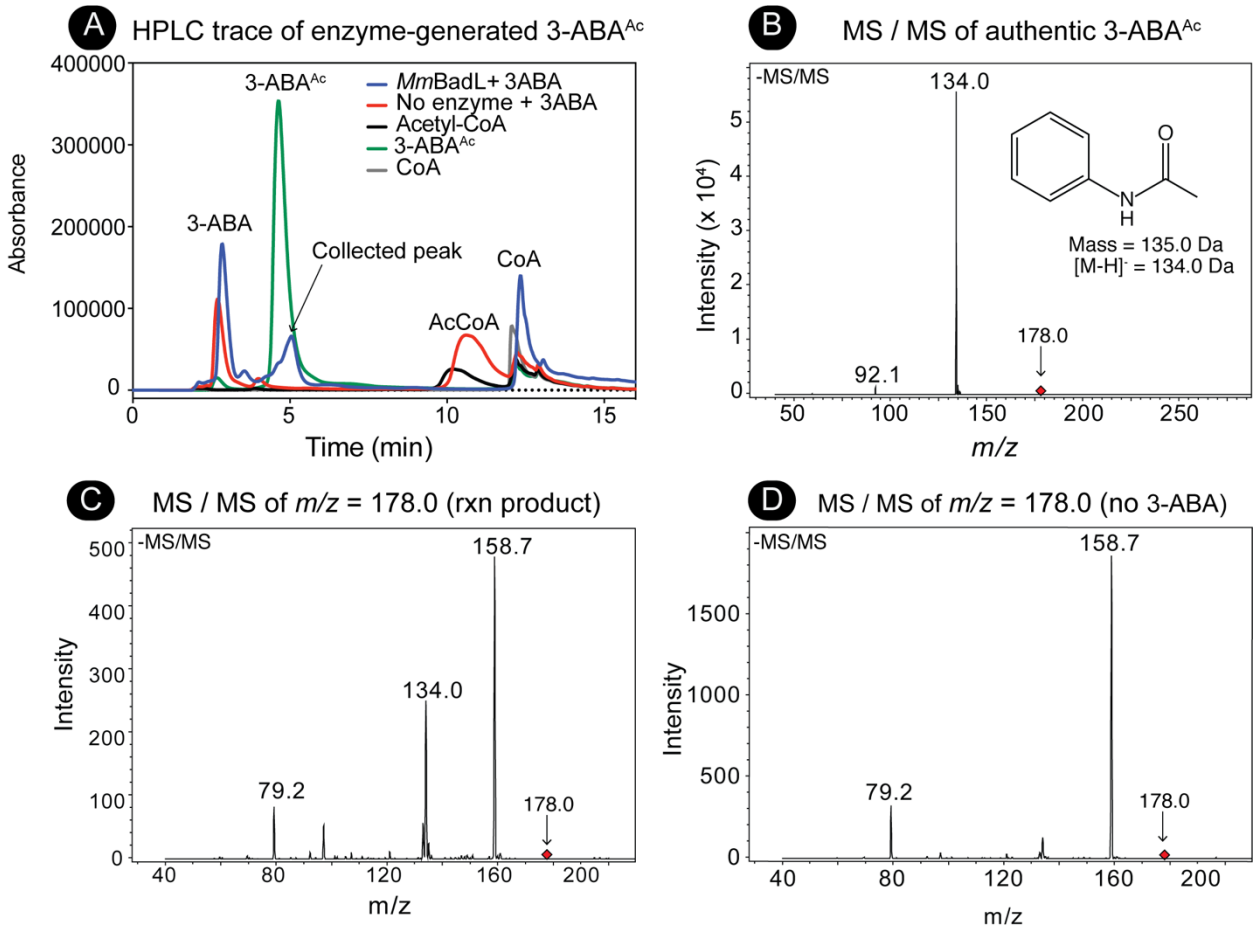


Figure 6.7. Isolation and identification of the *Mm*BadL reaction product by HPLC and mass spectrometry. A. Reaction mixtures containing *Mm*BadL, 3-ABA, and Ac-CoA were set up as described in materials and methods. Protein was removed from mixtures and material elution from a C-18 kinetex column was monitored at 254 nm. Reaction mixtures containing *Mm*BadL were compared to known standards (3-acetamidobenzoate, green peak) and arrow indicated fractions that were collected for MS/MS. B. Mass spectrometry was performed with commercially available, 3-acetamidobenzoate (1 mM) which was identified by the signal with a *m/z* of 178.0 Da (red diamond) in the negative ion spectrum shown. The 178.0 Da compound was fragmented (negative ion shown). The inset shows the structure of the compound with a mass of 134.0 Da, which was consistent with 3-acetamidobenzoate without the carboxylic acid. C. Mass spectrometry was performed on the fraction collected in panel A. The compound with a *m/z of* 178.0 Da (red diamond) corresponding to 3-acetomidobenzoate was fragmented (negative ion spectrum shown). The peak at 134.0 Da corresponds to the structure in panel B inset. Signals observed at *m/z* 158.7 Da and 79.2 Da were determined to be background from machine, as indicated in D.

spectrometry (MS and MS/MS). A signal corresponding to the mass of enzyme-generated 3-ABA^{Ac} (*m/z* = 178.0 Da) was fragmented (Fig. 6.7C) and compared to the authentic standard (Fig. 6.7B). *Mm*BadL-generated and commercially available 3-ABA^{Ac} retained the same fragmentation pattern (*m/z* = 134.0), a fragment that corresponded to the loss of the carboxylate moiety (Fig. 6.7B inset). These results confirmed the identity of the *Mm*BadL product as 3-ABA^{Ac}. A blank of the machine (solvent without HPLC product) was acquired and the signals with *m/z* = 79.2 and 158.7 in Fig. 6.7B were found to be unrelated to the product of the reaction (Fig. 6.7D).

Acetamidobenzoates bind to BadM lowering its affinity for the *badDEFGAB* promoter.

The binding of BadM to the *badDEFGAB* promoter has been analyzed by others (6). We synthesized the *badDEFGAB* promoter region (212 nucleotides, -191 to +21 of *badD* ATG) with a 6-FAM fluorescent label covalently attached to the 3' end of the probe. Under the conditions tested, BadM bound all of the fluorescently labeled probes at 6-fold molar excess, and this condition was used to test the effects of 2-, 3-, and 4-acetamidobenzoate. When BadM was incubated with 2-ABA, its binding to the *badDEFGAB* promoter probe did not differ from that of BadM incubated with the probe in the absence of ABAs (Fig. 6.8A, lane 2 *vs* 4). However, addition of 2-acetamidobenzoate blocked binding of BadM to its DNA target (Fig. 6.8A, lane 2 *vs* 6). This change was clear when the concentration of 2-acetamidobenzoate in the reaction mixture was 10 mM; at 5 mM 2-acetamidobenzoate, BadM/*badDEFGAB* promoter probe interactions were unaffected (Fig. 6.8B). This information helped us assess the effects of 3- and 4-acetamidobenzoate on BadM/*badDEFGAB* promoter probe interactions. Notably, at the concentrations necessary to affect BadM/*badDEFGAB* promoter probe interactions, the solubility of 3- and 4-acetoamidobenzoate was substantially reduced relative to that of 2-acetoamidobenzoate (Fig. 6.9A). When incubated under the same conditions, BadM/*badDEFGAB* promoter probe interactions changed upon incubation with 3- and 4-acetoamidobenzoate, albeit to a lesser degree than with 2-acetoamidobenzoate (Fig. 6.9B, see free probe band percentages). The alluded solubility issues with 3-acetoamidobenzoate and 4-acetoamidobenzoate prevented the quantification of ligand effects of BadM's DNA binding activity. It is possible that all three acetamidobenzoates may serve as ligands for BadM. In combination with qRT-PCR

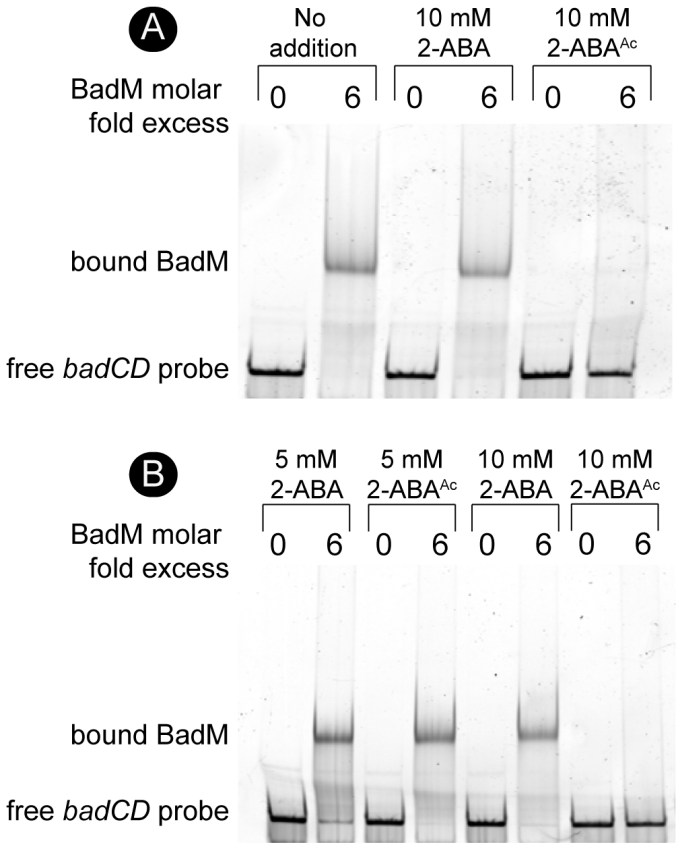


Figure 6.8. BadM DNA binding to the *badDEFGAB* promoter is reduced in the presence of acetylated 2-aminobenzoate. A. Binding of BadM (72 nM) to the *badDEFGAB* promoter (12 nM) was analyzed by electrophoretic mobility shift assays using 6-FAM 5'-labeled probes. A. Probe containing the intergenic region of *badCD* (212 bp, 0.303 pmol) was incubated without or with (1.8 pmol) BadM. Additions of 2-aminobenzoate or 2-acetamidobenzoate are indicated above respective lanes. B. Binding of BadM to the *badCD* intergenic region as in panel A, but with different concentrations of 2-aminobenzoate or 2-acetamidobenzoate.

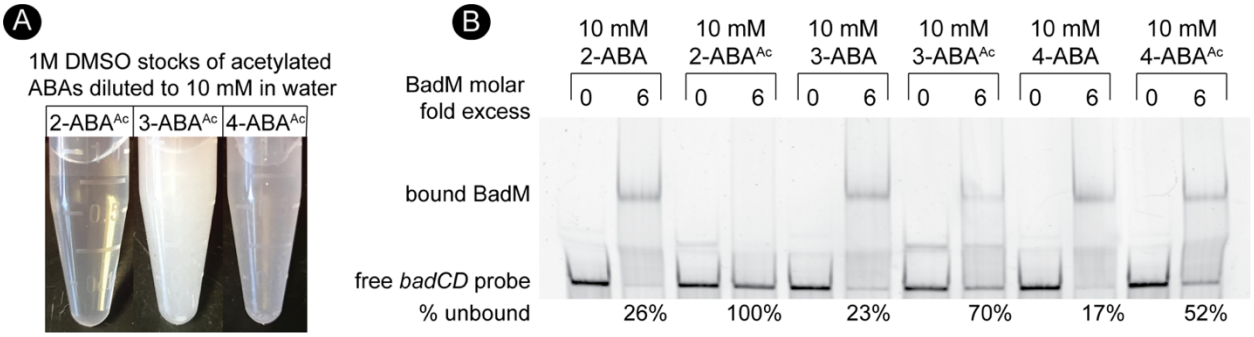


Figure 6.9. Binding of BadM to DNA is affected by all acetamidobenzoates. A. To meet reaction volumes, 2-, 3-, and 4-acetamidobenzoate were resuspended in 100% DMSO to a final concentration of 1M. Pictures show the solubility of each chemical when diluted to 10 mM in water. B. 6-FAM 5'-labeled probe containing the intergenic region of *badCD* (0.303 pmol, 12 nM) was incubated without or with (1.8 pmol, 72 nM) BadM. Various acetamidobenzoates were added to reaction mixtures to the final concentrations indicated above each lane. The percentage of unbound *badCD* probe was analyzed by comparing the intensities of upper and lower bands of samples containing BadM. Intensities were calculated using ImageQuant software and percentages correspond to lanes directly above.

data shown in Figure 6.4, these results indicate that BadL acetylates aminobenzoates, which in turn bind to BadM leading to de-repression of the *badDEFGAB* operon.

Addition of acetamidobenzoates to benzoate medium restores growth of cells devoid of BadL

To assess the substrate specificity of BadM *in vivo*, authentic 2-, 3-, or 4-acetamidobenzoate was added to cultures of a Δ*badL* strain and photoheterotropic growth with benzoate as the carbon source was compared to the response of the same strain to non-acetylated aminobenzoate. When 3- or 4-acetamidobenzoate was added to the medium, the final density of the Δ*badL* cultures was similar to that of the *badL*$^+$ strain, albeit growth rates were different (Fig. 6.10 B and C). Notably, the addition of 2-acetamidobenzoate to the medium did not restore growth of the Δ*badL* strain (Fig. 6.10A), but we cannot rule out transport problems or inhibitory effects. We note that restoration of growth of the Δ*badL* strain upon addition of 4-acetamidobenzoate was not due to alternative carbon source utilization, as *badL*$^+$ and Δ*badL* cells failed to grow when only 4-acetamidobenzoate was added to the medium (Fig. 6.10F). However, when 3-acetamidobenzoate was added to the medium, *badL*$^+$ and Δ*badL* cells grew in the absence of benzoate. Reasons for the differences in growth rates are discussed below. Of note, addition of acetamidobenzoates decreased the lag phase of *badL*$^+$ cells by 20-40% but did not alter the doubling time of these cells (from Fig.6.10).

BadL plays a role in the expression of genes encoding reaction center proteins

During photoheterotrophic growth with benzoate, we noticed a substantial difference in pigmentation between *badL*$^+$ and Δ*badL* strains. To determine whether this observation was due to differences in cell density or pigment synthesis, *R. palustris badL*$^+$ and Δ*badL* strains were grown on benzoate and peak intensities at wavelengths corresponding to the light harvesting 1 (LH1) reaction center complex (880 nm) and the light harvesting 2 (LH2) reaction center complex (808 nm, 863 nm) were analyzed. Absorbance scans ($A_{600\text{-}1000}$) of cultures of *badL*$^+$ and Δ*badL* strains were obtained during lag, log, and stationary growth phases. The *R. palustris* Δ*badL* strain grew at a much slower rate with benzoate compared to the *badL*$^+$ strain (9 *vs* 36 h doubling time, Fig. 6.2), and was therefore ensured that comparisons

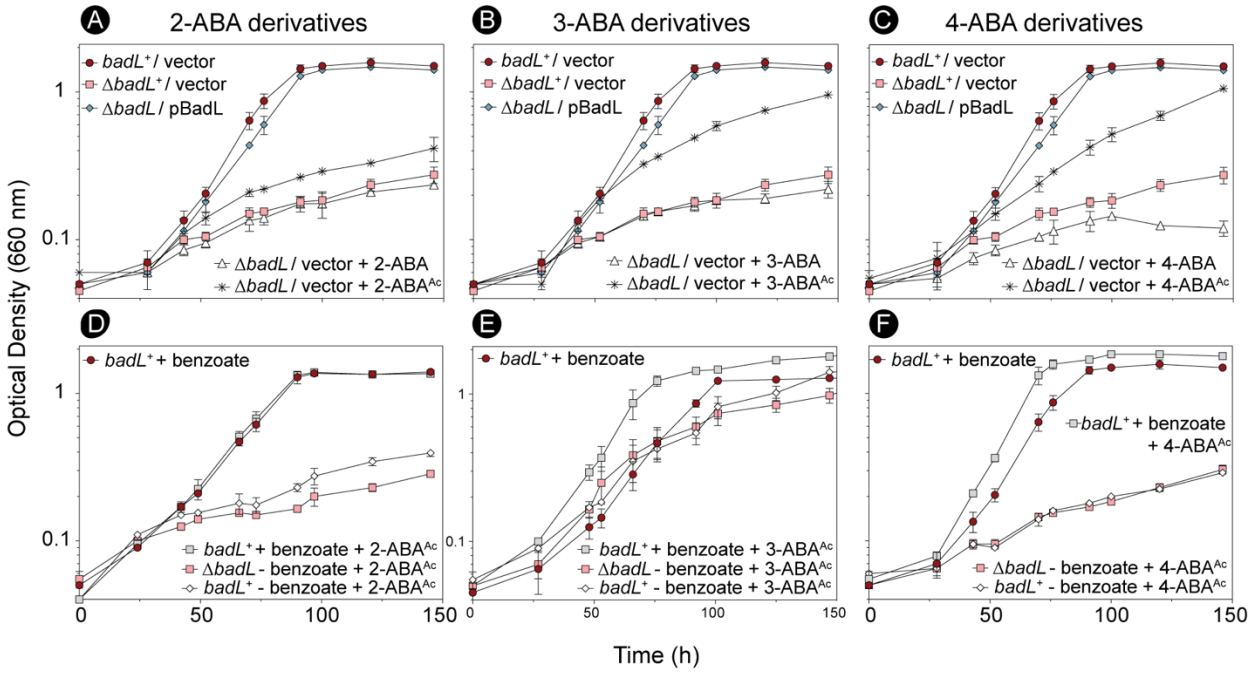


Figure 6.10. Addition of 4-acetamidobenzoate restores photoheterotrophic growth of Δ*badL* cells on benzoate. Cells were grown in biological duplicate photoheterotrophically on benzoate (3 mM) with or without acetamidobenzoates (5 mM) and growth was monitored at 660 nm. Growth analysis of Δ*badL* cells was assessed upon addition of 2-acetamidobenzoate (A), 3-acetamidobenzoate (B), or 4-acetamidobenzoate (C). To determine if growth restoration of Δ*badL* cells was due to alternative carbon source utilization, *badL*$^+$ and Δ*badL* cells were grown on 2-acetamidobenzoate (D), 3-acetamidobenzoate (E), or 4-acetamidobenzoate (F) in the presence or absence of benzoate, as indicated in legends within each panel. Each experiment was performed in triplicate with a representative growth curve shown. Error bars represent SD of a biological triplicate.

between *badL*$^+$ and Δ*badL* were presented at same cell densities. When the pigments of LH1 and LH2 complexes were analyzed, we found that Δ*badL* cells had less pigment in both complexes compared to *badL*$^+$ cells (Fig. 6.11). The lower level of pigment in the Δ*badL* strain was independent of growth phase and was seen in lag, log, and stationary growth phase (Fig. 6.11). This lighter pigmentation was most evident in stationary phase cells, as shown in the inset of Figure 6.11.

We looked further into an unexpected, possible role of BadL in the generation of light-driven proton motive force (pmf) in *R. palustris*. For this purpose, we used qRT-PCR to measure transcript levels of *pucC* and *pufM*. PucC is a probable chlorophyll major facilitator superfamily exporter, while PufM is a photosynthetic reaction center subunit. When cells were grown photoheterotrophically with succinate and benzoate, a strain lacking BadL had significantly lower amounts of *pucC* and *pufM* transcripts compared to the levels measured in the *badL*$^+$ strain (Fig. 6.12A and C). To determine whether or not this result was related to BadM function, *pucC* and *pufM* transcripts were analyzed in a Δ*badL* Δ*badM* strain. In cells lacking *badL*, the presence or absence of *badM* did not change the levels of *pucC* or *pufM* transcripts (Fig. 6.12B and D). In addition to qRT-PCR, spectral analysis of LH1 and LH2 complexes were obtained in *badL*$^+$, Δ*badL*, Δ*badM*, and Δ*badL* Δ*badM* strains. In agreement with the qRT-PCR results, the strain carrying a deletion of *badM* and *badL* did not display increased pigment synthesis relative to the *badL* strain (Fig. 6.13A), suggesting that the reduction in pigmentation in the Δ*badL* strain was unrelated to BadM function. To investigate whether acetamidobenzoates were involved in pigment synthesis or to determine if BadL had different acetylation targets, pigment scans were obtained with *badL* cultures grown photoheterotreophically on benzoate supplemented with 2-, 3-, or 4- acetamidobenzoate. Absorbance readings related to LH1 and LH2 were partially restored to *badL*$^+$ levels only when 4-acetamidobenzoate was added to the medium (Fig. 6.13B). This result would suggest a new role for BadL in energy conservation, or a secondary target for acetamidobenzoates that is directly or indirectly affecting pigment biosynthesis. These possibilities are under investigation.

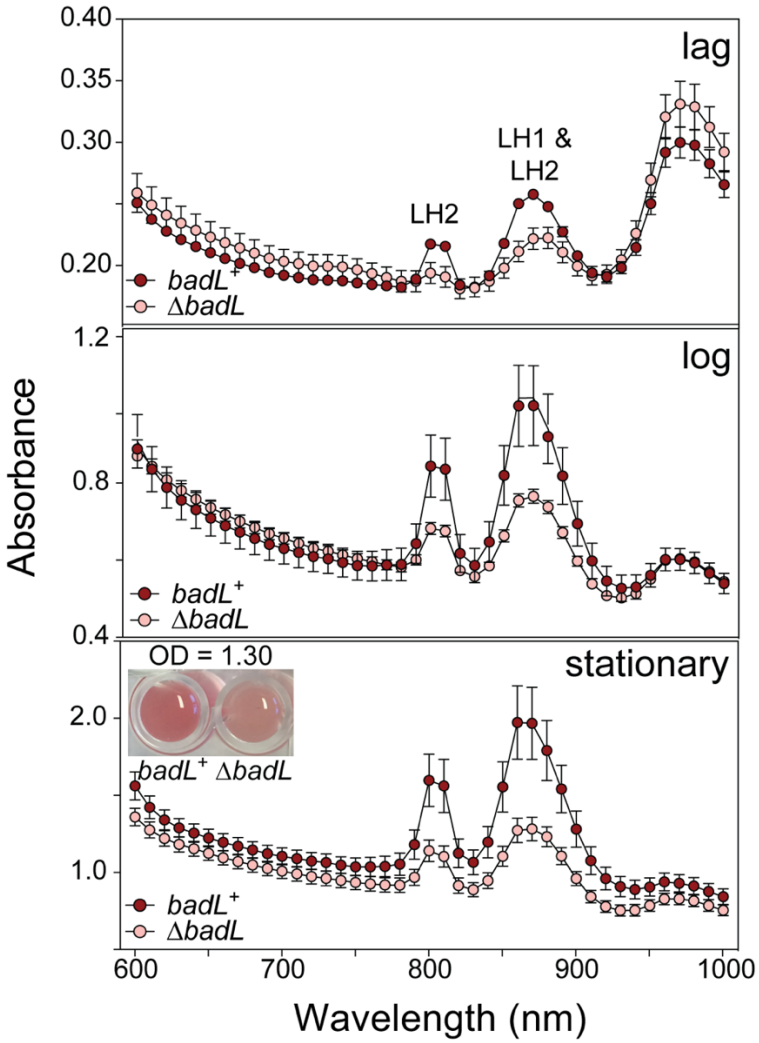


Figure 6.11. Cells lacking BadL have reduced light harvesting reaction center complex pigments. The amount of light harvesting 1 (LH1) and light harvesting 2 (LH2) reaction center complexes in *badL*$^+$ and Δ*badL* strains was monitored spectrophotometrically during lag, log, and stationary growth phases. LH1 complexes absorb at 880 nm and LH2 complexes at 808 and 863 nm. Cells were analyzed at similar cell densities (measured at 600 nm), to ensure that differences in complexes were not due to differences in cell number. Data were obtained in triplicate on a SpectraMax instrument with absorbance readings every 10 nm. For visualization purposes, the inset in the bottom panel shows *badL*$^+$ and Δ*badL* cells at identical stationary phase OD_{660}. Experiment was repeated in biological triplicate and error bars represent SD of a technical triplicate.

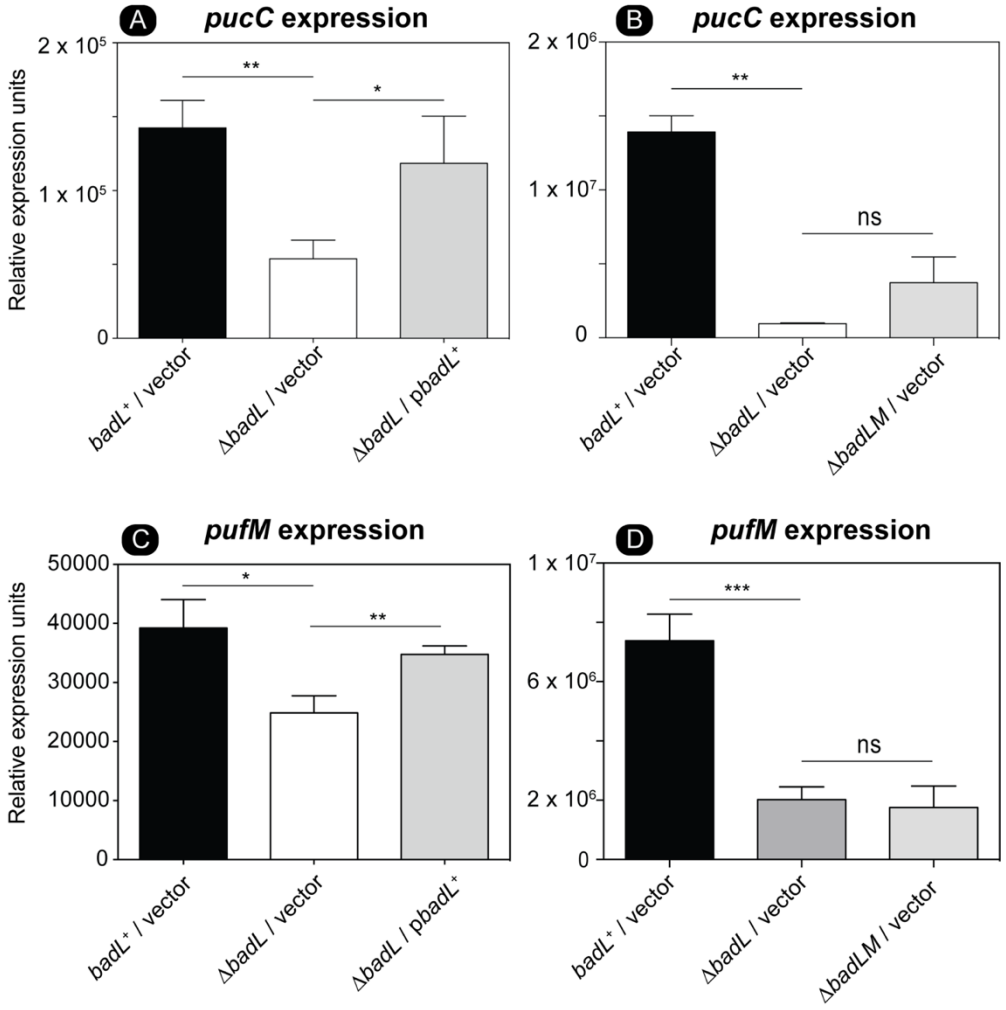


Figure 6.12. Deletion of *badL* lowers the level of *pucC* RNA transcripts, an effect that is independent of the presence or absence of BadM. Total RNA of cells grown photosynthetically with succinate + benzoate was harvested as described in materials and methods. Expression of *pucC* was assessed in *badL*$^{+}$, Δ*badL,* or Δ*badL* Δ*badM* strains using qRT-PCR. A. absence of *badL* lead to a decrease in *pucC* transcription. The gene for *pucC* codes for a probable major facilitator superfamily transporter for photosynthetic complex assembly. Transcription of *pucC* is restored when *badL* was provided on a plasmid. B. *pucC* RNA level was not significantly different in Δ*badL* and Δ*badL* Δ*badM* strains. C. Deletion of *badL* led to a decrease in *pufM* transcription. The *pufM* gene codes for the photosynthetic reaction center complex M. Transcription of *pufM* is restored with *badL* on a plasmid. D. *pufM* RNA level was not significantly different in a Δ*badL vs* Δ*badL* Δ*badM* strain. A *p* value of > 0.005 is represented by ** and *a p* value of less than 0.01 is represented by *.

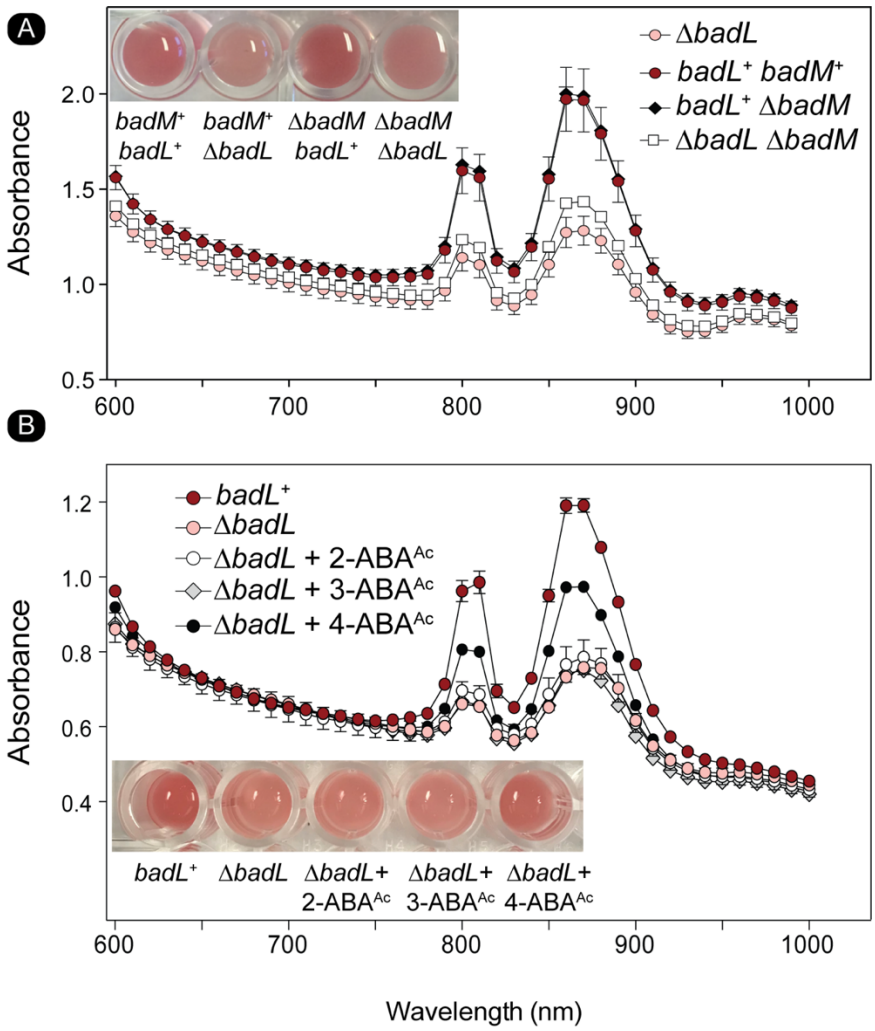


Figure 6.13. The absence of BadM does not restore pigment synthesis in a Δ*badL* strain. The content of light harvesting 1 (LH1) and light harvesting 2 (LH2) reaction center complexes was assessed spectrophotometrically. LH1 complexes absorb at 880 nm and LH2 complexes at 808 and 863 nm. Cells were analyzed at similar cell densities (measured at OD_{660}), to ensure that differences in complexes were not due to differences in cell numbers. Data were obtained in triplicate on a SpectraMax spectrophotometer with absorbance readings every 10 nm. A. LH1 and LH2 complex analysis for *badL*$^+$ and Δ*badL,* Δ*badM,* Δ*badL* Δ*badM* cells. Picture inset represents coloration of cells in a microtiter dish. B. LH1 and LH2 complex analysis for *badL*$^+$ and Δ*badL* cells with either 2-, 3-, or 4-acetamidobenzoate (5 mM).

DISCUSSION

The degradation of lignin-derived aromatics (*e.g.,* benzoate, hydroxybenzoate) under anoxic conditions in the presence of light has been extensively studied in *R. palustris.* The pathway is well defined, and so are the genes encoding the enzymes required to convert such compounds to acetyl-CoA (3). In addition, elegant system wide studies of the regulation of expression of the benzoate acid degradation (*bad*) genes of this bacterium have been reported (6, 7). In spite of this wealth of information, the function of one gene, *badL,* remained enigmatic. In genome databases, *badL* is annotated as encoding a homologue of the yeast Gcn5-type histone *N*-acetyltransferase (GNAT, PF00583).

Here, we present evidence in support of the physiological role of the BadL protein in the degradation of benzoate, and reveal an additional, unexpected role for this protein in the generation of the light-driven proton motive force of this bacterium. Importantly, we show that the role of BadL in photosynthesis is independent of its role in benzoate catabolism.

BadL is required for the expression of the genes encoding enzymes that dearomatize the benzene ring.

As shown by *in vivo* data (Fig. 6.2A, B), *R. palustris* requires BadL function to grow on benzoate as the source of carbon under photoheterotrophic conditions. Notably, BadL function is not required beyond the point of the pathway where intermediates have been dearomatized (Fig. 6.2C). On the basis of bioinformatics information (Fig. 6.5) we propose that the strategy of using an acetyltransferase to affect the function of the repressor is shared by *Geobacter metallireducens, Magnetospirillum magneticum*, and probably many other benzoate degraders. Prior to this work, it was known that AadR, a Crp-like regulator, activated transcription of the *badDEFGAB* operon and that AadR was itself regulated through an oxygen-sensing two component system (5). We have shown that the repression of *badDEFGAB* is relieved through small molecule acetylation and the binding of the acetylated product to BadM.

BadL function generates acetamidobenzoates, which bind to BadM leading to *badDEFGAB* operon de-repression.

The link between BadL function and the expression of genes encoding enzymes responsible for the activation of benzoate (EC 6.2.1.25) and its reduction to cyclohexa-1,5-diene-1-carbonyl-CoA (EC 1.3.7.8) suggested two possible functions for BadL. Either the putative acetyltransferase modified and altered the function of BadA, BadB, BadDEFG or BadM, or it acetylated a small molecule that would signal the presence of benzoate in the cell environment triggering the expression of the *badDEFGAB* genes. *In vivo* and *in vitro* evidence reported herein shows that the latter scenario is correct. We found that BadL acetylates aminobenzoates (2-, 3- or 4-ABA) yielding acetamidobenzoates (also referred throughout the paper as 2-, 3- or 4-ABA^{Ac}) (Fig. 6.6, 6.7), which in turn bind to the BadM repressor decreasing its affinity for its binding site upstream of the *badDEFGAB* operon (Fig. 6.8); this conclusion is supported by qRT-PCR data (Fig. 6.4A). Furthermore, consistent with the idea that acetamidobenzoates act through BadM is the fact that in the absence of BadM, BadL function is irrelevant to benzoate degradation (Fig. 6.4B). A model of this hypothesis can be seen in Fig. 6.14.

Phenotypic analyses validate the assigned functionality of BadL as an aminobenzoate acetyltransferase.

Results of growth behavior analyses of mutant and wildtype strains offer strong support to the conclusion that, in *R. palustris*, acetamidobenzoates generated by BadL are required to relieve the BadM-dependent repression of the *badDEFGAB* operon. In other words, in the absence of BadL, the *badDEFGAB* operon is not expressed unless BadM is absent (Fig. 6.3A, Fig. 6.4). The fact that certain acetamidobenzoates present in the growth medium can bypass the need for BadL (Fig. 6.10) validates the role of BadL in benzoate degradation. Concerning the physiological relevance of each acetamidobenzoate derivative, we cannot say for certain if a single acetamidobenzoate is the true ligand for BadM. All three acetamidobenzoates tested altered the affinity of BadM for the *badDEFGAB* promoter (Fig. 6.8, 6.9), but addition of each acetamidobenzoate to the growth medium of Δ*badL* cultures led to varying growth patterns. For example, 2-acetamidobenzoate did not restore growth of a Δ*badL* strain (Fig. 6.10A), while 3- and 4-acetamidobenzoate did. However, we noted that, unlike 2- and 4-acetamidobenzoate, 3-acetamidobenzoate

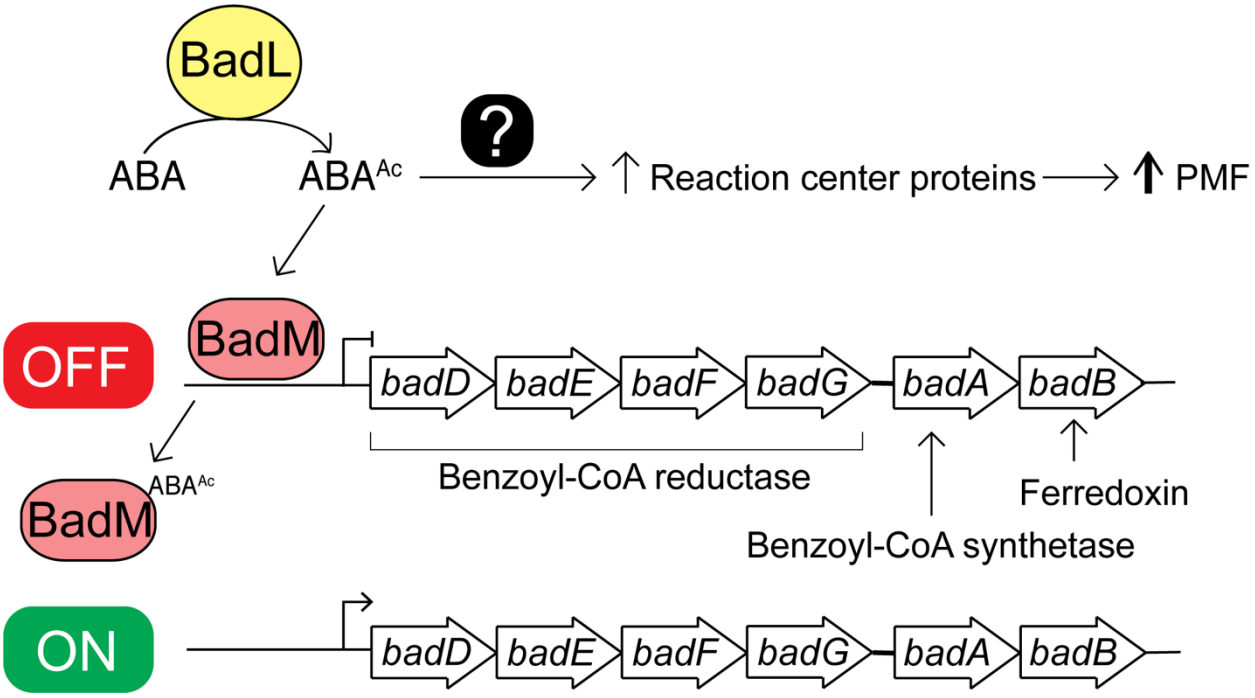


Figure 6.14. A model of the impact of acetamidobenzoates on benzoate degradation and photosynthesis. BadL acetylates aminobenzoates (ABAs) under photoheterotrophic growth. The resulting acetamidobenzoates (ABA^{Ac}) either directly or indirectly increase the presence of reaction center proteins. Additionally, ABA^{Ac} binds to the BadM repressor triggering expression of the *badDEFGAB* operon. The combined activities of benzoyl-CoA synthetase (BadA), benzoyl-CoA reductase (BadDEFG) and ferredoxin (BadB) activate and dearomatize the benzene ring.

was used as a carbon source for *badL*$^{+}$ and Δ*badL* cells in the absence of benzoate (Fig. 6.10 E). To explain this observation, we speculate that an enzymatic cleavage of the acetamide group of 3-acetamidobenzoate releases benzoate, which is subsequently catabolized by Bad enzymes. This explanation would not preclude the de-repression of the *badDEFGAB* operon by 3-acetamidobenzoate. Notably, 4-acetamidobenzoate was not utilized as a carbon source (Fig. 6.10F), allowed the Δ*badL* strain to grow photoheterotrophically on benzoate (Fig. 6.10C), altered DNA-binding of BadM (Fig. 6.9), and restored pigmentation of the Δ*badL* strain to wildtype levels (Fig. 6.13B). For these reasons, one could argue 4-aminobenzoate may be the preferred substrate for BadL *in vivo*. Addition of 2-, 3-, and 4-acetamidobenzoates decreased the lag phase of *badL*$^{+}$ cells indicating that these compounds may be used to allow for optimal growth of *R. palustris* under photosynthetic conditions.

BadL activity is required for the expression of light harvesting proteins

The initial observation regarding the difference in pigmentation between Δ*badL* and *badL*$^{+}$ strains was intriguing because there was no reason to suspect a link between BadL function and energy conservation at the gene expression level. Since benzoate is only catabolized by *R. palustris* under photosynthetic conditions, it is reasonable that benzoate catabolism and photosynthesis would be linked in this organism. Clearly, the extraction of reducing equivalents from carbon sources requires that reduced electron carriers be oxidized so they can continually participate in carbon catabolism. In microorganisms that have electron transport systems, the oxidation of reduced electron carriers results in proton extrusion with the concomitant generation of a proton motive force. Data presented in this paper (Fig. 6.10-6.13) uncover an unexpected role for the BadL acetyltransferase in the regulation of expression of genes encoding proteins that are later assembled into light harvesting complexes. To our knowledge, a connection between BadL function and *puc* and *puf* gene expression has not been reported. Even though at this point we do not understand how BadL function is connected to energy generation, two points are worth discussing. First, we note that the observed decrease in light harvesting 1(LH1) and light harvesting 2 (LH2) complexes in a Δ*badL* strain is not affected by the absence of the BadM repressor (Figs. 6.12, 6.13A), and second, 4-

acetamidobenzoate partially corrects the pigmentation phenotype (Fig. 6.13B), consistent with the need for the newly assigned BadL function as a source of acetamidobenzoates. This unexpected new role for acetamidobenzoates in the expression of genes encoding functions involved in photosynthesis is an exciting contribution to the field that warrants further investigation.

Importance of BadL in *Rhodopseudomonas palustris* physiology

From a physiological stand point, the most intriguing question raised by this work is why does *R. palustris* rely on an *N*-acetyltransferase to control the transcription of genes required for photosynthetic growth on benzoate? One possible explanation is that aminobenzoates are very abundant in the environments occupied by *R. palustris*, and that the high levels of such compounds likely exerted a strong selective pressure for the evolution of a sensing mechanism that would activate the expression of genes encoding benzoate catabolic enzymes. At present, it appears as if the lack of specificity of BadL for a given aminobenzoate may be an advantage since BadM also appears to bind diverse acetamidobenzoates. To begin addressing these questions, the presence of aminobenzoates within the environment must be investigated. The connection of acetamidobenzoates to the expression of genes whose protein products are needed for the generation of a light-driven proton motive force is equally exciting, and further work in this area of *R. palustris* physiology will likely uncover new knowledge regarding the physiology of other bacteria that also appear to use *N*-acetyltransferases to regulate gene expression in response to diverse environmental stimuli.

MATERIALS AND METHODS

Chemicals, bacterial strains, culture media and growth conditions.

All chemicals were purchased from Sigma Aldrich with the following exceptions: [1-^{14}C]-Acetyl-CoA (Moravek, 15 mCi mmol^{-1}), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, GoldBio), isopropyl β-D-1-thiogalactopyranoside (IPTG, GoldBio), ampicillin (GoldBio), 2-acetamidobenzoic acid (Alfa Aesar), 3- and 4-acetamidobenzoic acid (VWR). All strains and plasmids used in this study are listed

in Tables 6.1 and 6.2. *Escherichia coli* strains DH5α (New England Biolabs) or C41 (λDE3) (32) were grown on lysogeny broth (LB, Difco) at 37°C. When used, antibiotics were added at the following concentrations: kanamycin (50 μg mL^{-1}) and ampicillin (100 μg mL^{-1}). All *Rhodospeudomonas* strains used in this study were derivatives of *Rhodopseudomonas palustris* CGA009 (33).

Rhodopseudomonas palustris was grown at 30 °C on YP rich medium (10 g yeast extract L^{-1} and 20 g peptone L^{-1}) or photosynthetic medium [PM (34)] supplemented with succinate (10 mM, pH 7 at 25°C) or benzoate (3 mM, pH 7 at 25°C) with $NaHCO_3$ (10 mM). When specified, kanamycin was added at 75 μg mL^{-1}. When added, ABAs and acetamidobenzoates stock solutions (100 mM) were dissolved in water and the pH was adjusted with NaOH until all material was dissolved. ABAs were added to growth medium at a final concentration of 5 mM. For growth analysis, cells were grown aerobically in 5 mL of YP medium for three days and supplemented with kanamycin when needed. Cultures were inoculated (1:30, v/v; cells:fresh medium) in triplicate into 5 mL of PM in Balch tubes (35) supplemented with specified carbon sources and antibiotics, as listed in the Figure legends. Tubes were stoppered and aluminum seals were crimped in place. The head space was degassed with O_2-free N_2 gas for 15 min. Cells were incubated without shaking at 30 °C in a light chamber equipped with eight light bulbs (60W, 120V, 840 lumens) with tubes 50 cm from light source. Cell density was monitored at 660 nm using a Manostat Spec 20D and all growth curves were repeated at least three times. *Geobacter metallireducens* GS-15 genomic DNA was a gift from Daniel Bond (University of Minnesota) and *Magnetospirillum magneticum* AMB-1 genomic DNA was a gift from Arash Komeili (UC Berkeley).

Molecular techniques.

Primers were synthesized from Integrated DNA Technologies and are listed in Table 6.3. Genomic DNA was synthesized using ethanol precipitation (36). DNA manipulations were performed using standard techniques (37). DNA was amplified using Phusion High-Fidelity DNA polymerase (New England Biolabs) following manufactures protocol for amplification of high GC DNA (GC buffer and 3% DMSO, v/v). PCR products were analyzed on agarose (1% w/v) gels developed at 100 V. PCR products were

Table 6.1. Bacterial Strains and plasmids used in this study		
Strain	**Relative Genotype**	**Source[1]**
E. coli strains		
E. coli DH5α	Φ80d*lacZ*ΔM15 *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 (r_k-, m_k^+) *supE*44 *relA*1 *deoR* Δ(*lacZYA-argF)* U169 *phoA*	NEB
E. coli C41 (λDE3)	*pka12::kan*$^+$ *ompT hsdS* (r_Bm_B) *gal* (λDE3*)*	Laboratory collection
R. palustris strains		
JE11365	*R. palustris* CGA009 (wild-type strain)	C. Harwood
Derivatives of JE11365		
JE11529	*badL*$^+$ / pBBR1MCS-2	
JE19690	*badL*$^+$ / pAadR3	
JE12597	Δ*badL*	
JE13235	Δ*badL* / pBBR1MCS-2	
JE13236	Δ*badL* / pBadL3	
JE19220	Δ*badM*	
JE19534	Δ*badM* / pBBR1MCS-2	
JE19221	Δ*badL* Δ*badM*	
JE19535	Δ*badL* Δ*badM* / pBBR1MCS-2	
JE19537	Δ*badL* Δ*badM* / pBadM3	
JE19691	Δ*badL* / pAadR3	

[1]Unless otherwise noted, all plasmids and strains were constructed in this study

Table 6.2. Plasmids used in this study		
Plasmid	**Genotype**	**Source**[1]
pTEV6	*lacI*$^+$ *malE*$^+$ *bla*$^+$	(38)
pTEV18	*lacI*$^+$ *bla*$^+$	(39)
pTEV20	*lacI*$^+$ *bla*$^+$	(39)
pBBR1MCS-2	*kan*$^+$	(40)
pK18mobsacB	*kan*$^+$ *sacB*$^+$	(41)
Overexpression plasmids		
p*Rp*BadL1	*R. palustris badL*$^+$ cloned into pTEV6, *bla*$^+$	
pBadM8	*R. palustris badM*$^+$ cloned into pTEV20, *bla*$^+$	
p*Mm*BadL	*M. magneticum badL*$^+$ cloned into pTEV18, *bla*$^+$	
p*Gm*BadL	*G. metallireducens badL*$^+$ cloned into pTEV18, *bla*$^+$	
Complementation plasmids		
pBadL3	*R. palustris badL*$^+$ cloned into pBBR1MCS-2, *kan*$^+$	
pBadM3	*R. palustris badM*$^+$ cloned into pBBR1MCS-2, *kan*$^+$	
pAadR3	*R. palustris aadR*$^+$ cloned into pBBR1MCS-2, *kan*$^+$	
***R. palustris* deletion constructs**		
pBadL2	Upstream and downstream regions of *badL* cloned into pK18mobsacB	
pBadM2	Upstream and downstream regions of *badM* cloned into pK18mobsacB	
pBadLM1	Upstream and downstream regions of *badLbadM* cloned into pK18mobsacB	

[1]Unless otherwise noted, all plasmids and strains were constructed in this study

Table 6.3. Primers used in this study	
Primer Name	**Primer Sequence[3] 5' → 3'**
Overexpression primers	
5' *badL* pTEV6 KpnI	TTA**GGTACC**ATGAGTGGCGCAAGCAGCATC
3' *badL* pTEV6 HindIII	GGCG**AAGCTT**TCAGTCGAGATTCTTTTCCAT
5' *badM* pTEV20	NN**GCTCTTC**NTACATGCCGATGCGGCTGCAGAAATC
3' *badM* pTEV20	NN**GCTCTTC**NTTCCTAAGCCCGTCCGGCGGAGGAAT
5' *M.m. badL* pTEV18	NN**GCTCTTC**NTTCATGGCAAATCCACAGGCTCGGCAAT
3' *M.m. badL* pTEV18	NN**GCTCTTC**NTTATCAGAGATCCTTCTCGAACTGGATGAAG
5' *G.m. badL* pTEV18	NN**GCTCTTC**NTTCATGAGCGATCAATTATCCATCCGTA
3' *G.m. badL* pTEV18	NN**GCTCTTC**NTTATCAGTCGAGCGGTTTTTCCAACTCGAT
Complementation primers	
5' *badL* pBBR1MCS-2	TTA**GGTACC**ATGAGTGGCGCAAGCAGCATC
3' *badL* pBBR1MCS-2	GGCG**AAGCTT**TCAGTCGAGATTCTTTTCCAT
5' *badM* pBBR1MCS-2 HindIII	NNNNNN**AAGCTT**CGGTCGAGACCAACACGAAGATTG
3' *badM* pBBR1MCS-2 XbaI	NNNNNN**TCTAGA**CTAAGCCCGTCCGGCGGAGGAAT
5' *aadR* pBBR1MCS-2 HindIII	NNNNNN**AAGCTT**ATGCCGCATCTCGCTTATCCGA
3' *aadR* pBBR1MCS-2 XbaI	NNNNNN**TCTAGA**TCAGGCCGCGGCGAGCGCGT
***R. palustris* deletion construct primers**	
5' upstream *badL* EcoRI	GTC**GAATTC**GATCGAGCACGACATGGAT
3' upstream *badL*	GACAACGCGATGAGTGGCGCAAGCTGATCG AAACGCCCGACACGCTCG
5' downstream *badL*	CGAGCGTGTCGGGCGTTTCGATCAGCTTGCGCCAC TCATCGCGTTGTC
3' downstream *badL* HindIII	GAT**AAGCTT**GAACACGGCCTGACCAAG
5' upstream *badM* EcoRI	NNNNNN**GAATTC**CCGAACCTGTCGGTCGAAGAGAACC
3' upstream *badM*	CAATCTTCGTGTTGGTCTCGACCGCAAACCGCGCGA GCGTGTCGG
5' downstream *badM*	ACCTTAGCTGAACTTTCAGTCGGCAGCG
3' downstream *badM* BamHI	NNNNNN**GGATCC**TGCTGCACATCTTCCTGTCGAATTTGC
5' upstream *badLM* XbaI	NNNNNN**TCTAGA**TGCAGATCGACGAGCTCGACGTCA
3' upstream *badLM*	CGCGTTGTCCTCCCTCGTGGTGCCGCGTGATCCCGGC TCCTGATG
5' downstream *badLM*	CACCACGAGGGAGGACAACGCGACCTTAGCTGAAC TTTCA GTC GGCAGCG
3' downstream *badLM* HindIII	NNNNNN**AAGCTT**TGCTGCACATCTTCCTGTCGA ATTTGC
5' *badLM* del check	ACGAGATTCGGCGCAACGAAG
3' *badLM* del check	CAAGAGCGATCACGAGGTGCAG
Electrophoretic Mobility Shift Assay Primers	
5' *badCD* intergenic	AAGATCGCGATCGATATCTGAG
3' *badCD* intergenic	GTGATGCGGCCTGTCGAGC
RT-qPCR primers	
5' *badD*	ACCACGGCGGACATCATC
3' *badD*	GGTGAAACTGAGATCGGAGAACA
5' *fixJ*	GCGATGCGGGAGTCGAT
3' *fixJ*	TGTCTGCGCGGATTCGTA
5' *pucC*	CTTTACGCAGGTCCAGGTTG
3' *pucC*	GTTGAAGGTCGTCTGCGTTT
5' *pufM*	CGGCAATTGGTTACTGCTTC
3' *pufM*	CGTAGTTCGTTCGTGTCGTG

purified using the Wizard SV Gel and PCR Clean-Up System (Promega) and plasmids were purified using the Wizard Plus SV Miniprep kit (Promega). DNA sequencing was performed at the Georgia Genomics Facility (Athens, GA, USA).

The following genes with their locus tags were PCR amplified as described above from purified gDNA: *R. palustris badM* (RPA0663, TX73_RS03425), *R. palustris badL* (RPA0664, TX73_RS03430), *G. metallireducens badL* (gmet_2096), and *M. magneticum badL* (amb3392, AMB_RS17160). Gene products (*badM, G.m. badL* and *M.m. badL*) and overexpression vectors (pTEV20 or pTEV18) were digested with restriction enzyme BspQI and cloned using published protocols (39, 42). *R. palustris badL* gene product and overexpression vector (pTEV6) were digested with KpnI and HindIII restriction sites as described (37). Plasmids were transformed into *E. coli* DH5α and plated on LB agar supplemented with ampicillin. Positive clones were screened via colony PCR using T7 promoter forward and reverse primers. The resulting plasmids were: p*Rp*BadM8 (BadM fused to a His_6 recombinant tobacco etch virus (rTEV) prtotease-cleavable *C*-terminal tag), p*Rp*BadL1 (BadL fused to a His_6-MBP rTEV-cleavable *N*-terminal tag), p*Gm*BadL1, and p*Mm*BadL2 (*Gm or Mm*BadL fused to a His_6 rTEV-cleavable *N*-terminal tag). rTEV was overproduced and purified in house using described protocols (43).

R. palustris badL, badM, and *aadR* (RPA4234, TX73_RS21595) were amplified and purified as described above. The *badM* and *aadR* gene products were cloned into HindIII and XbaI sites of pBBR1MCS-2, resulting in p*Rp*BadM3 and p*Rp*AadR3. The *badL* gene product was cloned into KpnI and HindIII sites of pBBR1MCS2, resulting in p*Rp*BadL3. Ligated plasmids were transformed into *E. coli* DH5α and plated on LB agar supplemented with kanamycin. Clones of interest were screened via colony PCR using M13 forward and reverse primers.

In-frame deletions of *R. palustris badL* and *badM* genes were constructed using described methods (41). Briefly, 1-kb regions upstream and downstream of targeted genes were amplified and fused using overlap extension PCR (44). The PCR products were digested with either EcoRI and HindIII (*badL*), EcoRI and BamHI (*badM*), or XbaI and HindIII (*badL badM*) and cloned into pK18*mobsacB* (41). Ligated

plasmids were transformed into *E. coli* DH5α and plated on LB agar supplemented with kanamycin. Clones of interest were screened via colony PCR using M13 forward and reverse primers. Deletion construct plasmids were electroporated into exponentially growing *R. palustris* (washed twice with 10% glycerol, v/v) and plated on PM-succinate agar plates + kanamycin. Cells were grown photosynthetically in an Annoxamat Jar (Spiral Biotech) until single colonies appeared. Cells were streaked to isolation on PM-succinate agar plates + kanamycin and grown photosynthetically. Multiple colonies were then streaked to isolation for counterselection on PM-succinate + sucrose (10%, v/v, added after autoclaving) and grown under photosynthetic conditions. Single colonies were screened for acquisition of the deletion and deletions were confirmed via DNA sequencing. For construction of Δ*badL* Δ*badM*, both genes were deleted simultaneously in a *badL*$^{+}$ *badM*$^{+}$ *R. palustris* strain. For complementation of genes, plasmids constructed as described above were electroporated into *R. palustris* and selected for transformation on PM-succinate + kanamycin agar plates.

Purification of BadM and BadL proteins.

Plasmids p*Rp*BadM8, p*Rp*BadL1, p*Gm*BadL1, and p*Mm*BadL2 were transformed into *E. coli* C41 (λDE3) *pka12*::*kan*$^{+}$ (JE9314) cells. The resulting strains were grown overnight in 50 mL LB + ampicillin. Overnights were sub-cultured (1:100) into two liters of LB + ampicillin and grown at 25°C to an OD_{650} of 0.5, after which transcription of genes was induced by the addition of IPTG (0.5 mM). Cells were harvested the next day by centrifugation at 6000 x *g* for 15 min at 4°C in an Avanti J-2 XPI centrifuge equipped with rotor JLA-8.1000 (Beckan Coulter). Cell pellets were stored at -80°C until used.

Cell pellets were thawed and resuspended in 30 mL of buffer A [HEPES (50 mM, pH 7.5 at 4°C), NaCl (0.5 M), imidazole (20 mM), and glycerol (20% v/v)] with lysozyme (1 mg mL^{-1}), DNase (1 μg mL^{-1}), and protease inhibitor phenylmethanesulfonyl fluoride (PMSF, 0.5 mM). Cells were lysed on ice with two rounds of sonication [1 min (2 s, 50% duty)] using a Q500 Sonicator (VWR) at 60% amplitude. Cell lysates were clarified via centrifugation at 40,000 x *g* for 30 min at 4°C in an Avanti J-25I centrifuge (Beckman Coulter) equipped with rotor JA-25.50. Supernatants were filtered through a 0.45 μm filter

(Millipore) and samples were applied at 4°C to a pre-equilibrated 1-mL HisPur nickel-nitrilotriacetic acid (Ni-NTA) resin (Thermo Scientific). The column was washed with 10 column volumes (CV) of buffer A, 5 CV of buffer B [HEPES (50 mM, pH 7.5 at 4°C), NaCl (0.5 M), imidazole (80 mM), and glycerol (20% v/v)], and 5 CV of buffer C [HEPES (50 mM, pH 7.5 at 4°C), NaCl (0.5 M), imidazole (0.5 M), and glycerol (20% v/v)]. Proteins were separated by SDS-PAGE and concentrations of target protein containing fractions were quantified on a NanoDrop (Thermo Fischer) using the extinction coefficient and molecular weight of the protein (ExPASy, ProtParam).

Proteins were cleaved at 25°C for 3 h with a 1:50 mg:mg ratio of rTEV to target protein. After cleavage, proteins were dialyzed in buffer D [HEPES (50 mM, pH 7.5 at 4°C), NaCl (0.5 M), EDTA (1 mM), and glycerol (20% v/v)], and twice in buffer A. After dialysis, cleaved proteins were applied to a pre-equilibrated 1 mL HisPur Ni-NTA resin. Proteins that did not bind to the column were collected and dialyzed against three buffers with decreasing concentrations of NaCl (400 mM, 0.25 M, 150 mM), with a final buffer composition of [HEPES (50 mM, pH 7.5 at 4°C), NaCl (0.15 M), and glycerol (20% v/v)]. Protein concentrations were determined with a NanoDrop and were flash frozen in liquid N_2 and stored at -80°C until used.

Thin-layer chromatography (TLC).

Reaction mixtures included HEPES (50 mM, pH 7.5 at 25°C), TCEP (0.5 mM), [1-^{14}C]-Acetyl-CoA (20 μM), aromatic substrates (as listed in Figures, 5 mM), and BadL protein (3 μg). Reaction mixtures were incubated at 37°C for 1 h and 5 μL of each reaction was spotted and dried onto a pre-scored silica gel plate with aluminum backing (Whatman Ltd). Plates were developed for 2-3 h in a pre-equilibrated chamber with *n*-butanol, acetic acid, and water (3:1:1). Plates were dried, exposed to a phosphor screen, and imaged using a Typhoon Trio+ Variable Mode Imager (GE Healtcare) and analyzed with ImageQuant v5.2 software.

HPLC and MS/MS analyses.

Reactions containing BadL (10 μM), acetyl-CoA (50 μM), aminobenzoates (100 μM), TCEP (0.5 mM), and sodium phosphate buffer (pH 8 at 25° C) were incubated for 2 h at 37°C. BadL was removed by passing reactions over an Amicon Ultra 0.5 mL 10kDa molecular cut off centrifugal filter (Millipore). The product of BadL aminobenzoate acetylation was resolved by RP-HPLC using a Shimadzu Prominence UFLC with a Kinetex 5μm C18 column (150 mm x 4.6 mm; Phenomenex). The column was equilibrated at a flow rate of 0.5 ml min^{-1} with 5 column volumes of elution buffer [50 % (v/v, in water) acetonitrile] and 5 column volumes of wash buffer [10 mM sodium phosphate in water (pH 8 at 25° C)]. Samples (100 μl) were injected and the column was developed with wash buffer for 5 min and then a gradient to elution buffer over 10 min. Compounds were detected and compared to standards at 254 nm using a computer-controlled Shimadzu Nexera X2 SPD-30A diode array detector. Data were analyzed using Prism v6 (GraphPad). Fractions were collected in 0.5 mL increments and fractions containing peaks corresponding to acetamidobenzoate were analyzed by MS and MS/MS (Protein and Mass Spectrometry Facility, UGA, Athens, GA, USA). Electrospray ionization (ESI)-MS was performed in acetonitrile and resolved on an Esquire 3000 Plus (Bruker) Ion Trap Mass Spectrometer at 0.3 ml h^{-1}. Pure standards (3-ABA and 3-ABA^{Ac}, 1 mM) were suspended in DMSO and run as described above.

DNA-binding assays.

Electrophoretic mobility shift assays (EMSAs) were performed using DNA probes containing a 6-carboxyfluorescein (6-FAM) label covalently attached to the 3' end of synthesized PCR products. Probes were generated using primers listed in Table 6.3 and included the intergenic region (212 bp) of *badC* and *badD* as described (6). Binding reaction mixtures (25 μL) contained 6-FAM double-stranded DNA probe (50 ng, 0.303 pmol), BadM buffer [Tris (20 mM, pH 7.5 at 25°C), KCl (50 mM), dithiothreitol (1 mM), and glycerol (8% v/v)], bovine serum albumin (BSA, 100 μg mL^{-1}), BadM (0 or 1.82 pmol), and acetamidobenzoates (10 mM). All acetamidobenzoates used (*i.e.,* 100 mM stock solutions of 2-, 3-, and 4-acetamidobenzoates were prepared in DMSO (100%), added to the reaction tube (to generate a master mix

for three reactions), and immediately diluted with the appropriate amount of water followed by the addition of buffer, BSA, and protein. If DMSO stocks were added to water, precipitation occurred, hence order of addition was important. Reactions were incubated at 37°C for 10 mins and 5 μL of glycerol (50%, v/v) was added to the mixture. A non-denaturing gel with Tris-HCl (365 mM, pH 8.6) and polyacrylamide (7.5% w/v) (Criterion, BioRad) was pre-developed for 30 min at 110 V with 0.5x TBE (Tris base, boric acid, EDTA) buffer. Reactions (20 μL) were pipetted into the wells of the pre-developed gel and resolved at 110 V for 2-3 h. Gels were imaged using a Typhoon Trio+ variable mode imager (GE Healthcare) at wavelength 488 nm (blue setting) and analyzed with ImageQuant v5.2 software.

RNA isolation.

Strains JE11529 (*badL*$^{+}$ / pBBR1-MCS2), JE13235 (Δ*badL* / pBBR1-MCS2), and JE13236 (Δ*badL* / p*Rp*BadL3) were each grown in quadruplicate in 5 mL of YP medium + kanamycin for four days. Cells were diluted (1:20; v/v) into 5 mL of fresh PM + succinate and grown photosynthetically until cells reached mid-exponential (OD_{660} ~0.5, 24 h later). Exponentially growing cells were back diluted into 15 mL of fresh PM + succinate to an OD_{660} of 0.03 and grown photosynthetically at 30 °C until cultures reached an OD of 0.1, at which point benzoate (3 mM) and $NaHCO_3$ (10 mM) were added to the cultures with a sterile syringe and needle. Cells were grown for 24 h after the addition of benzoate, and all 15 mL was harvested by centrifugation at 4,000 x *g* for 10 min. Supernatants were decanted, cells were immediately centrifuged at 4,000 x *g* for 1 min and excess medium was aspirated off and cells were flash frozen and stored at -80°C until used. RNA was isolated using the RNAsnap method (45).

Cell pellets were resuspended in 150 μL of boil solution [EDTA (18 mM), sodium dodecyl sulfate (SDS, 0.025% v/v), RNA-grade formamide (95% v/v), 2-mercaptoethanol (1 % v/v)] in water and transferred to 1.7-mL Eppendorf tubes. Mixtures were incubated at 100°C for 7 min and immediately centrifuged at 16,000 X *g* for 5 min. A sample (100 μL) of supernatant was transferred to a fresh 2.0-mL Eppendorf tube containing 400 μL of RNase-free water and 50 μL of 3M sodium acetate, pH 5.2. Following RNA dilution, 1.65 mL of ice-cold ethanol (100%) was added to the reaction mixtures and tubes were

incubated at -80°C for at least 1 h. After DNA/RNA precipitation, reaction mixtures were poured into 1.7-mL Eppendorf tubes and centrifuged at 16,000 x *g* for 1 h. Supernatants were decanted and pellets were washed with 300 µL of ice-cold ethanol (70% v/v). Supernatants were decanted, tubes were immediately centrifuged, excess supernatant was decanted and tubes were allowed to dry upside down on Kimwipes for 20 min. DNA/RNA pellets were resuspended in 100 µL of RNase-free water and centrifuged at 16,000 x *g* to remove non-soluble debris. Subsequent DNase I treatment was carried out with 90 µL of the DNA/RNA suspension and the Ambion Turbo DNA-free kit (Thermo Fischer Scientific) according to manufacturer's instructions for rigorous DNase treatment. After DNA cleavage, a final sodium acetate-ethanol precipitation was performed as described above. Dried RNA pellets were resuspended in 100 µL of RNase-free water and frozen at -80°C in 20-µL aliquots. A small sample for each RNA prep was sent to the Georgia Genomics Facility (Athens, GA, USA) for quality control analysis using the RNA 600 Nano kit of the Agilent 2100 bioanalyzer. Any RNA samples that had a RNA Integrity Number (RIN) of above 8.0 were used for subsequent qRT-PCR experiments.

cDNA synthesis and quantitative reverse-transcription PCR.

Primers for qRT-PCR were designed using primer 3 software and were evaluated for specificity and melting curve. Total RNA (150 ng) from each sample was used with the iScript cDNA synthesis kit (BioRad) following the manufacturer's protocol. The final RNA concentration of the cDNA reaction mixture was 7.5 ng/µL and was used as a template for qRT-PCR. Master mixes were prepared in RNase-free water with FastSYBR green master mix (1X, Applied Biosystems) and gene-specific primers (1 µL of each primer from a 10 µM stock). qRT-PCR reactions (20 µL) were assembled with 15 ng of cDNA (2 µL of the 7.5 ng/µL cDNA) and 18 µL of the above-mentioned master mix into a MicroAmp™ Fast Optical 96-well reaction plate (Thermo Fischer Scientific). The real-time PCR was performed using a 7500 Fast real-time PCR system (Applied Biosystems). Cycle threshold (C_T) data were normalized to the *fixJ* (RPA4248) gene as described elsewhere (46). The normalized ΔC_T values were transformed using $2(e-\Delta C_T)/10^{-6}$ (47) and were reported as arbitrary expression units (EU). Mean EU values calculated from

technical triplicates of biological triplicates were used to calculate the standard error of the mean in Prism6 software. Statistically significant differences between expression across strains were determined using a Welch's *t*-test with GraphPad Prism6 software. Figure legends report *p* values for each sample.

***R. palustris* pigment analysis.**

R. palustris strains were grown in triplicate for three days in YP medium aerobically. Cells were sub-cultured (1:30) into 10 mL of PM + benzoate (3 mM) and $NaHCO_3$ (10 mM). The headspace of Balch culture tubes was flushed with O_2-free N_2 gas, and cultures were grown photosynthetically as described above. At lag (OD_{660} ~0.20), log (OD_{660} ~0.6), and stationary phase (OD_{660} >1) of each replicate, 200 μL was removed from the Balch tube with a sterile needle and syringe. Culture samples (100 μL) were pipetted into a 96-well plate and scanned ($A_{600\text{-}1000}$) using a Spectramax UV-vis spectrophotometer. Replicates were plotted using GraphPad Prism6 software to determine mean and the standard error of the mean (SEM) for each strain.

ACKNOWLEDGMENTS

We thank Heidi Crosby for fruitful discussions and to the Proteomics and Mass Spectrometry Core Facility of The University of Georgia for the performance and analysis of LC/MS/MS.

COMPETING INTERESTS. The authors do not have any competing interests with the contents of this paper.

REFERENCES.

1. Hunter N.C. DF, Thurnauer M.C., Beatty J.T. 2009. The Purple Phototrophic Bacteria, vol 28. Springer, Netherlands.
2. Heider J, Fuchs G. 1997. Anaerobic metabolism of aromatic compounds. Eur J Biochem 243:577-596.

3. Harwood CS, Burchhardt G, Herrmann H, Fuchs G. 1999. Anaerobic metabolism of aromatic compounds via the benzoyl-CoA pathway. FEMS Microbiol Rev 22:439-458.

4. Harwood CS, Gibson J. 1997. Shedding light on anaerobic benzene ring degradation: A process unique to prokaryotes? J Bacteriol 179:301-309.

5. Rey FE, Harwood CS. 2010. FixK, a global regulator of microaerobic growth, controls photosynthesis in *Rhodopseudomonas palustris*. Mol Microbiol 75:1007-1020.

6. Hirakawa H, Hirakawa Y, Greenberg EP, Harwood CS. 2015. BadR and BadM proteins transcriptionally regulate two operons needed for anaerobic benzoate degradation by *Rhodopseudomonas palustris*. Appl Environ Microbiol doi:10.1128/AEM.00377-15.

7. Dispensa M, Thomas CT, Kim MK, Perrotta JA, Gibson J, Harwood CS. 1992. Anaerobic growth of *Rhodopseudomonas palustris* on 4-hydroxybenzoate is dependent on AadR, a member of the cyclic AMP receptor protein family of transcriptional regulators. J Bacteriol 174:5803-5813.

8. Dutton PL, Evans WC. 1969. The metabolism of aromatic compounds by *Rhodopseudomonas palustris*. A new, reductive, method of aromatic ring metabolism. Biochem J 113:525-536.

9. Harwood CS, Gibson J. 1986. Uptake of benzoate by *Rhodopseudomonas palustris* grown anaerobically in light. J Bacteriol 165:504-509.

10. Geissler JF, Harwood CS, Gibson J. 1988. Purification and properties of benzoate-coenzyme A ligase, a *Rhodopseudomonas palustris* enzyme involved in the anaerobic degradation of benzoate. J Bacteriol 170:1709-1714.

11. Egland PG, Pelletier DA, Dispensa M, Gibson J, Harwood CS. 1997. A cluster of bacterial genes for anaerobic benzene ring biodegradation. Proc Natl Acad Sci U S A 94:6484-6489.

12. Harrison FH, Harwood CS. 2005. The *pimFABCDE* operon from *Rhodopseudomonas palustris* mediates dicarboxylic acid degradation and participates in anaerobic benzoate degradation. Microbiology 151:727-736.

13. Pelletier DA, Harwood CS. 1998. 2-Ketocyclohexanecarboxyl coenzyme A hydrolase, the ring cleavage enzyme required for anaerobic benzoate degradation by *Rhodopseudomonas palustris*. J Bacteriol 180:2330-2336.

14. Pelletier DA, Harwood CS. 2000. 2-Hydroxycyclohexanecarboxyl coenzyme A dehydrogenase, an enzyme characteristic of the anaerobic benzoate degradation pathway used by *Rhodopseudomonas palustris*. J Bacteriol 182:2753-2760.

15. Boll M, Albracht SS, Fuchs G. 1997. Benzoyl-CoA reductase (dearomatizing), a key enzyme of anaerobic aromatic metabolism. A study of adenosinetriphosphatase activity, ATP stoichiometry of the reaction and EPR properties of the enzyme. Eur J Biochem 244:840-851.

16. Gibson KJ, Gibson J. 1992. Potential early intermediates in anaerobic benzoate degradation by *Rhodopseudomonas palustris*. Appl Environ Microbiol 58:696-698.

17. Hentchel KL, Escalante-Semerena JC. 2015. Acylation of biomolecules in prokaryotes: a widespread strategy for the control of biological function and metabolic stress. Microbiol Mol Biol Rev 79:321-346.

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

19. Thao S, Chen CS, Zhu H, Escalante-Semerena JC. 2010. N(epsilon)-Lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. PLoS One 5:e15123.

20. Crosby HA, Heiniger EK, Harwood CS, Escalante-Semerena JC. 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.

21. Tucker AC, Escalante-Semerena JC. 2013. Acetoacetyl-CoA synthetase activity is controlled by a protein acetyltransferase with unique domain organization in *Streptomyces lividans*. Mol Microbiol 87:152-167.

22. VanDrisse CM, Escalante-Semerena JC. 2018. In *Streptomyces lividans*, acetyl-CoA synthetase activity is controlled by O-serine and N(epsilon)-lysine acetylation. MolMicrobiol 107:577-594.

23. You D, Wang MM, Ye BC. 2017. Acetyl-CoA synthetases of *Saccharopolyspora erythraea* are regulated by the nitrogen response regulator GlnR at both transcriptional and post-translational levels. Mol Microbiol 103:845-859.

24. Xu JY, You D, Leng PQ, Ye BC. 2014. Allosteric regulation of a protein acetyltransferase in *Micromonospora aurantiaca* by the amino acids cysteine and arginine. J Biol Chem 289:27034-27045.

25. Song L, Wang G, Malhotra A, Deutscher MP, Liang W. 2016. Reversible acetylation on Lys501 regulates the activity of RNase II. Nucleic Acids Res doi:10.1093/nar/gkw053.

26. Hentchel KL, Escalante-Semerena JC. 2015. In *Salmonella enterica*, the Gcn5-related acetyltransferase MddA (formerly YncA) acetylates methionine sulfoximine and methionine sulfone, blocking their toxic effects. J Bacteriol 197:314-325.

27. VanDrisse CM, Parks AR, Escalante-Semerena JC. 2017. A toxin involved in *Salmonella* persistence regulates its activity by acetylating its cognate antitoxin, a modification reversed by CobB sirtuin deacetylase. MBio 8.

28. Cheverton AM, Gollan B, Przydacz M, Wong CT, Mylona A, Hare SA, Helaine S. 2016. A *Salmonella* toxin promotes persister formation through acetylation of tRNA. Mol Cell 63:86-96.

29. Burckhardt RM, Escalante-Semerena JC. 2017. In *Bacillus subtilis*, the SatA (formerly YyaR) acetyltransferase detoxifies streptothricin via lysine acetylation. Appl Environ Microbiol doi:10.1128/AEM.01590-17.

30. Vetting MW, Hegde SS, Javid-Majd F, Blanchard JS, Roderick SL. 2002. Aminoglycoside 2'-*N*-acetyltransferase from *Mycobacterium tuberculosis* in complex with coenzyme A and aminoglycoside substrates. Nat Struct Biol 9:653-658.

31. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, de Crecy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Ruckert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 33:5691-5702.

32. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289-298.

33. Larimer FW, Chain P, Hauser L, Lamerdin J, Malfatti S, Do L, Land ML, Pelletier DA, Beatty JT, Lang AS, Tabita FR, Gibson JL, Hanson TE, Bobst C, Torres JL, Peres C, Harrison FH, Gibson J, Harwood CS. 2004. Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. Nat Biotechnol 22:55-61.

34. Kim M-K, Harwood CS. 1991. Regulation of benzoate-CoA ligase in *Rhodopseudomonas palustris*. FEMS Microbiol Lett 83:199-203.

35. Balch WE, Wolfe RS. 1976. New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. Appl Environ Microbiol 32:781-791.

36. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning: A Laboratory Manual, Second ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

37. Elion EA, Marina P, Yu L. 2007. Constructing recombinant DNA molecules by PCR, p Unit 3.17.1-3.17.12. *In* Ausubel FM, R. Brent RE, Kingston DD, Moore JG, Seidman JA, Smith a, Struhl K (ed), Curr Protoc Mol Biol. Greene Publishing Associates & Wiley Interscience, New York, N.Y.

38. Rocco CJ, Dennison KL, Klenchin VA, Rayment I, Escalante-Semerena JC. 2008. Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*. Plasmid 59:231-237.

39. VanDrisse CM, Escalante-Semerena JC. 2016. New high-cloning-efficiency vectors for complementation studies and recombinant protein overproduction in Escherichia coli and Salmonella enterica. Plasmid 86:1-6.

40. Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM, 2nd, Peterson KM. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166:175-176.

41. Schafer A, Tauch A, Jager W, Kalinowski J, Thierbach G, Puhler A. 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145:69-73.

42. Galloway NR, Toutkoushian H, Nune M, Bose N, Momany C. 2013. Rapid cloning for protein crystallography using Type IIS restriction enzymes. Crystal Growth & Design 13:2833-2839.

43. Blommel PG, Fox BG. 2007. A combined approach to improving large-scale production of tobacco etch virus protease. Protein Expr Purif 55:53-68.

44. Horton RM, Ho SN, Pullen JK, Hunt HD, Cai Z, Pease LR. 1993. Gene splicing by overlap extension. Methods Enzymol 217:270-279.

45. Stead MB, Agrawal A, Bowden KE, Nasir R, Mohanty BK, Meagher RB, Kushner SR. 2012. RNAsnap: a rapid, quantitative and inexpensive, method for isolating total RNA from bacteria. Nucleic Acids Res 40:e156.

46. Egland PG, Gibson J, Harwood CS. 1995. Benzoate-coenzyme A ligase, encoded by *badA*, is one of 3 ligases able to catalyze benzoyl-coenzyme A formation during

anaerobic growth of *Rhodopseudomonas palustris* on benzoate. J Bacteriol 177:6545-6551.

47. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402-408.

CHAPTER 7

DIACETYLSERINE, A MISSING LINK IN CYSTEINE BIOSYNTHESIS REGULATION IN *SALMONELLA ENTERICA,* IS SYNTHESIZED FROM *O*-ACETYL-SERINE BY OATA (FORMERLY YJGM), AN *N*-ACETYLTRANSFERASE.

[1]VanDrisse C.M. and Escalante-Semerena J.C. 2018

To be submitted to Molecular Microbiology.

ABSTRACT

L-cysteine biosynthesis has been extensively analyzed in *Salmonella enterica*. The cysteine regulon contains the genes whose protein products are necessary to convert sulfate to sulfide, which is eventually reacted with *O*-acetyl-serine to generate cysteine. The LysR type regulator, CysB, is required for activation of the cysteine regulon, and its interaction with various *cys* genes has been thoroughly characterized. Through these previous studies, it was hypothesized that *O*-acetyl-serine undergoes a spontaneous *O*- to *N*- migration to produce *N*-acetyl-serine, that acts as a ligand for CysB. Up until this point, it was unknown if this migration occurs spontaneously *in vivo* or if *N*-acetyl-serine was generated enzymatically. This work characterizes a *S. enterica N*-acetyltransferase, OatA (previously YjgM), that acetylates the N_{α}-amine group of *O*-acetyl-serine, producing diacetyl-serine. Characterization of OatA was carried out through enzyme activity assays, kinetics, size exclusion chromatography, isothermal titration calorimetry, and verification of product via mass spectrometry. *In vivo* work with *S. enterica* showed that overexpression of *oatA* led to shorter lag times on sulfate-limiting medium and that these delayed lag times were due to increased expression of the cysteine regulon (as indicated by RT-qPCR).

INTRODUCTION

L-Cysteine (L-Cys) is one of the 22 proteinogenic amino acids in nature, and one of only two amino acids whose side chains contain a sulfur atom. The sulfhydryl group of L-Cys is critical to maintaining redox homeostasis in cells from all domains of life. The biosynthesis of this important amino acid has been extensively studied (1), and *Salmonella enterica* synthesizes L-Cys *de novo*. Briefly, the pathway reduces sulfate (oxidation state 6^{+}) to sulfide (oxidation state 2^{-}) and then condenses *O*-acetyl-serine (OAS) with sulfide to yield L-Cys and acetate (Fig. 7.1) (1). Expression

of several genes encoding enzymes of the pathway are regulated by a LysR type regulator, CysB (2, 3). CysB has a complex mechanism of transcriptional regulation. Previous literature has characterized binding of CysB to the promoter regions of *cysK* (encodes *O*-acetylserine (thiol)-lyase), *cysJIH* (encode sulfite reductase), *cysP* (encodes thiosulfate-binding protein), and to *cysB* (encodes a LysR regulator) (4-7).

In the literature, CysB binding to its target promoters was allosterically modulated by *N*-acetyl-serine (hereafter (NAS) and *O*-acetyl-serine (hereafter OAS). For years, OAS, the product of CysE (serine transacetylase) and a substrate for CysK, was suspected to be the inducer of the cysteine regulon (8, 9). However, later studies specified that the more likely inducer was *N*-acetyl-serine (10), which is produced from intramolecular *O*- to *N*-acetyl migration that was suggested to occur at a very fast rate (1% per minute at pH 7.6 (11) (Fig. 7.2). It was not known, however, whether this conversion was enzymatic *in vivo* or whether the cell relied on this non-enzymatic conversion to induce the expression of the cysteine regulon. Regardless, studies characterizing CysB ligand and promoter interactions utilized OAS as the source of NAS.

In vitro electrophoretic mobility shift assays, DNA footprinting, and *in vitro* transcription assays revealed activation regions in *cysJIH*, *cysK*, and *cysP* and CysB binding to these regions is apparently altered upon interaction with acetyl-serine (1). CysB interactions with acetyl-serine release it from accessory sites of *cysK*, *cysJIH*, and *cysB* (6, 7, 12). Additionally, the *cysB* promoter region has a repression site, to which affinity of CysB is lowered in the presence of acetyl-serine (5, 12). Overall, CysB is a complex regulator that binds to accessory sites and activation sites of *cysK*, *cysJIH*, *cysP*, and *cysB* promoters and binding is altered in the presence of acetyl-serine.

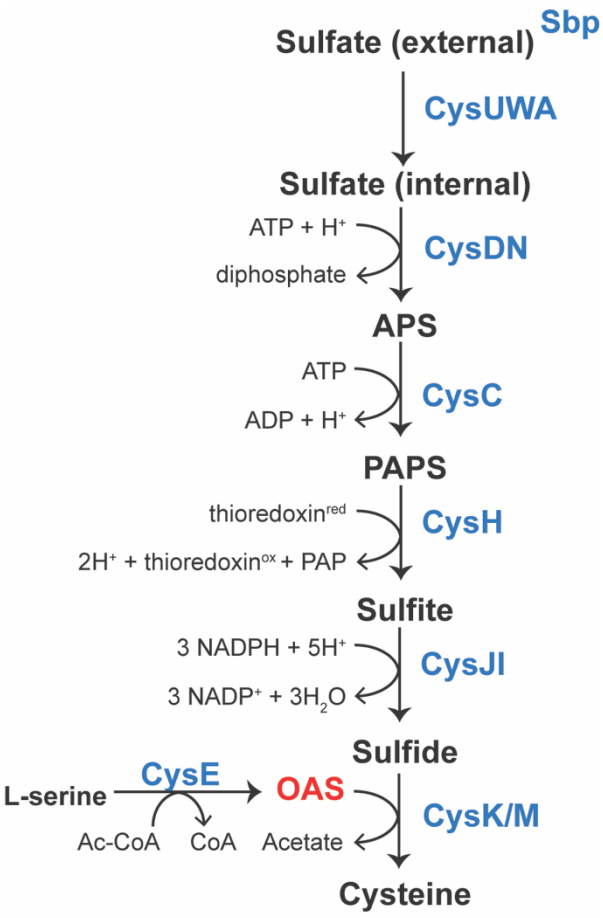


Figure 7.1. Pathway for biosynthesis of L-cysteine in *Salmonella enterica.* Sulfate is sequestered by the sulfate-binding protein (Sbp) and uptake is mediated by the CysUWA sulfate/thiosulfate transporter. The ATP sulfurylase [CysD (catalytic subunit) / CysN (GTP-binding subunit) EC 2.7.7.4] activates sulfate to adenosine 5'-phosphate (APS), which is a substrate for APS kinase (CysC, EC 2.7.1.25). 3-phosphoadenosine 5'-phosphosulfate (PAPS) is reduced to sulfite by the PAPS reductase (CysH, EC 1.8.4.8) and sulfite is reduced to sulfide by a NADPH-sulfite reductase (CysJI, EC 1.8.1.2). A serine acetyltransferase (CysE, EC 2.3.1.30) produces *O*-acetyl-serine (OAS) from L-serine and acetyl-CoA (Ac-CoA). OAS and sulfide are then substrates for cysteine synthase (CysK, EC 2.5.1.47), which produces L-cysteine. Transcription of *cysJIH* and *cysK* is upregulated upon binding of CysB and acetyl-serine.

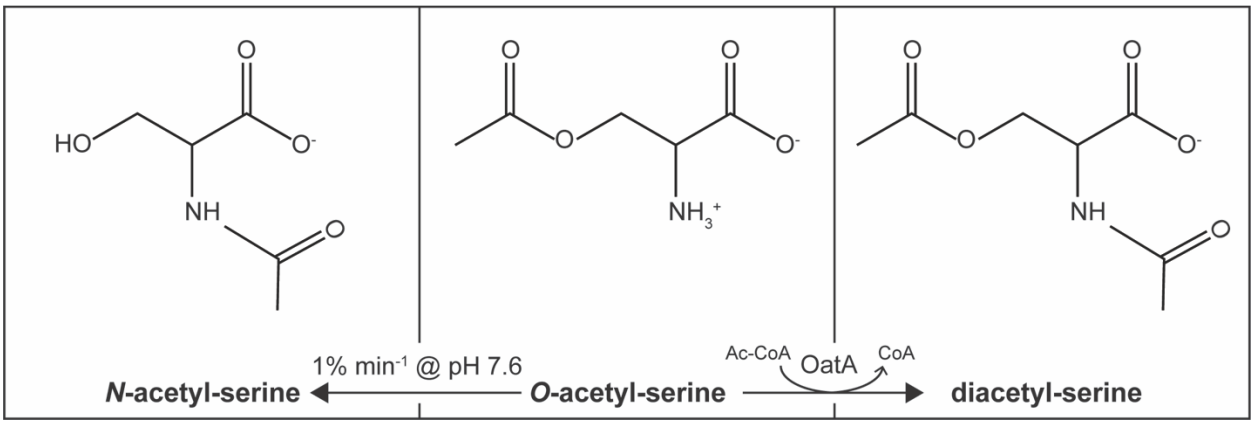


Figure 7.2. Acetyl-serine analogs: *N*-acetyl-serine, *O*-acetyl-serine, and diacetyl-serine. *N*-acetyl-serine (left) is thought to be a ligand for CysB, which is produced from non-enzymatic *O* to *N*-intramolecular migration of *O*-acetyl-serine (center) at pH 7.6 at a rate of 1% min^{-1}. Diacetyl-serine (left) is produced from acetylation of *O*-acetyl-serine by the *Salmonella enterica* OatA *N*-acetyltransferase.

Molecular details on how this change in CysB promoter binding alters transcription remains unclear, although models predict changes in DNA bending are thought to be at play (4, 6, 12).

Acetylation in *Salmonella enterica* and its potential role in cysteine biosynthesis

In *S. enterica*, efforts to identify substrates for the 26 *N*-acetyltransferases coded in the genome has been ongoing, with about half already characterized (13). *N*-acetyltransferases target *N*ε-amine groups of lysines, *N*α-amine groups of proteins, and *N*α-amine groups of small molecules. Acetylation changes the chemical properties of the substrate by changing a protein's activity or changing the way a small molecule interacts with a target. Initially, it was thought many of the *S. enterica* acetyltransferases targeted protein substrates, although much work has been completed that has identified many small molecule substrates such as toxic methionine-derivatives (14, 15), aminoacyl-tRNAs (16-18), and antibiotics (19). Considering *O*-acetyl-serine possess a free *N*α-amine group (Fig. 7.2), there was precedent to investigate acetyltransferases in *S. enterica* that could enzymatically acetylate *O*-acetyl-serine. Using a broad small molecule acetylation substrate screen (20), we purified uncharacterized *S. enterica N*-acetyltransferases and identified an acetyltransferase, OatA (previously YjgM), that acetylated *O*-acetyl-serine *in vitro*, producing diacetyl-serine (Fig. 7.2). The work provided herein characterized OatA through specific activity, kinetics, substrate affinity, and size exclusion chromatography. Through growth analysis and RT-qPCR, we show that OatA is beneficial for growth under conditions that require cysteine biosynthesis, and that overexpression of *oatA*, leads to an increase in transcription of the cysteine regulon. Interactions of CysB and diacetyl-serine proved difficult to establish, and at time of this publication, it is unclear whether or not the product of OatA directly interacts with CysB as a ligand, or if diacetyl-serine positively affects cysteine biosynthesis through an indirect mechanism.

RESULTS

The OatA acetyltransferase (previously YjgM) acetylates *O*-acetyl-serine.

A previously published broad substrate screen was utilized to find putative targets of OatA (for OAS acetyltransferase A) (20) and it was found that OatA acetylated *O*-acetyl-L-serine. A continuous spectrophotometric assay using DTNB was used to assess acetylation of substrates by OatA as described in materials and methods. Activity of OatA was seen for *O*-acetyl-L-serine (438 nmol AMP min^{-1} mg^{-1}) and minimal activity was detected for *N*-acetyl-L-serine (25 nmol AMP min^{-1} mg^{-1}) or L-serine (10 nmol AMP min^{-1} mg^{-1}) (Fig. 7.3). To determine kinetic parameters of OatA for *O*-acetyl-L-serine the same continuous spectrophotometric assay was utilized with acetyl-CoA in excess (1 mM). The apparent K_m for *O*-acetyl-L-serine was 1.1 mM and the catalytic efficiency (k_{cat}/K_m) of OatA was 1.14 x 10^3 M^{-1} s^{-1} (Fig. 7.4). Through spectrophotometric assays, it was determined that OatA converted 50% of OAS to DAS, regardless of enzyme or substrate concentration.

Acetylation of *O*-acetyl-L-serine by OatA produces *N,O*-1,2,-acetyl-serine (DAS).

To verify the acetylated product of OatA, *O*-acetyl-L-serine was incubated with OatA and acetyl-CoA and subsequently purified via HPLC (Fig. 7.5). The diacetyl-serine (DAS, retention time = 13.89 min) was resolved away from OAS (retention time = 6.5 min) and acetyl-CoA/CoA (retention time = 6.5 min). The peak containing DAS was determined not to be spontaneous migration of OAS to NAS due to differences in retention times (DAS, 13.89 min and NAS 12.0 min). This product was analyzed via mass-spectrometry and was found to have a *m/z* ($[MS]^+$ = 190.1 Da and $[MS]^-$ = 188.0 Da) corresponding to an acetyl-group on both the hydroxyl and amine group of L-serine (Fig. 7.6). This, combined with above activity assays, indicated OatA acetylates the amine-group of OAS, producing DAS.

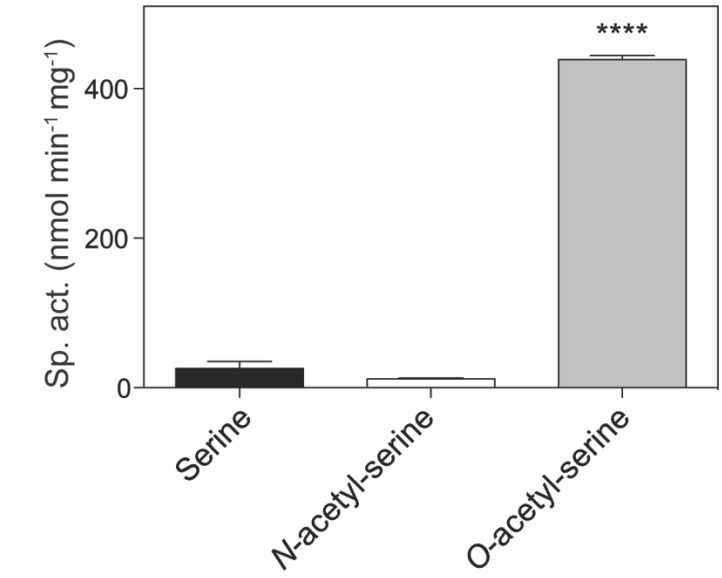


Figure 7.3. OatA acetylation of *O*-acetyl-serine *in vitro.* Using a continuous spectrophotometric assay as described in the Material and Methods, OatA (0.5 μM), acetyl-CoA (500 μM), and either L-serine, *N*-acetyl-L-serine, or *O*-acetyl-L-serine (200 μM) were incubated and specific activity for each substrate was calculated. Error bars represent standard deviation and **** represents a p-value < 0.0001.

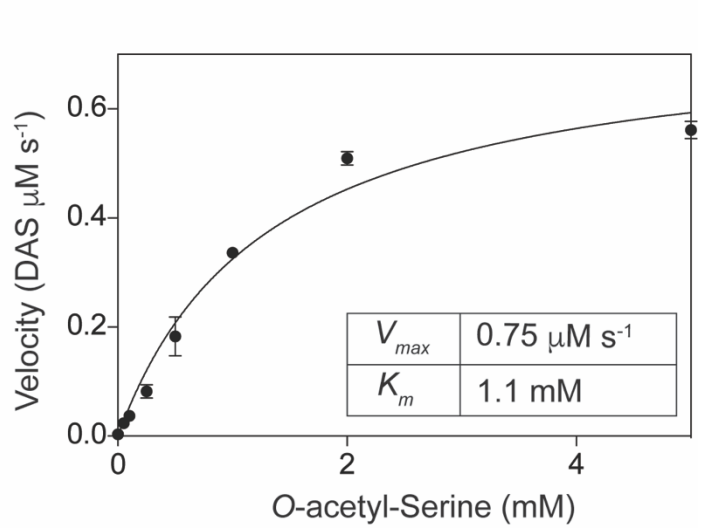


Figure 7.4. Michaelis-Menton saturation curves for OatA-mediated *O*-acetyl-serine acetylation. Using a continuous spectrophotometric assay as described in the Material and Methods, OatA (0.5 μM), acetyl-CoA (1 mM), and *O*-acetyl-serine (concentrations indicated on *x*-axis) were incubated and monitored to calculate V_{max} at each *O*-acetyl-serine concentration. Vmax values were plotted against OAS concentrations and an apparent K_M was calculated by fitting data to a first-order rate equation as indicated by black lines. Each data point represents the mean of three replicates and error bars represent standard deviation.

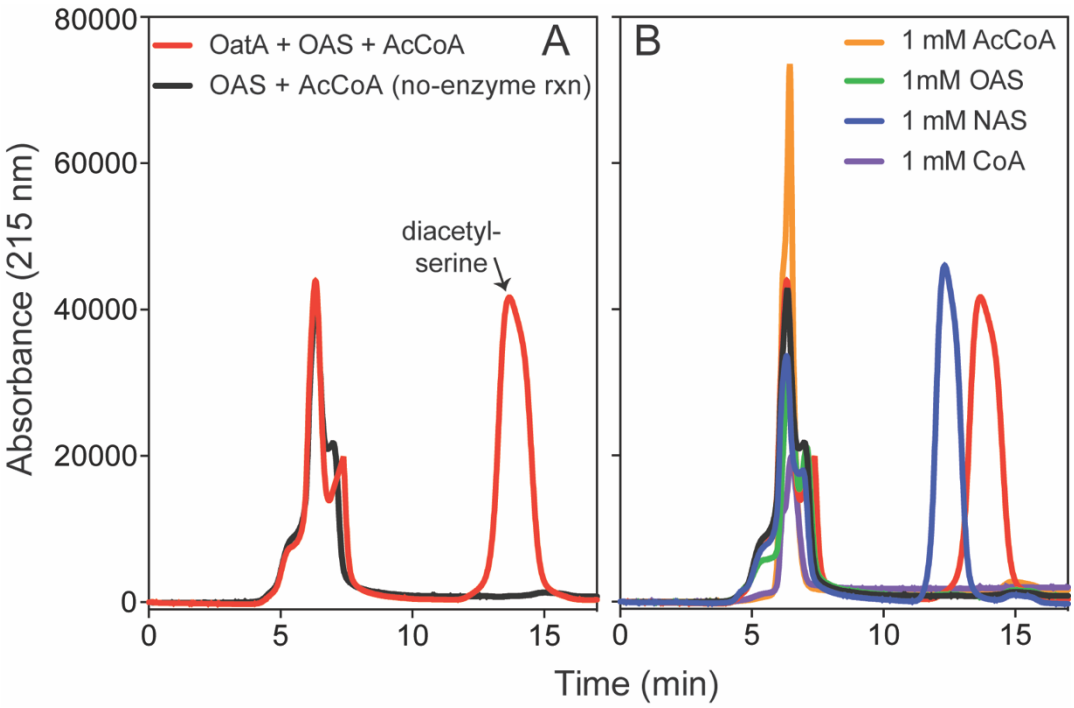


Figure 7.5. HPLC traces of OatA reaction products. A) Reactions containing ± OatA (80 μM), *O*-acetyl-serine (OAS, 1 mM), and acetyl-CoA (Ac-CoA, 1.2 mM) were incubated for 1 h at 37 °C and OatA was removed from reaction mixture, which was then injected (100 μL) onto a Aminex HPX-87H column (BioRad) as described in Materials and Methods. Separation of desired product (diacetyl-serine, retention time = 13.9 min, red peak) was monitored using an absorbance of 215 nm **B)** Proper separation of diacetyl-serine was compared to known standards: acetyl-CoA (retention time = 6.5 min), CoA (retention time = 6.5 min), *O*-acetyl-serine (OAS, retention time = 6.5 min), and *N*-acetyl-serine (NAS, retention time = 12.3 min). All standards were resuspsended for a final concentration of 1 mM. Legends in panel A also apply to panel B. Experiment was repeated in three independent experiments.

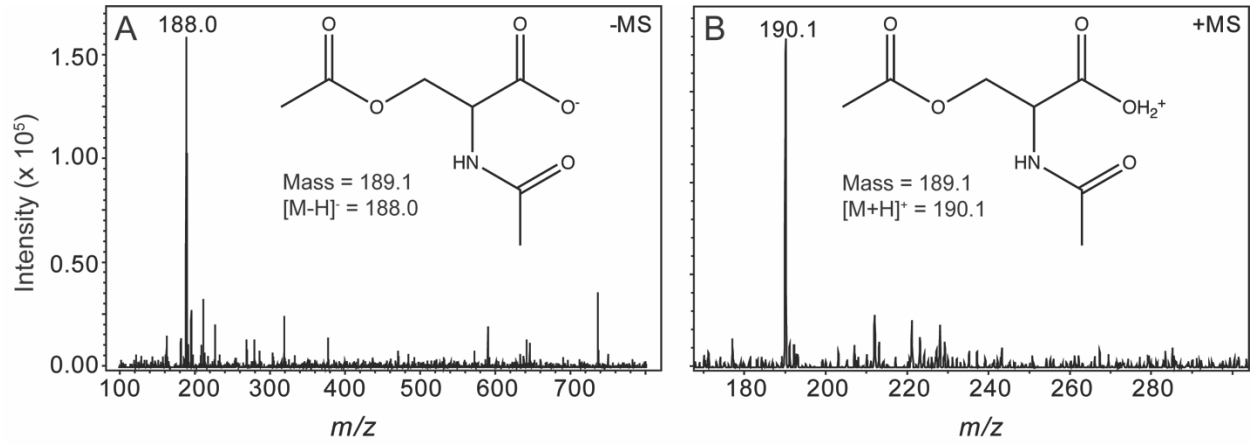


Figure 7.6. Mass spectrometry analysis of OatA reaction product, diacetyl-serine. Mass spectrometry was performed from product collected from HPLC separation (red line, peak indicated with arrow, Fig. 5). Unfragmented diacetyl-serine has a mass of 189.1 Da. **A.** $[MS]^-$ ionization of OatA reaction product displayed a *m/z* of 188.0 Da, which was indicative of a negatively charged diacetyl-serine molecule (inset). **B.** $[MS]^+$ ionization of OatA reaction product displayed a *m/z* of 190.1 Da, which was indicative of a postively charged diacetyl-serine molecule (inset). Experiment was repeated in triplicate.

OatA is a monomer in solution and binds acetyl-CoA before binding *O*-acetyl-L-serine.

To determine the oligomeric state of OatA, purified protein was analyzed by FPLC gel filtration using commercially available molecular mass standards. Under the conditions tested, OatA eluted at a retention time consistent with that of a protein whose mass was approximately 19.51 kDa (Fig. 7.7). Since the calculated molecular mass of OatA is 18.39 kDa, it was inferred that OatA was a monomer in solution. Dissociation constants of OatA for acetyl-CoA and OAS were determined using isothermal calorimetry. It was found that OatA had a K_D for acetyl-CoA of 21.84 μM with an n-value corresponding to 1 molecule of acetyl-CoA binding per dimer (Fig. 7.8). OatA did not bind *O*-acetyl-L-serine under conditions tested perhaps eluding to a binding order for OatA substrates.

Overexpression of *oatA* in *S. enterica* reduces lag phase of cells growing under sulfur limiting conditions.

To probe for OatA activity *in vivo*, a plasmid encoding an arabinose-inducible promoter linked to *oatA* was introduced into a strain lacking *oatA* on the chromosome. Overnight cultures were washed to rid cells of excess sulfur sources and inoculated into minimal medium containing $MgCl_2$ in lieu of $MgSO_4$ (Fig. 7.9A). It was found when sulfate was limiting, cells expressing *oatA* had a lag phase of 5 h, compared to a strain lacking *oatA* which had a lag phase of 10 h (Fig. 7.9C). These differences in lag phase were completely abolished when cysteine was added exogenously to the medium in lieu of $MgSO_4$ (Fig. 7.9B). It was determined that the decreased lag phase associated with the overexpression of *oatA* was most likely linked to cysteine biosynthesis, because addition of other sulfur compounds (*i.e.* methionine or glutathione) did not change *oatA* growth patterns (Fig. 7.10A-C).

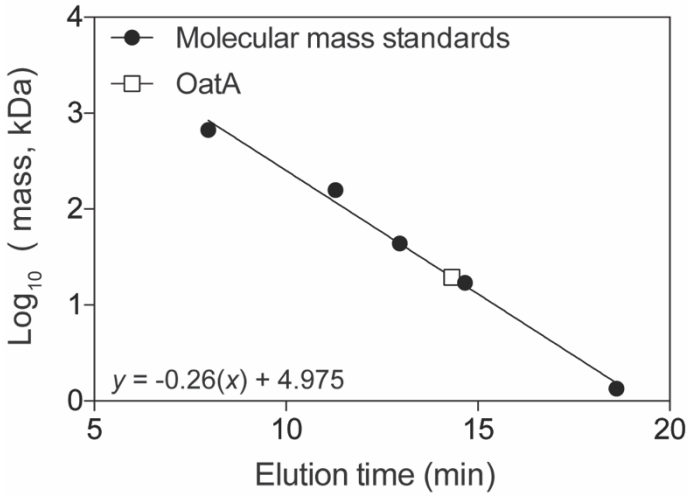


Figure 7.7. Determination of OatA oligomeric state using size exclusion chromatography. The molecular mass of OatA was determined by using a Superose 12 10/300 GL size exclusion chromatography as described in Materials and Methods. Moelcular mass standards (black circles) were purchased from BioRad with the following masses: thyroglobulin (bovine; 670,00 Da), gamma globulin (bovine; 158,000 Da), ovalbumin (chicken; 44,000 Da), myoglobin (horse; 17,000 Da), and vitamin B_{12} (1,350 Da). OatA represented as white square and molecular mass (19,500 Da) was calculated using equation displayed on graph ($R^2 = 0.9928$). Predicted molecular mass for OatA is 18,329 Da.

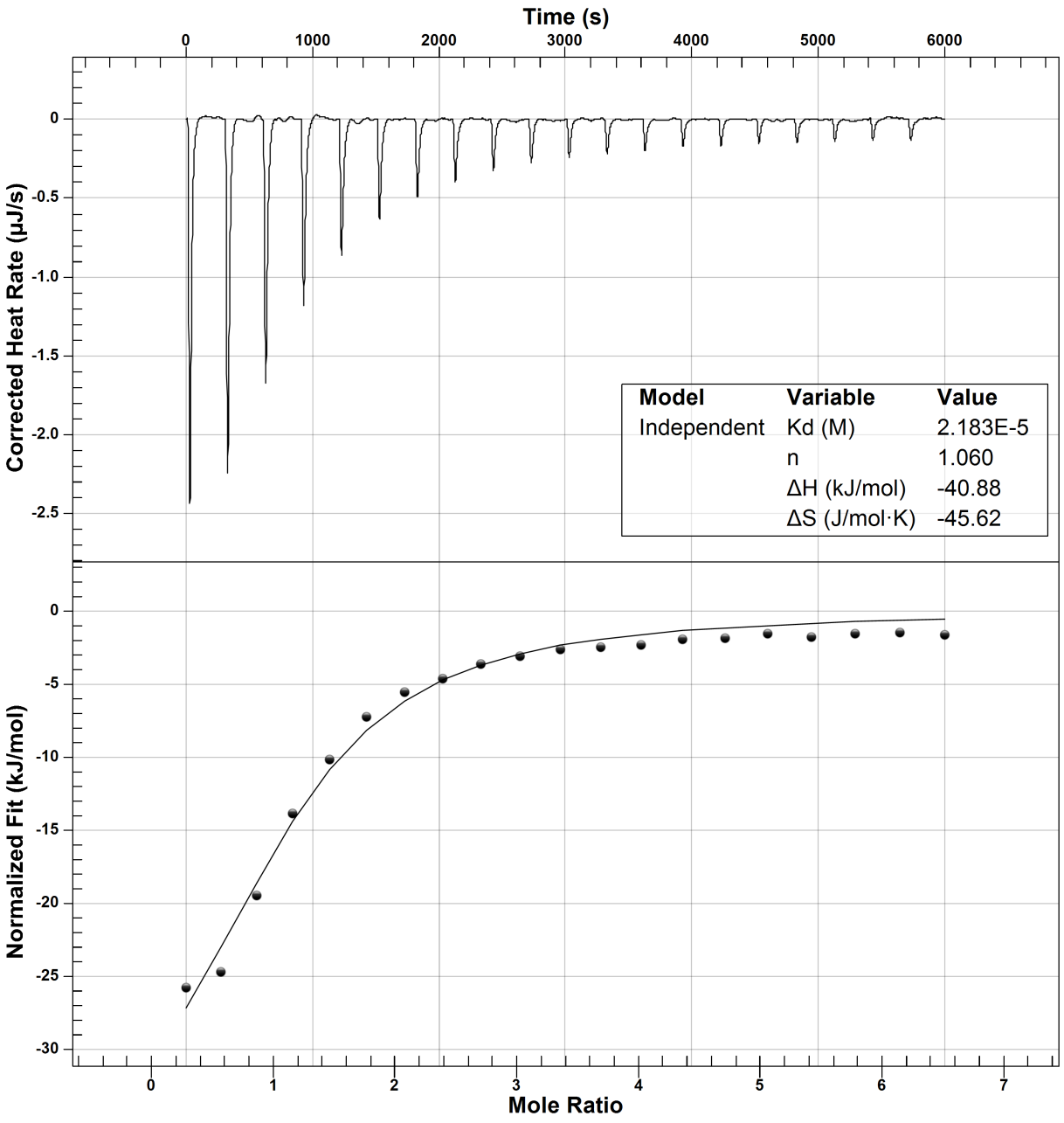


Figure 7.8. Isothermal Titration Calorimetry analysis of OatA binding to acetyl-CoA. ITC was used to determine binding of acetyl-CoA to OatA. Top panel is the raw data of heat release from 20 consecutive 2.48 µL injections of acetyl-CoA (1 mM) into a 170 µL cell volume containing OatA (50 µM). Bottom panel represents binding isotherms obtained by integrating the areas under the curve of each injection peak. Data were acquired using a NanoITC (TA instruments) and data were analyzed using NanoAnalyze 1.2 software. Molar ratio of acetyl-CoA to OatA was 1:1 (represented by n) and dissociation constant (K_d) was found to be 2.18 x 10^{-5} M.

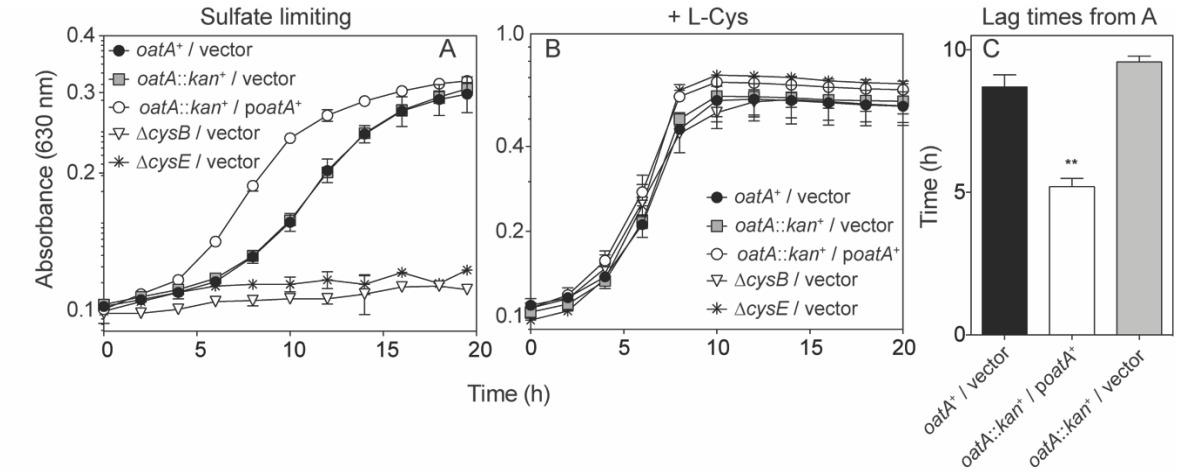


Figure 7.9. Growth analysis of cells overexpressing *oatA* under sulfur-limiting and exogenously provided cysteine conditions. The gene coding for *oatA* was cloned into a complementation vector and introduced into an *oatA*::*kan*$^+$ background. The same empty vector was introduced into strains as indicated in symbol legends. Cells from overnight cultures (lysogeny broth) were washed twice and grown on minimal medium with glycerol (22 mM), chloramphenicol, and transcription of plasmids was induced by addition of arabinose (1 mM). Growth curves were obtained using a microplate reader (BioTek Instruments). Error bars represent standard deviation and each experiment was completed in three independent experiments with strains in triplicate, with a representative graph shown. **A.** Growth of strains under sulfur limiting conditions ($MgCl_2$ added in lieu of $MgSO_4$, 1mM). **B.** Lag times of strains listed on *x*-axis, calculated from growth analysis of cells in panel A. **C.** Growth of strains under sulfur limiting conditions ($MgCl_2$ added in lieu of $MgSO_4$, 1mM) with exogenous L-cysteine (200 μM).

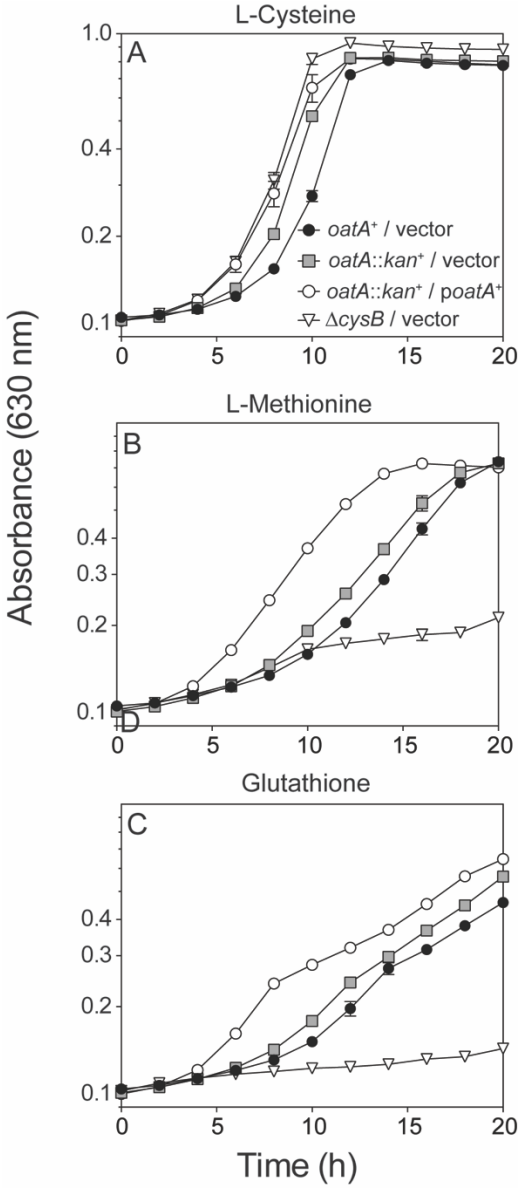


Figure 7.10. Growth analysis of cells overexpressing *oatA* after addition of various sulfur sources. The gene coding for *oatA* was cloned into a complementation vector and introduced into an *oatA*::*kan*$^+$ background. The same empty vector was introduced into strains as indicated in symbol legends. Cells from overnight cultures (lysogeny broth) were washed twice and grown on minimal medium with glycerol (22 mM), chloramphenicol, and transcription of plasmids was induced by addition of arabinose (1 mM). Growth of strains was assessed under sulfate-limiting conditions ($MgCl_2$ added in lieu of $MgSO_4$, 1mM). Growth curves were obtained using a microplate reader (BioTek Instruments). Error bars represent standard deviation and each experiment was completed in three independent experiments with strains in triplicate, with a representative graph shown. **A.** Growth after addition of exogenous L-cysteine (200 μM), **B.** L-methionine (300 μM), and **C.** oxidized glutathione (2.5 mM).

Overexpression of *oatA* leads to increased transcription of *cys* genes.

O-acetyl-L-serine is known to be a product of CysE, a serine transacetylase. This product along with sulfide is a substrate for CysK or CysM, both *O*-acetyl-serine (thiol)-lyases, that are involved in the last step of cysteine biosynthesis. Additionally, *O*-acetyl-serine has been predicted to migrate to *N*-acetyl-serine at a rate of 1% min^{-1}, serving as an effector for CysB, leading to transcriptional activation of genes needed to synthesize cysteine. This information was utilized to further analyze the effects of OatA and DAS on cysteine biosynthesis *in vivo*.

In order to determine if the decreased lag phase of cells overexpressing *oatA* was due to increased effector for CysB and subsequent activation of *cys* genes, RT-qPCR was performed under sulfur limiting conditions from above. Transcription was increased for nearly all genes that are positively regulated by CysB upon binding of *O*-acetyl-serine (Fig. 7.11). Transcription was increased for *cysB* (5-fold), *cysK* (5-fold), *cysD* (3-fold), and *cysI* (3-fold). Transcription was not affected with relation to *cysM* (Fig. 7.11).

Diacetyl serine doesn't change CysB binding patterns *in vitro*.

In order to determine effects of diacetyl-serine on CysB binding *in vitro*, we purified *cysJIH* and *cysK* promoter fragments containing a 5'-6FAM fluorescent label and incubated CysB with or without the diacetyl product of OatA. The promoters of *cysK* and *cysJIH* were chosen due to their gene changes in transcription (Fig. 7.11), and because CysB binding to these promoters has been extensively characterized. OatA reactions were set up as described in materials and methods, and a negative control included incubation of OatA without ac-CoA. OatA was removed from reaction mixtures and OatA products were added to EMSA reactions containing CysB and promoter regions. It was found that binding of CysB to *cysJIH* was altered upon incubation with diacetyl-serine, although the affects were small and therefore determined to be insignificant (data

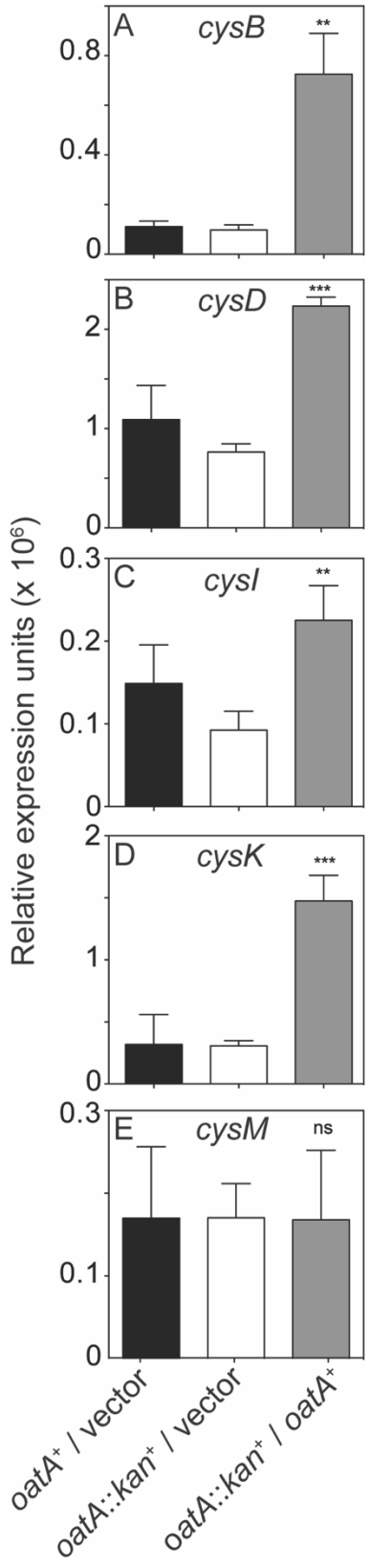


Figure 7.11. Effects of *oatA* overexpression on transcription of cysteine regulon genes. Total RNA was extracted from cells grown on glycerol (22 mM) minimal medium with $MgCl_2$ (1 mM) in lieu of $MgSO_4$. Cells were harvested in mid-log, where *oatA* overexpression caused a growth phenotype, as shown in Figure 9. Expression of *cysB* (**A.**), *cysD* (**B.**), *cysI* (**C.**), *cysK* (**D.**), and *cysM* (**E.**) was assessed using RT-qPCR and transcripts were normalized to the *gyrB* housekeeping gene of *S. enterica*. Changes in transcription when *oatA* was overexpressed (induction with 1 mM L-arabinose) were observed *cysB* (5-fold), *cysK* (5-fold), *cysD* (3-fold), and *cysI* (3-fold). Transcription was not affected with relation to *cysM*. Experiment was completed in three independent experiments, with RNA harvested in biological triplicate for each experiment and RT-qPCR completed in technical triplicate for each RNA replicate. Values shown are averages of biological triplicates with error bars representing standard deviation. P-values were calculated and represented by ** < 0.005 and *** < 0.0005

not shown). Binding of the *cysK* promoter was not altered upon binding of diacetyl-serine. Reasons for these observations are discussed below. Isothermal titration calorimetry was attempted, but was not fruitful due to solubility issues with CysB, which has been previously seen (21).

DISCUSSION.

OatA (previously YjgM) is a bona fide *O*-acetyl-serine *N*-acetyltransferase that produces diacetyl-serine.

Through specific activity (Fig. 7.3), kinetics (Fig. 7.4) and mass spectrometry (Fig. 7.6), we have shown that OatA acetylates the free amine group on the serine derivative, *O*-acetyl-serine, which leads to the production of diacetyl-serine (Fig. 7.2). This activity is specific for OAS, because OatA is unable to acetylate NAS or L-serine *in vitro* (Fig. 7.3). This study is the first example of enzymatic production of diacetyl-serine.

Overexpression of *oatA* is beneficial for *S. enterica* growth on sulfate-limiting medium and is due to increased expression of the cysteine regulon.

Sulfate binding to the LysR regulator, CysB, causes repression of the cysteine regulon. Under sulfate-limiting conditions, it is hypothesized that CysB binds to NAS and induces expression of the cysteine regulon. Previous studies have shown that *O*-acetyl-serine (an intermediate of cysteine biosynthesis) undergoes a spontaneous non-enzymatic *O*- to *N*- migration that produces NAS. It was hypothesized that this occurs *in vivo*, leading to CysB de-repression. We utilized these CysB dependent growth conditions to determine if OatA production of diacetyl-serine was beneficial for cysteine biosynthesis. Under sulfate-limiting conditions, overexpression of *oatA* decreased the lag time in half (Fig. 7.9) compared to cells not overexpressing *oatA*. Addition of cysteine to the sulfate-limiting medium abolished the beneficial growth phenotypes associated with overexpression of *oatA*. This was specific to cysteine, because addition of other

sulfur sources, such as methionine or glutathione, did not have the same affect (Fig. 7.10). Upon completion of RT-qPCR analyses from RNA extracted under the same conditions, it was determined that overexpression (and therefore presence of diacetyl-serine) lead to increased expression of *cysB, cysD, cysI,* and *cysK* (Fig. 7.11). What we know from these studies is OatA mediated production of diacetyl-serine is beneficial during cysteine biosynthesis. However, due to solubility issues with $CysB^{EBD}$ and the inability to show differences in CysB DNA binding *in vitro*, we are unsure if the effects of diacetyl-serine are direct or indirect. Due to the complex binding patterns of CysB and other LysR regulators, we cannot rule out that binding of CysB to diacetyl-serine causes changes in DNA bending or binding, that cannot be visualized via electrophoretic mobility shift assays. Due to the fact that diacetyl-serine cannot be purchased caused limitations to the amounts of diacetyl-serine that we were able to produce enzymatically, and may not have been a high enough concentration to detect changes in CysB DNA-binding.

What is the physiological role of OatA?

Deletion of *oatA* is not detrimental to *S. enterica* growth. Therefore, we would argue that there are other means of accumulating NAS or DAS *in vivo*. It is possible that OatA serves as a sensor of stress, and when additional cysteine biosynthesis is required, OatA is able to fine tune expression of the cysteine regulon. Previous work has shown that OatA (previously YjgM), is upregulated during planktonic growth when compared to biofilm formation (22). It is possible that cysteine biosynthesis or sulfur metabolism is related to biofilm formation and that expression of *oatA* is able to modulate cellular activity under these conditions. More work is needed to determine the direct or indirect involvement of OatA in cysteine biosynthesis. Regardless, this study characterizes a putative acetyltransferase in *Salmonella enterica* and further elucidates the mechanistic importance of acetylation in cellular physiology.

MATERIALS AND METHODS.

Bacterial Strains, culture media and chemicals.

All bacterial strains are listed in Table 7.1. All strains for growth analysis were derivatives of *Salmonella enterica* subsp. *enterica* serovar Typhimurium LT2 (hereafter *S. enterica)* and constructed using the Wanner in-frame gene deletion method (23). *S. enterica* strains were grown at 37°C on lysogeny broth (LB, Difco) or no-carbon essential (NCE) minimal medium (24) supplemented with glycerol (22 mM), $MgSO_4$ (concentrations varied and noted in figure legends), and Wolfe's trace minerals (1x) (25). *Escherichia coli* C41 (λDE3) (26) was used for protein overexpression and *E. coli* DH5α (New England Biolabs) was used for plasmid construction. All *E. coli* strains were grown at 37°C in LB. Antibiotics for all media were used at the following concentrations: ampicillin, 100 μg mL^{-1}; chloramphenicol, 20 μg mL^{-1}. All chemicals were purchased from Sigma-Aldrich unless otherwise noted; isopropyl β-D-1-thiogalactopyranpside (IPTG, Gold BioTechnology), glycerol (Fisher Scientific), [1-^{14}C] acetyl-CoA (Moravek, Inc.), HEPES (Gold BioTechnology). *O*-acetyl-L-serine and L-cysteine was made fresh in HEPES (pH 7 at 25°C) immediately before each *in vitro* or *in vivo* experiment.

Plasmid construction.

All plasmids used in this work are listed in Table 7.2. Primers were synthesized by Integrated DNA Technoligies, Inc. (IDT [Coralville, IA]) and are listed in Table 7.3. Genes were amplified from *S. enterica* genomic DNA using Pfu Ultra II fusion DNA polymerase (Stratagene) per manufactures instructions. A high efficiency cloning method described elsewhere (27, 28) was utilized to clone *oatA* and *cysB* genes into pCV3 (complementation vector, L-(+)-arabinose inducible) and pTEV16 or pTEV18 (overexpression vectors, IPTG inducible). BspQI was

Table 7.1. Bacterial strains used in this study		
Strain	**Relative Genotype**	**Source[1]**
E. coli strains		
E. coli DH5α	Φ80d*lacZ*ΔM15 *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 (r_k-, m_k^+) *supE*44 *relA*1 *deoR* Δ(*lacZYA-argF)* U169 *phoA*	NEB
E. coli C41 (λDE3)	*pka12::kan*$^+$ *ompT hsdS* (r_Bm_B) *gal* (λDE3*)*	Laboratory collection
S. enterica strains		
JE10079	*ara-9*	
Derivatives of JE10079		
JE23873	*oatA*$^+$ / pCV3	
JE23874	*oatA*::*kan*$^+$ / pCV3	
JE23875	*oatA*::*kan*$^+$ / p*oatA*$^+$	
JE19460	Δ*cysE* / pCV3	
JE24366	*cysB*::*kan*$^+$ / pCV3	

[1]Unless otherwise noted, all strains were constructed in this study

Table 7.2. Plasmids used in this study		
Plasmid	**Genotype**	**Source**[1]
pTEV16	*lacI*$^+$ *bla*$^+$	(28)
pTEV20	*lacI*$^+$ *bla*$^+$	(28)
pCV3	*araC*$^+$ *cat*$^+$	(28)
Overexpression plasmids		
pCysB2	*S. enterica cysB*$^+$ cloned into pTEV20, *bla*$^+$	
pOatA4	*S. enterica oatA*$^+$ cloned into pTEV16, *bla*$^+$	
Complementation plasmids		
pOatA6	*S. enterica oatA*$^+$ cloned into pCV3, *cat*$^+$	

[1]Unless otherwise noted, all plasmids and strains were constructed in this study

Table 7.3. Primers used in this study	
Primer Name	**Primer Sequence[3] 5' → 3'**
Overexpression primers	
5' *oatA* pTEV16	NN**GCTCTTC**NAGCATGAATAATG TCGCCTCGCCAACGC
3' *oatA* pTEV16	NN**GCTCTTC**NTTATCAGAGATCCTTTAACATCCTCACTTCG
5' *cysB* pTEV20	NN**GCTCTTC**NTTCATGAAATTGCAGCAGCTTCGTTAC
3' *cysB* pTEV20	NN**GCTCTTC**NTTATTACTTTTCAGGCAGCTTTATATCCTG
Complementation primers	
5' *oatA* pCV3	NN**GCTCTTC**NTTCATGAATAATG TCGCCTCGCCAACGC
3' *oatA* pCV3	NN**GCTCTTC**NTTATCAGAGATCCTTTAACATCCTCACTTCG
***S. enterica* wanner deletion primers**	
5' *cysB* wanner *kan*$^+$	ACGATGGCCTGATGGCGCTAATCTGGATGATGTATT GTGTAGGCTGGAGCTGCTTC
3' *cysB* wanner *kan*$^+$	GTCTGCCATGCCACTACGACACAAACCGACGGTGAT ACATATGAATATCCTCCTTAG
5' *cysE* wanner *kan*$^+$	ACGGGTTGGTCGTTTTCTGCCCGTCTGGAGTAAGCC ATG GTGTAGGCTGGAGCTGCTTC
3' *cysE* wanner *kan*$^+$	CCATCGGAACAGCGTTTTTTAGTTGTACCGCGCAA TTCACATATGAATATCCTCCTTAG
5' *oatA* wanner *kan*$^+$	GTCAGCTTCCGGCGTGGCCGCGGATAACAAGAGA GAGTGTAGGCTGGAGCTGCTTC
3' *oatA* wanner *kan*$^+$	AGATGCCTTCATCGAGTAGTTGGATATGTCCAGCT ACATATGAATATCCTCCTTAG
Electrophoretic Mobility Shift Assay Primers	
5' *cysJIH*	TCCAACCCTTCTTTAATTGTTATTCCTC
3' *cysJIH*	/56-FAM/CGTCATGCGTCGTTATGTTCCAG
RT-qPCR primers	
5' *cysB*	AAGCGGTATCGAAAGGGAAC
3' *cysB*	TAGCATGGCAACATCACCAG
5' *cysD*	GCTTCCGTACACTTGGCTGT
3' *cysD*	CTGATCGCGGTCAATCATC
5' *cysI*	CCGCCCGTATGAGTTTACC
3' *cysI*	AGCGTCAGGTGCCATTTATC
5' *cysK*	TGCGGTAGAACCCACTGACT
3' *cysK*	GGCCTGGTTTGATCTCTTCA
5' *cysM*	TACGCCCACTACACGACCAC
3' *cysM*	GAATGCCCGGAATACTGCT

purchased from New England Biolabs. The resulting complementation vector was referred to as pOatA6. The resulting overexpression vectors were referred to as pOatA4 and pCysB2 (full length CysB).

Strain construction and growth analysis.

In frame deletions of *S. enterica* genes were constructed using the phage lambda Red recombinase system as described elsewhere (18, 23). Plasmids were transformed into strains for complementation studies as described elsewhere (18). For growth analyses, starter cultures were grown overnight at 37°C shaking in LB containing appropriate antibiotic. Cultures (1% [vol / vol]) where used to inoculate 198 μL of fresh minimal medium placed in each well of a 96-well microtiter plate. When necessary before inoculation, cells were washed (x 2) with NCE (1x) by centrifuging 1 mL of culture at 6,000 X *g* for 3 min. Figure legends indicate if inocula were larger than 1% (vol / vol). L-(+)-Arabinose was used as an inducer when indicated. Under sulfur limiting conditions, $MgCl_2$ (1 mM) was used instead of $MgSO_4$. The following chemicals had varying concentrations depending on experiment and are indicated in figures or figure legends: *O*-acetyl-L-serine, L-cysteine, and $MgSO_4$. Microtiter plates were incubated at 37°C inside the temperature-controlled chamber of a PowerWave microtiter plate reader (Bio-Tek Instruments), and plates were continuously shaken using the slow setting of the instrument. Time points were taken for 24 h every 15 min with data shown for every 2 h. Cell density was monitored at 630 nm and data were analyzed using Prism v6 (GraphPad). Growth studies were performed in triplicate in three independent experiments, with a representative growth curve shown with standard deviation of technical triplicates.

Protein overproduction and purification.

Plasmids coding for proteins of interest were transformed into *E. coli* C41 (λDE3). Overnight cultures (50 mL) of transformants were subcultured (2% [vol / vol]) into 1 L of LB containing ampicillin. Cultures were grown at 25 °C with shaking to an optical density of 0.6 (OD_{600nm}) after which plasmid expression was induced by addition of IPTG (0.5 mM) and shaken overnight. Cells were harvested by centrifugation at 6,000 X *g* for 10 min in a Becman Coulter Avanti J-20 XOI refrigerated centrifuge with a JLA-8.1000 rotor at 4°C. Cell pellets were stored at -80°C until used. For purification, cell pellets were thawed and re-suspended in 30 mL of buffer A (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HEPES [50 mM, pH 7.0 at 4°C], NaCl [500 mM], and Imidizole [20 mM]) containing lysozyme (1 mg/mL) DNase (1μg/mL), and protease inhibitor phenylmethanesulfonyl fluoride (PMSF, 0.5 mM). Cells were lysed by two rounds of sonication on ice (1 min total (2s on 2s off, 60% duty) using a Sonicator. Clarified lysates were obtained after centrifugation for 30 min at 4°C at 40,000 X *g* in a Beckman Coulter Avanti J-25I centrifuge with JA-25.50 rotor followed by filtration of supernatant through a 0.45 μm filter (Millipore). Samples were applied at 4°C to a 1 mL HisPur nickel-nitrilotriacetic acid (Ni-NTA) resin (Thermo Scientific) that was prequilibrated with 10 column volumes (CV) of water and 10 CV of buffer A. After lysate was applied to the column, it was washed with 10 CV of buffer A, 5 CV of buffer B (HEPES [50 mM, pH 7.0 at 4°C], NaCl [500 mM], Imidizole [60 mM]). Proteins were eluted with 5 CV of buffer C (HEPES [50 mM, pH 7.0 at 4°C], NaCl [500 mM], Imidizole [500 mM]). Fractions were run on an SDS-PAGE gel and fractions containing the desired protein were combined and dialyzed for 3 h at 4°C in HEPES (50 mM, pH 7.0) with decreasing concentrations of NaCl (400 mM, 200 mM, and 150 mM) with a final buffer composition of HEPES (50 mM, pH 7.0 at 4°C), NaCl (150 mM), and glycerol (20% v/v). Final protein

concentration was determined using the extinction coefficient and molecular weight (ExPASy, ProtParam) of each protein with a NanoDrop. Proteins were drop-frozen in liquid nitrogen and stored at -80°C.

***In vitro* acetylation spectrophotometric assays.**

To determine specific activity or kinetic parameters of OatA for L-serine, *N*-acetyl-DL-serine, or *O*-acetyl-L-serine a continuous spectrophotometric assay that employed 5,5'-dithiobus-(2-nitrobenzoic acis) (DTNB, Ellman's reagent) was utilized as described elsewhere (15). Briefly, assays were performed at 25°C in 100 μL reaction volumes in 96-well plates and reactions were monitored at 412 nm. Reaction mixtures contained HEPES buffer (50 mM, pH 7.2), DTNB (0.3 mM), acetyl-CoA (100 μM), protein (1 μg) and substrate (200 μM). To obtain data for enzyme efficiency (k_{cat} / K_m), acetyl-CoA was added at saturating levels (1 mM) with sweeps of *O*-acetyl-L-serine (0-5 mM). Data were acquired using the Soft Max Pro 6.2 Spectramax software every 10 s over 5 min. Path lengths were calculated using Soft Max Pro 6.2 endpoint readings and absorbance data were corrected for path lengths of 1 cm.

Specific activity was calculated from the slope of the linear range of (ΔA_{415} s^{-1}) using Beer's Law ($A = [\varepsilon]lc$) with a path length of 1 cm and the molar extinction coefficient of the TNB^{2-} thiolate anion (14,150 M^{-1} cm^{-1}) (29). Equation was solved for *c*, giving specific activity in μmol min^{-1} mg^{-1} of YjgM.

For calculating kinetic parameters, graphs of initial velocity (μM s^{-1}) versus substrate concentration (μM) were plotted using Prism v6 (GraphPad). Data were fitted to the Michaelis-Menton kinetics model to determine K_m and V_{max}. The turnover number (k_{cat}) was determined using the following equation: $V_{max} = k_{cat}[E]$, where [*E*] was the concentration of protein added. All spectrophotometric assays mentioned above were completed thrice, each in technical triplicate

with a representative data set shown. Error bars represent standard deviation as calculated by Prism v6 (GraphPad).

Size exclusion chromatography.

A Superose 12 10/300 GL gel filtration column (GE Healthcare Life Sciences) was equilibrated as per manufacturer's protocol using water and buffer (50 mM HEPES, pH 7.0 @ 4°C, 150 mM NaCl). Samples were applied to column using a 100-μL superloop. Gel filtration standards (Bio-Rad) were applied to column first and a standard curve was calculated using the MW_{log} of each standard against retention time of each protein. Purified OatA was eluted from the column in the same manner and retention times were recorded and molecular weights were determined using equation calculated from the standard curve.

HPLC purification of diacetyl-L-serine.

Reactions containing OatA (80 μM), acetyl-CoA (1.2 mM), *O*-acetyl-serine (1 mM) and sodium phosphate buffer (pH 7.0 at 25° C) were incubated for 1 h at 37°C. OatA was removed by passing reactions over an Amicon Ultra 0.5 mL 10kDa molecular cut off centrifugal filter (Millipore). The product of OatA OAS acetylation was resolved by RP-HPLC using a Shimadzu Prominence UFLC with a Aminex HPX-87H column (BioRad). The column was equilibrated at a flow rate of 0.5 ml min^{-1} with 5 column volumes of buffer [5 mM H_2SO_4 in water]. Samples (100 μl) were injected and products were resolved isocratically. Compounds were detected and compared to standards at 215 nm using a computer-controlled Shimadzu Nexera X2 SPD-30A diode array detector. Data were analyzed using Prism v6 (GraphPad). Fractions were collected in 0.5 mL increments and fractions containing peaks corresponding to diacetyl-serine were analyzed by MS (Protein and Mass Spectrometry Facility, UGA, Athens, GA, USA). Electrospray ionization

(ESI)-MS was performed in methanol and resolved on an Esquire 3000 Plus (Bruker) Ion Trap Mass Spectrometer at 0.3 ml h^{-1}.

Growth of cultures for RNA isolation.

For cultures grown under ample sulfur conditions, ($MgSO_4$ 1 mM), five individual colonies were inoculated into five independent tubes containing LB + Cm (2 mL) and grown overnight at 37°C. Stationary phase ODs were calculated by diluting (1:10) and measuring absorbance (OD_{600}). Cells were then inoculated into minimal medium (10 mL, conditions described above) with a starting OD of 0.1. Cultures were grown for 6 hours, spun down in conical tubes (15 mL), flash frozen and stored at -80°C until use.

For cultures grown under sulfur limiting conditions, cells were grown as just described with the following exceptions. Before calculating stationary phase OD, cells were washed (x 2) with NCE (1 x). Cells were inoculated into minimal medium (two 10 mL cultures) with $MgCl_2$ (1 mM) in lieu of $MgSO_4$.

RNA isolation.

Cell pellets were thawed and RNA was isolated using a protocol described previously (cite Kushner RNASnap). In detail, cells were re-suspended (150 μL) in boil solution [EDTA (18 mM), SDS (0.025% v / v), 2-mercaptoethanol (1% v / v), formamide (95% v / v), RNase-free water (up to final volume)] and heated (95°C) for 7 min. Suspensions were spun down (16,000 *x* g) for 5 min at room temperature and supernatant (100 μL) was transferred to a new tube. Solution was diluted with RNase-free water (400 μL) and sodium acetate (50 μL, final concentration 0.3 M). Ice-cold ethanol (95% v /v, 1650 μL) was added and solutions were incubated at -80°C for 1 h. Solutions were centrifuged at 16,000 *x* g for 30 min at 4°C and supernatant was poured off. RNA/DNA pellets were washed with 300 μL cold ethanol (70 %, v / v) and centrifuged at 8,000 *x* g for 5 min.

Supernatant was poured off, tubes were centrifuged briefly and liquid was aspirated. Pellets were air-dried upside down on Kim-Wipes for 30 min and re-suspended in 100 μL RNase-free water. Tubes were centrifuged at 16,000 *x* g for 1 min (to rid of water insoluble material) and 90 μL was transferred to a new tube for subsequent DNase digestion. DNA was digested using TURBO DNase (Thermo Fischer Scientific) following the protocol for rigorous DNase treatment. After DNase treatment and DNase inactivation, mixtures were centrifuged at 10,000 *x* g for 1 min and 100 μL of supernantent was transferred to a fresh tube. RNA was sodium acetate / ethanol precipitated as described above. RNA was re-suspended in 100 μL, aliquoted into 4 tubes and frozen at -80°C for later use. RNA quality and concentration was assessed using the Agilent 2100 Bioanalyzer with the RNA 6000 nano kit (Georgia Genomics Facility, University of Georgia). Samples with RIN values over 6 were considered adequate for subsequent RT-qPCR experiments.

cDNA synthesis and quantitative reverse transcription PCR.

RNA samples were diluted in RNase free water (final concentration 40 ng/μL) and used as template (400 ng RNA final) to synthesize cDNA using the iScript cDNA synthesis kit (Bio-Rad Laboratories) per manufactures protocol. Synthesized cDNA was diluted to 7.5 ng/μL to be used as a template for PCR. For real-time PCR, 20 μL reaction mixtures were prepared in 96-well plates with 10 μL of FastSYBR green master mix (Applied Biosystems), gene-specific primers (500 nM), and cDNA (15 ng or 2 μL of above diluted cDNA). The RT-qPCR was performed using a 7500 Fast real-time PCR system (Applied Biosystems). The *S. entercia gyrB* gene was used as an internal control and gene-specific primers were designed using Primer 3 software. Each biological replicate (3 for each strain) was analyzed in technical triplicate. Cycle threshold (C_T) data were normalized to *gyrB* and these values were transformed using $2(e-\Delta C_T)/10^{-6}$ (30) and reported as arbitrary gene expression units (EU). Mean EU values of each technical triplicate were used to

calculate the standard error of the mean (SEM) of three biological replicates using the Prism v6 software. Differences between *yjgM*$^{+}$ and *yjgM*::*kan*$^{+}$ / pYjgM^{WT} or *yjgM*::*kan*$^{+}$ / pCV3 and *yjgM*::*kan*$^{+}$ / pYjgM^{WT} were compared using Welch's *t* test with GraphPad Prism v6 software.

Isothermal titration calorimetry (ITC).

Proteins (OatA) were purified as described above and concentrated to at least 100 μM before dialyzing three times in filtered HEPES (50 mM, pH 7.0 @ 25°C). The final dialysis buffer was saved and used to clean and equilibrate the cell and syringe system, for re-suspension of ligands, and dilution of protein. OatA was diluted to 50 μM and added to the ITC cell following instruments instructions. Acetyl-CoA (1 mM) or *O*-acetyl-serine (0, 0.5, 1, 2, and 20 mM) was re-suspended in dialysis buffer and used as titrant (50 μL) in the stirring syringe. After cell and syringe were loaded, stirring (350 rpm) was initiated until the baseline was re-equilibrated. Titration was initiated using a program for 20 injections (2.48 μL each, 300 s apart) with continuous stirring (350 rpm). ITC data were analyzed using NanoAnalyze software (TA instruments). Each condition was repeated in triplicate with a representative data set shown in figures.

REFERENCES.

1. Kredich NM. 2008. Biosynthesis of Cysteine. EcoSal Plus 3:1-30.
2. Ostrowski J, Jagura-Burdzy G, Kredich NM. 1987. DNA sequences of the cysB regions of *Salmonella typhimurium* and *Escherichia coli*. J Biol Chem 262:5999-6005.
3. Jagura-Burdzy G, Kredich NM. 1983. Cloning and physical mapping of the cysB region of Salmonella typhimurium. J Bacteriol 155:578-85.
4. Hryniewicz MM, Kredich NM. 1994. Stoichiometry of binding of CysB to the *cysJIH*, *cysK*, and *cysP* promoter regions of *Salmonella typhimurium*. J Bacteriol 176:3673-3682.

5. Ostrowski J, Kredich NM. 1991. Negative autoregulation of *cysB* in *Salmonella typhimurium*: in vitro interactions of CysB protein with the *cysB* promoter. J Bacteriol 173:2212-2218.

6. Hryniewicz MM, Kredich NM. 1991. The *cysP* promoter of *Salmonella typhimurium*: characterization of two binding sites for CysB protein, studies of *in vivo* transcription initiation, and demonstration of the anti-inducer effects of thiosulfate. J Bacteriol 173:5876-5886.

7. Monroe RS, Ostrowski J, Hryniewicz MM, Kredich NM. 1990. In vitro interactions of CysB protein with the *cysK* and *cysJIH* promoter regions of *Salmonella typhimurium*. J Bacteriol 172:6919-6929.

8. Kredich NM. 1971. Regulation of L-cysteine biosynthesis in *Salmonella typhimurium*. I. Effects of growth of varying sulfur sources and *O*-acetyl-L-serine on gene expression. J Biol Chem 246:3474-3484.

9. Jones-Mortimer MC, Wheldrake JF, Pasternak CA. 1968. The control of sulphate reduction in *Escherichia coli* by *O*-acetyl-L-serine. Biochem J 107:51-53.

10. Ostrowski J, Kredich NM. 1989. Molecular characterization of the cysJIH promoters of *Salmonella typhimurium* and *Escherichia coli*: regulation by CysB protein and N-acetyl-L-serine. J Bacteriol 171:130-140.

11. Flavin M, Slaughter C. 1967. Enzymatic synthesis of homocysteine or methionine directly from O-succinyl-homoserine. Biochim Biophys Acta 132:400-405.

12. Hryniewicz MM, Kredich NM. 1995. Hydroxyl radical footprints and half-site arrangements of binding sites for the CysB transcriptional activator of *Salmonella typhimurium*. J Bacteriol 177:2343-2353.

13. Hentchel KL, Escalante-Semerena JC. 2015. Acylation of biomolecules in prokaryotes: a widespread strategy for the control of biological function and metabolic stress. Microbiol Mol Biol Rev 79:321-346.

14. VanDrisse CM, Hentchel KL, Escalante-Semerena JC. 2016. Identification of phosphinothricin acetyltransferases using in vivo, in vitro and bioinformatics analyses. Appl Environ Microbiol 82:7041-7051.

15. Hentchel KL, Escalante-Semerena JC. 2015. In *Salmonella enterica*, the Gcn5-related acetyltransferase MddA (formerly YncA) acetylates methionine sulfoximine and methionine sulfone, blocking their toxic effects. J Bacteriol 197:314-325.

16. Rycroft JA, Gollan B, Grabe GJ, Hall A, Cheverton AM, Larrouy-Maumus G, Hare SA, Helaine S. 2018. Activity of acetyltransferase toxins involved in *Salmonella persister f*ormation during macrophage infection. Nat Commun 9:1-11.

17. Cheverton AM, Gollan B, Przydacz M, Wong CT, Mylona A, Hare SA, Helaine S. 2016. A *Salmonella* toxin promotes persister formation through acetylation of tRNA. Mol Cell 63:86-96.

18. VanDrisse CM, Parks AR, Escalante-Semerena JC. 2017. A toxin involved in *Salmonella* persistence regulates its activity by acetylating its cognate antitoxin, a modification reversed by CobB sirtuin deacetylase. MBio 8.

19. Burckhardt RM, Escalante-Semerena JC. 2017. In *Bacillus subtilis*, the SatA (formerly YyaR) acetyltransferase detoxifies streptothricin via lysine acetylation. Appl Environ Microbiol doi:10.1128/AEM.01590-17.

20. Kuhn ML, Majorek KA, Minor W, Anderson WF. 2013. Broad-substrate screen as a tool to identify substrates for bacterial Gcn5-related N-acetyltransferases with unknown substrate specificity. Protein Sci 22:222-230.

21. Tyrrell R, Verschueren KH, Dodson EJ, Murshudov GN, Addy C, Wilkinson AJ. 1997. The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. Structure 5:1017-1032.

22. MacKenzie KD, Wang Y, Shivak DJ, Wong CS, Hoffman LJ, Lam S, Kroger C, Cameron AD, Townsend HG, Koster W, White AP. 2015. Bistable expression of CsgD in *Salmonella enterica* serovar Typhimurium connects virulence to persistence. Infect Immun 83:2312-2326.

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

24. Berkowitz D, Hushon JM, Whitfield HJ, Jr., Roth J, Ames BN. 1968. Procedure for identifying nonsense mutations. J Bacteriol 96:215-220.

25. Atlas R. 1995. Handbook of Media for Environmental Microbiology, p 6. CRC Press, Boca Raton

26. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289-298.

27. Galloway NR, Toutkoushian H, Nune M, Bose N, Momany C. 2013. Rapid cloning for protein crystallography using Type IIS restriction enzymes. Crystal Growth & Design 13:2833-2839.

28. VanDrisse CM, Escalante-Semerena JC. 2016. New high-cloning-efficiency vectors for complementation studies and recombinant protein overproduction in Escherichia coli and Salmonella enterica. Plasmid 86:1-6.

29. Riddles PW, Blakeley RL, Zerner B. 1983. Reassessment of Ellman's reagent. Methods Enzymol 91:49-60.

30. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402-408.

CHAPTER 8

ANALYSIS OF ACETYLATION TARGETS OF THE PROTEIN ACETYLTRANSFERASE ACT-PAT IN *STREPTOMYCES LIVIDANS*.[1]

[1*]VanDrisse C.M., *Tucker, A.C., and Escalante-Semerena J.C. 2018

To be submitted to *Proteomics*.

*Both authors contributed equally to this work.

ABSTRACT

Acetylation of lysine residues affects protein function in all domains of life and acetyltransferases target lysyl residues of substrate proteins. Gcn5-related *N*-acetyltransferases (GNATs) are the only family of acetyltransferases identified in bacteria. GNATs acetylate a diverse set of substrates, with the paradigm of acetylation being modification of the acetyl-CoA ligase, Acs. Previously we identified a Type III ACT-domain protein acetyltransferase, PatB, in *Streptomyces lividans* that acetylated Acs of this organism. Generally, protein acetyltransferases that target Acs are also able to acetylate other CoA ligases of the cell. Through proteomics studies, we identified new CoA ligases from *Streptomyces* that were acetylated *in vivo*. Upon purification of these CoA ligases, we found that they were acetylated by the ACT-domain *PatB* and that their activities were reduced similar to that of acetylated Acs. More work is needed to confirm the effects of acetylation on all the CoA ligases characterized.

INTRODUCTION

Reversible lysine acetylation is a post-translational modification that can affect enzyme activity, protein-protein interactions, protein-DNA interactions, and protein stability (1-3). Protein acetylation occurs in all domains of life (4). 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 (5). 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 (6). 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 (7-10). Gcn5-related *N*-acetyltransferases (GNATs) are the only acetyltransferases identified in all domains of life (11). Acetylation of the catalytic lysine of acetyl-CoA ligase (Acs) inactivates this enzyme (3). Acs is a member AMP-forming CoA ligase family and this family of enzymes activates organic acids to CoA thioesters via an acyl-AMP intermediate (12). 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 (8, 13-17, 18l, 19).

GNATs exhibit substrate specificity. The α-proteobacterium *Rhodopseudomonas palustris* produces a two-domain protein acetyltransferase (Pat) that can distinguish between members of the AMP-forming CoA ligase family (8). *Rp*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 (2, 11, 20). Together, these findings suggest that Pat homologs can recognize a diverse but specific set of substrates.

The Gram-positive actinomycete *Streptomyces lividans* encodes a Type I Pat homolog, namely PatA, in which the two domains are reversed relative to the Pat homologs in *E. coli*, *S. enterica*, and *R. palustris* (Fig. 8.1). *S. lividans* PatA contains an *N*-terminal GNAT domain linked to a large *C*-terminal domain of unknown function. PatA acetylates acetoacetyl-CoA ligase (AacS), an AMP-forming CoA ligase in *S. lividans* (18). The Acs homolog in the closely related

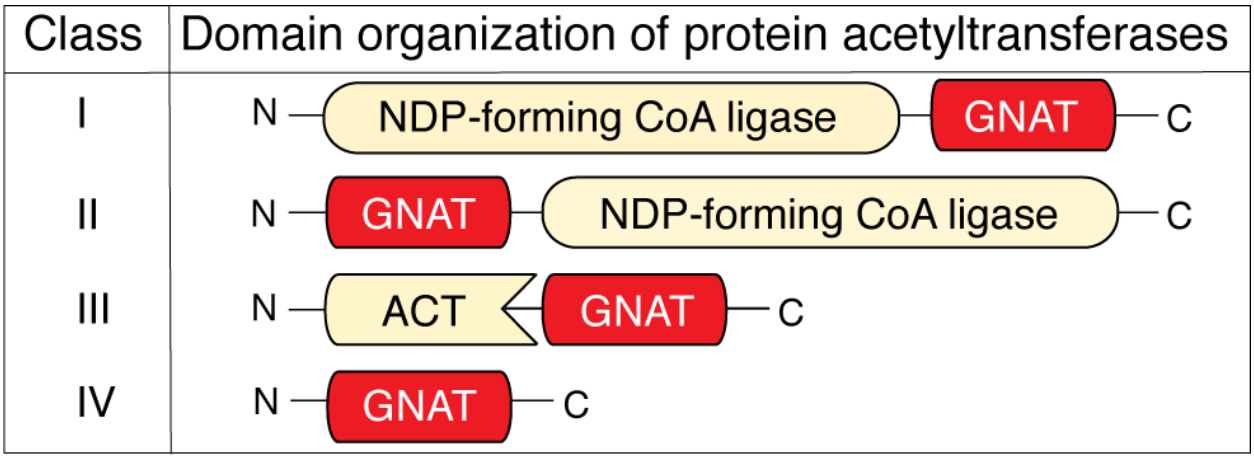


Figure 8.1. Classification of prokaryotic protein acetyltransferases based on domain organization.

actinomycete *Streptomyces coelicolor* is acetylated *in vivo* (21), but the acetyltransferase responsible for the acetylation has not been reported. In *S. lividans*, it was found that a Type III ACT-domain acetyltransferase (PatB), acetylated Acs of this organism (19). While generally one Pat protein targets all the CoA ligases in the cell, it is not unprecedented for an organism to have two acetyltransferases that target similar and dissimilar CoA ligases (8). How and why these specificities overlap is of interest. *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 (22-25). To characterize the role protein acetylation plays in the physiology of *S. lividans*, we have identified targets of acetylation *in vivo* using using 2-dimensional gel electrophoresis, acetyl-lysine western blotting, and mass peptide fingerprinting. Through this we identified 4 AMP- forming CoA ligases that are acetylated in *S. lividans*. We demonstrate that these uncharacterized AMP-forming CoA ligases recognize distinct substrates, and that they are acetylated by PatB on their active site lysyl residues. Additionally, we purify and characterize other putative CoA ligases and find they are also acetylated by PatB. Interestingly, none of these CoA ligases are acetylated by PatA, providing some interesting insight into the specificity of acetyltransferases in *Streptomyces.*

MATERIALS AND METHODS

Bacterial strains and growth conditions.

All strains and plasmids used in this study are listed in Tables 8.1 and 8.2, respectively. *Streptomyces* strains are derivatives of *Streptomyces lividans* TK24. ISP- 2 medium (26) or R2YE medium (27) was used to culture *S. lividans* on solid medium. Liquid cultures of *S. lividans* were

Table 8.1. Strains used in this study.

Strain or Plasmid	**Relevant Genotype**	**Reference[2]**
E. coli strains		
JE9314	C41λDE3) *pka12*::*kan*$^{+}$	Lab Collection
S. lividans strains		
TK24	Wild type	Lab Collection

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

Table 8.2. Plasmids used in this study.

Plasmids	Relevant Genotype	Reference[3]
pTEV5	N-terminal, rTEV-cleavable His_6-tag overexpression vector, *bla*$^+$	{Rocco, 2008 #11866}
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*$^+$	
pEFD68037-3	*S. lividans* MlaL^{K538A} allele in pTEV5, *bla*$^+$	
pEFD65795-3	*S. lividans* LbuL^{K527A} (EFD65795) allele in pTEV5, *bla*$^+$	
pEFD66106-1	*S. lividans* EFD66106 allele in pTEV5, *bla*$^+$	
pEFD67678-1	*S. lividans* EFD67678 allele in pTEV5, *bla*$^+$	
pEFD64965-1	*S. lividans* EFD64965 allele in pTEV5, *bla*$^+$	
pEFD64524-1	*S. lividans* EFD64524 allele in pTEV5, *bla*$^+$	

inoculated with 1 x 10^9 spores (28). *S. lividans* grown in either yeast extract-malt extract (YEME) rich medium (27). *S. lividans* liquid cultures were grown in baffled flasks with marine-grade stainless steel springs to aid in cell dispersion. Strains were cultured for 72 h 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) (29). When necessary, antibiotics were used at the following concentrations: ampicillin, 100 μg ml^{-1}; chloramphenicol, 12.5 μg ml^{-1}, thiostrepton, 10 μg ml^{-1}.

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, pH 7.5 @ 25°C) 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). Samples were centrifuged at 40,000 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).

Two-dimensional Western blot analysis.

Two-dimensional electrophoresis was performed according to the carrier ampholine method of isoelectric focusing (30, 31) 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 μg 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 μL 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 (32). 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 8.3.

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 8.3. 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 (33).

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} ~ 0.4 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 (20 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). The extract was cleared by centrifugation at 4°C

Table 8.3. Primers used in this study.

Name	Sequence[4]
EFD65795 5’ NheI pTEV5	GTAGCTAGCATGACCGCACCCGCGCCCCAGC
EFD65795 3’ EcoRI pTEV5	ACTGAATTCTCAGGGGCGCGTTCCGTACCGC
EFD65796 5’ NheI pTEV5	GTAGCTAGCATGACGACGGCCACGGAGCTG
EFD65796 3’ EcoRI pTEV5	ACTGAATTCTCACCGGAAGTCCTCCTCGCGG
EFD68037 5’ NheI pTEV5	GTAGCTAGCATGTCGCCCCGGGAGGACAC
EFD68037 3’ EcoRI pTEV5	ACTGAATTCTCAGAGCCGGGTGACGTCGAG
EFD64737 NheI pTEV5	GTAGCTAGCATGAGCAGCGAGCCGACGAC
EFD64737 EcoRI pTEV5	ACTGAATTCTCACGCGCCCCGCTGGTC
EFD66106 5’ NheI pTEV5	GTAGCTAGCATGTCGTCACCAACACGATACGCTCG
EFD66106 3’ EcoRI pTEV5	ACTGAATTCTTATGTCGTCGGCGTCGCGC
EFD67678 5’ NheI pTEV5	GTAGCTAGCATGTTCCGCAGCGAGTACGC
EFD67678 3’ EcoRI pTEV5	ACTGAATTCTCATCGCGGCTCCCTGAGC
EFD64965 5’ NheI pTEV5	GTAGCTAGCATGACGCCCGGACACGGCAGC
EFD64965 3’ EcoRI pTEV5	ACTGAATTCTCAGGCACCGGCGAAGCGGTCC
EFD64524 5’ NheI pTEV5	GTAGCTAGCATGAGCAAGACTCAACGGACC
EFD64524 3’ EcoRI pTEV5	ACTGAATTCTCACGGCCGGGGCAACTGC
EFD65795 5’ K527A	CACCTTGCGCACCGCCCCCGAGACGGTC
EFD65795 3’ K527A	GACCGTCTCGGGGGCGGTGCGCAAGGTG
EFD68037 5’ K538A	CACCTTCTTGTCGAACGCGCCGACGCTCGTCTTC
EFD68037 3’ K538A	GAAGACGAGCGTCGGCGCGTTCGACAAGAAGGTG

for 30 min at 40,000 × *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 (80 mM)]. H_6-tagged proteins were eluted with 5 CV Buffer C [Tris- HCl buffer (50 mM, pH 8), NaCl (500 mM), and imidazole (500 mM)]. All fractions containing H_6- tagged proteins were combined. rTEV protease (1:50 mg:mg ratio) 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 (20 mM) for 3 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 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 ligase assays.**

Activities of *Sl*MlaL, *Sl*LbuL, *Sl*IbaL, and *Sl*PaaL were measured using an NADH-consuming assay (13). *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, 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 (34). Specific activity data are presented in triplicate

***In vitro* acetylation assays.**

CoA ligases (3 μM) and PatB (1 μM) were incubated with or without [1-^{14}C]-acetyl-CoA (20 μM) in HEPES (50 mM, pH 7.5) and *tris*(2-carboxyethyl)phosphine (TCEP, 1 mM), for 1 h at 37 ºC in a total volume of 25 μl. Reactions were quenched by the addition of SDS-loading buffer (60% (v/v) glycerol, Tris-HCl pH 6.8 (0.3 M), EDTA (12 mM), 12% SDS, 2-mercaptoethanol (0.87 mM), bromophenol blue (0.05%, w/v) and reaction mixtures were resolved by SDS-PAGE on a 15% (w/v) polyacrylamide gel with Tris-HCl buffer pH 8.8 (resolving gel)/Tris-HCl pH 6.8 (stacking gel). Samples were run at 200V for 45 min. Transfer of the radiolabel onto TacA was visualized using a Typhoon Trio+ variable mode imager (GE Healthcare).

RESULTS.

Identification of acetylated proteins in *S. lividans*. We identified the acetylated in *S. lividan*s using 2-D gel electrophoresis followed by Western Blotting. Soluble proteins from *S. lividans*

wild-type strains grown on β-hydroxybutyrate were subject to isoelectric focusing followed by SDS-PAGE. Gels were stained to confirm equal loading (Fig. 8.2A,B). *S. lividans* wild-type protein extracts were also transferred to PVDF membrane for Western blotting. Acetylated proteins were detected using acetyllysine antibodies. We detected 6 distinct acetyllysine spots in lysates of *S. lividans* wild-type strain (Fig. 8.3A,B). Acetylated protein spots were excised from the corresponding pI and MW position in the duplicate, coomassie-stained protein gels (Fig. 8.4A,B) for mass peptide fingerprint. Shown in Table 8.4, 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.

Identification of the acetyltransferase that acetylated *S. lividans* CoA ligases.

Attempts to acetylate any of the identified CoA ligases with *Sl*PatA were unsuccessful under the conditions tested. However, incubation of the CoA ligases, EFD68037 and EFD65795, with an ACT-Pat domain acetyltransferase, *Ma*PatB, and [1-^{14}C]-acetyl-CoA, lead to acetylation on the active site lysine of these proteins (Fig. 8.5). The *Micromonospora aurantiaca* ACT-domain *Ma*PatB is homolgous to *Sl*PatB and was used for *in vitro* acetylation assays due to known inactivity of *Sl*PatB when isolated from *E. coli* (19). It was found that *Ma*PatB acetylated EFD68037 and EFD65795 on the active site lysine of these proteins, because transfer of the acetyl-group was not detected on the EFD68037^{K538A} or EFD65795^{K527A} protein variants (Fig. 8.5). Acetylation of EFD65796 and EFD64737 by *Ma*PatB was detected, but isolation of the lysine variants of these proteins was unsuccessful and therefore the data are not shown.

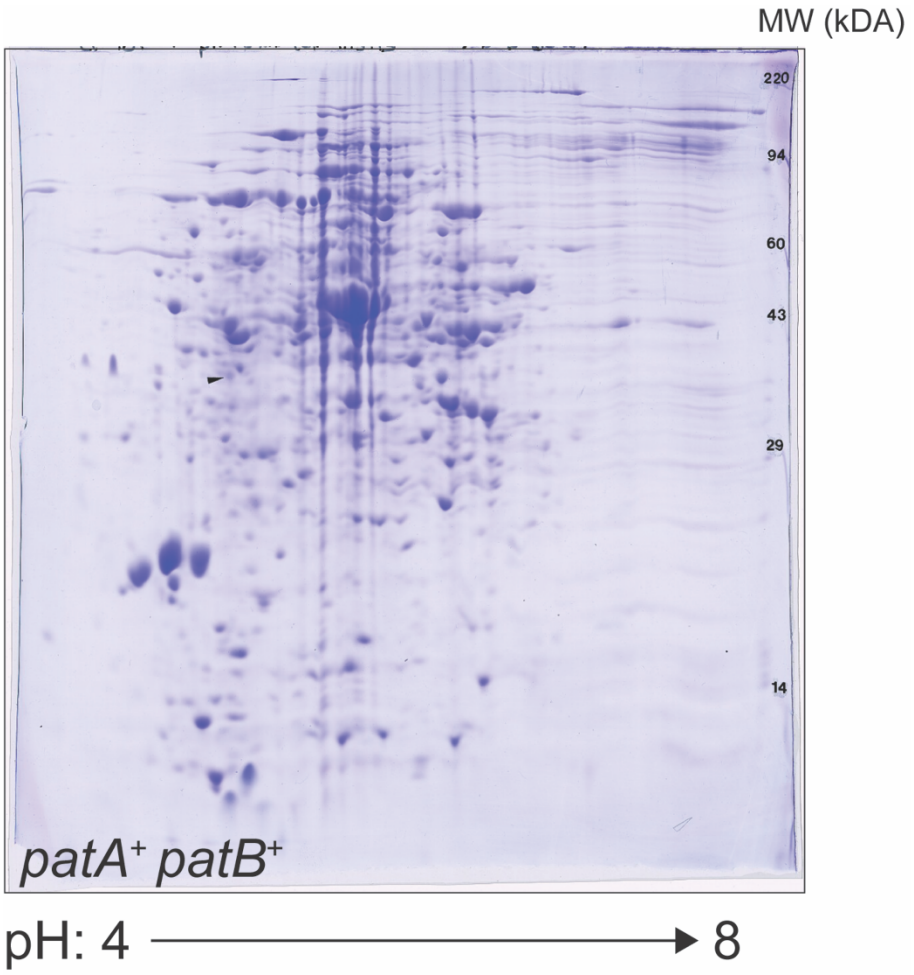


Figure 8.2. Resolution of *S. lividans* wild-type soluble proteins. Two-dimensional gel electrophoresis was used to resolve the soluble proteins of a wild-type strain of *S. lividans*. Two technical replicates were run. Shown here is one replicate from the sample stained with Coomassie Blue. The second replicate for each protein sample was transferred to a PVDF membrane for Western blot (see Figure 8.3). MW, molecular weight standards.

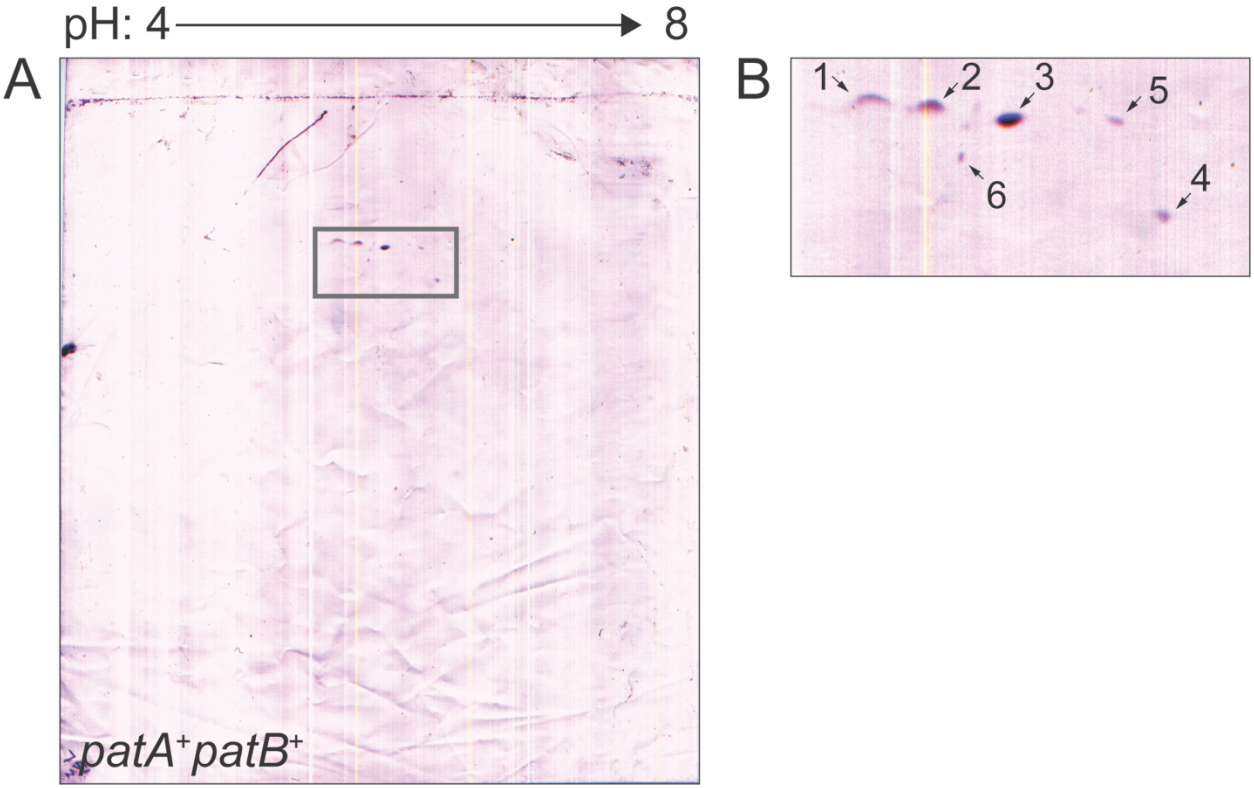


Figure 8.3. Detection of lysine-acetylated proteins in *S. lividans* wild-type strain soluble proteins. Identical gels to those described in figure 8.2 were transferred to PVDF membranes for Western blot and probed with anti-acetyllysine antibodies to detect proteins with acetylated lysines. (A) Anti- 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.

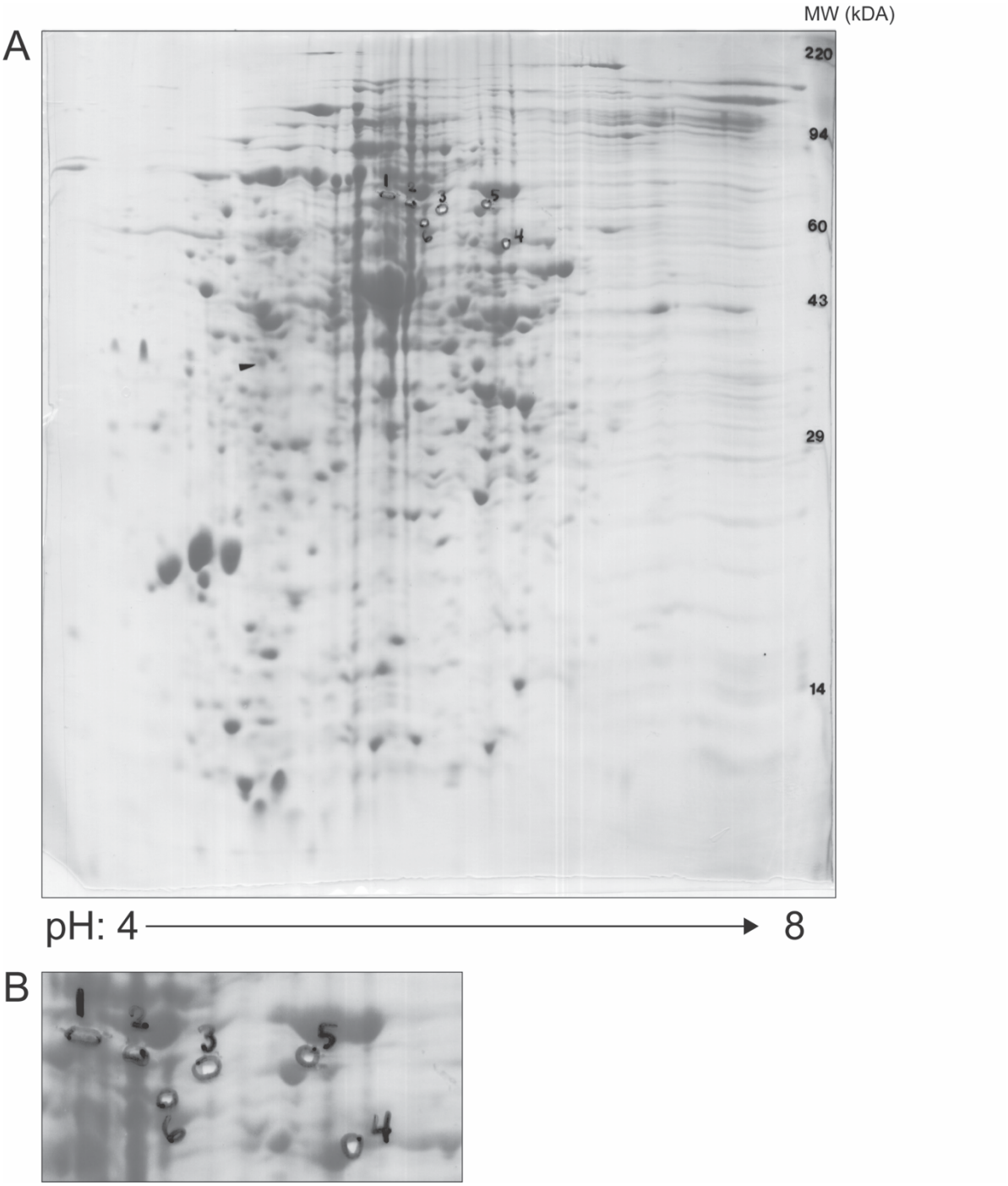


Figure 8.4. Protein spots corresponding to lysine-acetylated proteins were excised after SDS-PAGE. (A) The 6 acetyllysine spots identified by Western blot analysis (Figure 8.3A,B) were excised from the corresponding pI and molecular weight locations on a duplicate gel (Figure 8.2A). (B) The region corresponding to the location of the spots cut was magnified from clarity. MW, molecular weight standards.

Table 8.4. Identification of acetylated proteins.

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

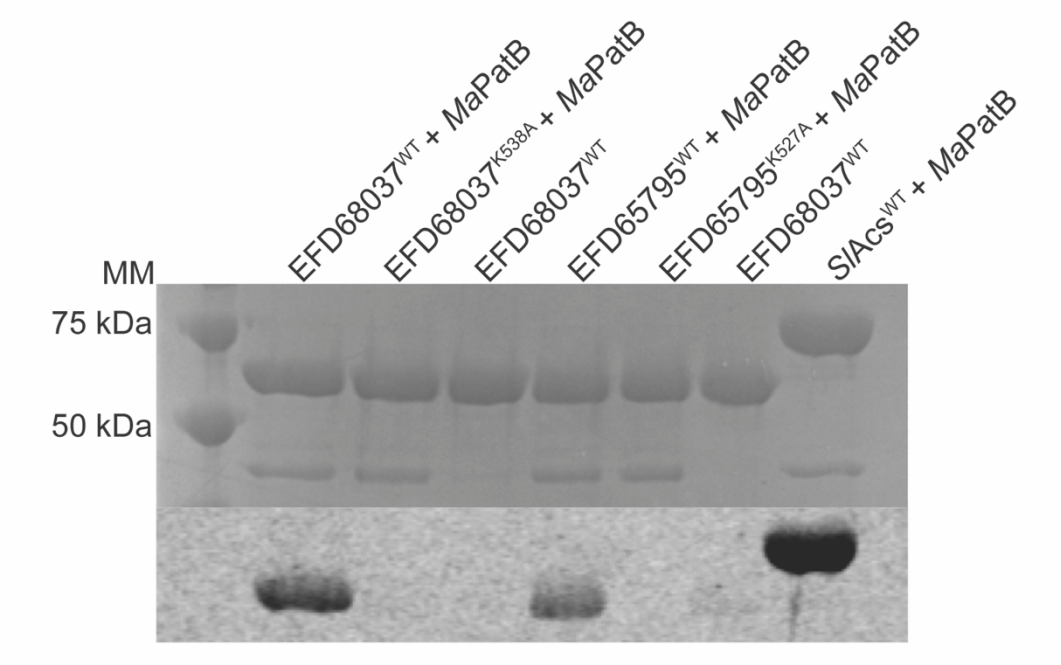


Figure 8.5. Acetylation of EFD68037 and EFD65795 by *Ma*PatB. CoA ligases (3 μM) were incubated with *Micromonospora* PatB (*Ma*PatB, 1 μM) and [1-^{14}C]-acetyl-CoA. Proteins were separated by SDS-PAGE and stained with Coomassie Brilliant Blue R to visualize proteins. Precision Plus Protein™ (BioRad) standard was used as a molecular marker (MM) with relevant molecular masses indicated in kDa. Acetylation was visualized by phosphor imaging. EFD68037 and EFD69795 are 60 kDa , *Sl*Acs is 70 kDa, and *Ma*PatB is 37 kDa.

Characterization of putative 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 Materials and methods (Fig. 8.7). 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.

Changes in CoA ligase activity upon acetylation by *Ma*PatB.

To further characterize the acetylation of EFD65795 and EFD68037, both CoA ligases were incubated with *Ma*PatB with or without acetyl-CoA, as described in materials and methods. After incubation, EFD68037 was tested for CoA ligase activity with hexanoate and EFD65795 with isovalerate (Fig. 8.8). Upon acetylation by *Ma*PatB, activity of EFD68037 was decreased by 10% and activity of EFD65795 was decreased by 20%. These changes were minimal, but similar to those seen after *Ma*PatB acetylation of *Sl*Acs (19). With regards to *Sl*Acs, a serine residue within the acetylation motif was needed for lysine acetylation by PatB. This serine modification only occurred *in vivo* in *S. lividans* and it is possible that these CoA ligases will need to be purified and tested for their acetylation ability by PatB. To rule out a pH or buffer condition needed for *Ma*PatB to be active, we acetylated EFD65795 across a range of buffers (Fig. 8.9). However, this did not appear to make a difference in activity of EFD65795 after acetylation, and pursuing the above

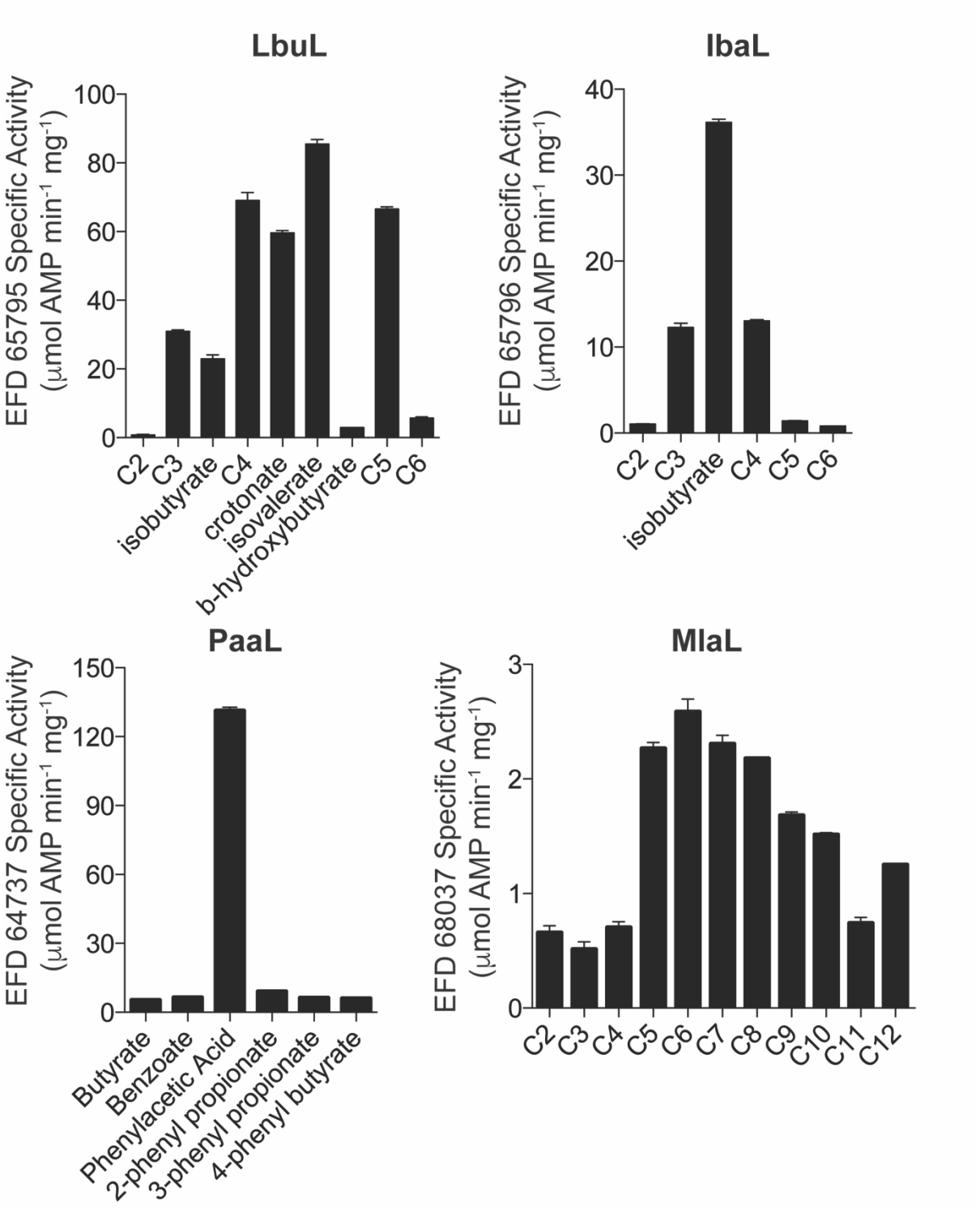


Figure 8.7. Characterization of new AMP-forming CoA ligases. AMP-forming CoA ligases identified as acetylated by mass peptide fingerprinting were tested with a library of organic acid substrates (see materials and methods) to identify the substrate preferences. Activity of each CoA ligase was measured in an NADH consumption assay. Data represent mean ± S.D. of triplicate experiments.

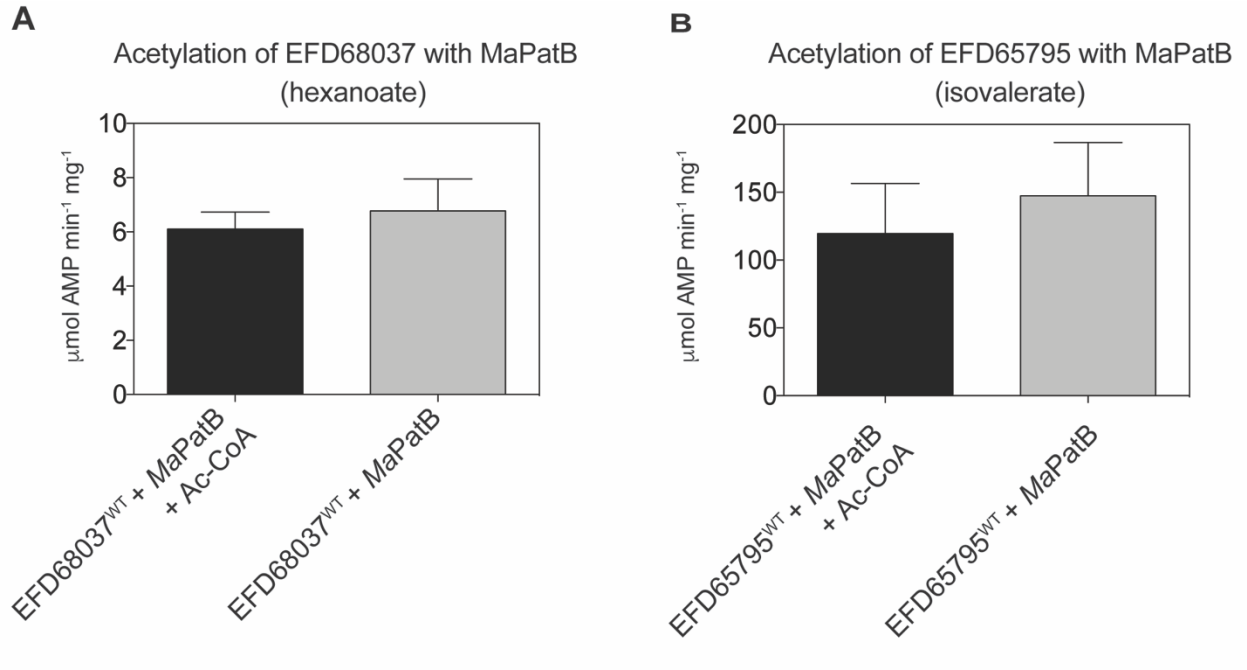


Figure 8.8. CoA ligase activity after acetylation by *Ma*PatB. AMP-forming CoA ligases EFD68037 (A) and EFD65795 (B) were incubated with *Ma*PatB with or without acetyl-CoA (see materials and methods). After incubation, CoA ligase activity was determined for acetylated and unacetylated proteins using a CoA ligase activity assay that monitors NADH consumption. Data represent mean ± S.D. of triplicate experiments.

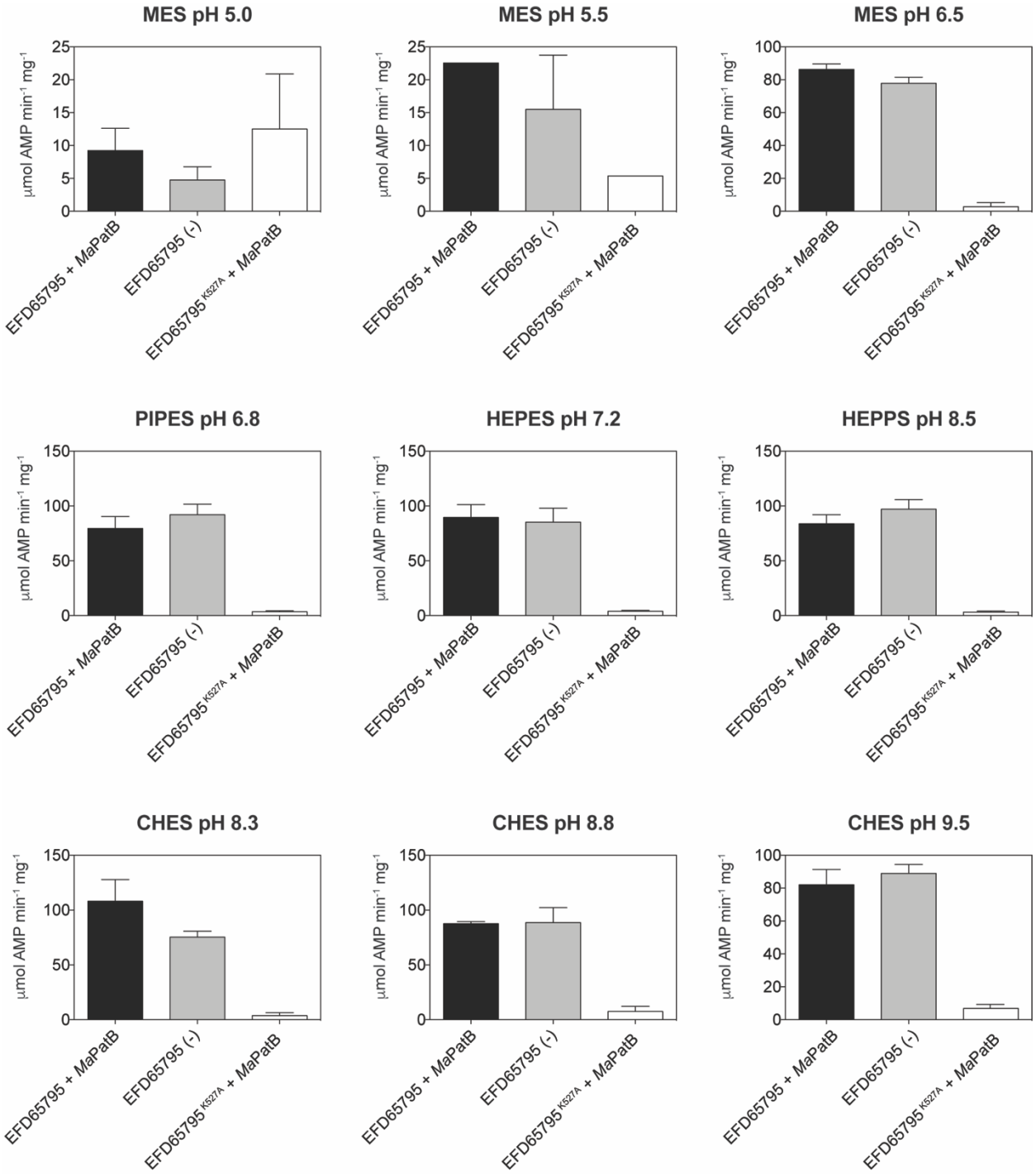


Figure 8.9. EFD65795 CoA ligase activity after acetylation by *Ma*PatB in a range of buffers. AMP-forming CoA ligase EFD65795 was incubated with *Ma*PatB with or without acetyl-CoA (see Materials and methods) in MES, PIPES, HEPES, HEPPS, and CHES. EFD65795 represents incubation without acetyl-CoA and EFD65795^{K527A} is an inactive variant of this protein. After incubation, CoA ligase activity was determined for acetylated and unacetylated proteins using a CoA ligase activity assay that monitors NADH consumption. Data represent mean ± S.D. of triplicate experiments.

Acetylation of CoA ligases not identified in mass spectrometry.

In order to determine if there were additional CoA ligases in *S. lividans* that could be acetylated by *Ma*PatB, we cloned and purified putative CoA ligases that had acetylation motifs and lysyl-residue active sites (Table 8.5). Interestingly, incubation of *Ma*PatB with EFD66106 and EFD64524 resulted in transfer of the acetyl group from acetyl-CoA (Fig. 8.10). However, incubation with EFD64965 and EFD67678 did not yield acetylated CoA ligase protein. It would be interesting and worthwhile to determine if these CoA ligases are acetylated by PatA and their range of substrate specificity for various fatty-acids.

DISCUSSION

In this study, we expand the scope of protein lysine acetylation in *S. lividans*. By following acetylation during growth of *S. lividans*, we identified 4 proteins whose acetylation was correlated with the presence of acetyltransferases *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.

PatB 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* (3, 8). Interesting, in *S. lividans,* another acetyltransferase, PatB, acetylated Acs of this organism (19). Our study identified 4-6 AMP-forming CoA ligases that are regulated by PatB, rather than PatA. Substrates were identified for each of the initial 4 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

Table 8.5. Additional CoA ligases coded for in *S. lividans* genome.

Accession number	Annotation	Acetylation motif
EFD66106	Acyl-CoA ligase	PRTATG**K**LQRYRLLD
EFD67678	Acyl-CoA ligase	PRAASG**K**ILRRQPAG
EFD64965	Acyl-CoA ligase	PRNASG**K**ILKRELRD
EFD64524	Acyl-CoA ligase	PLTAVG**K**VDKAALAR

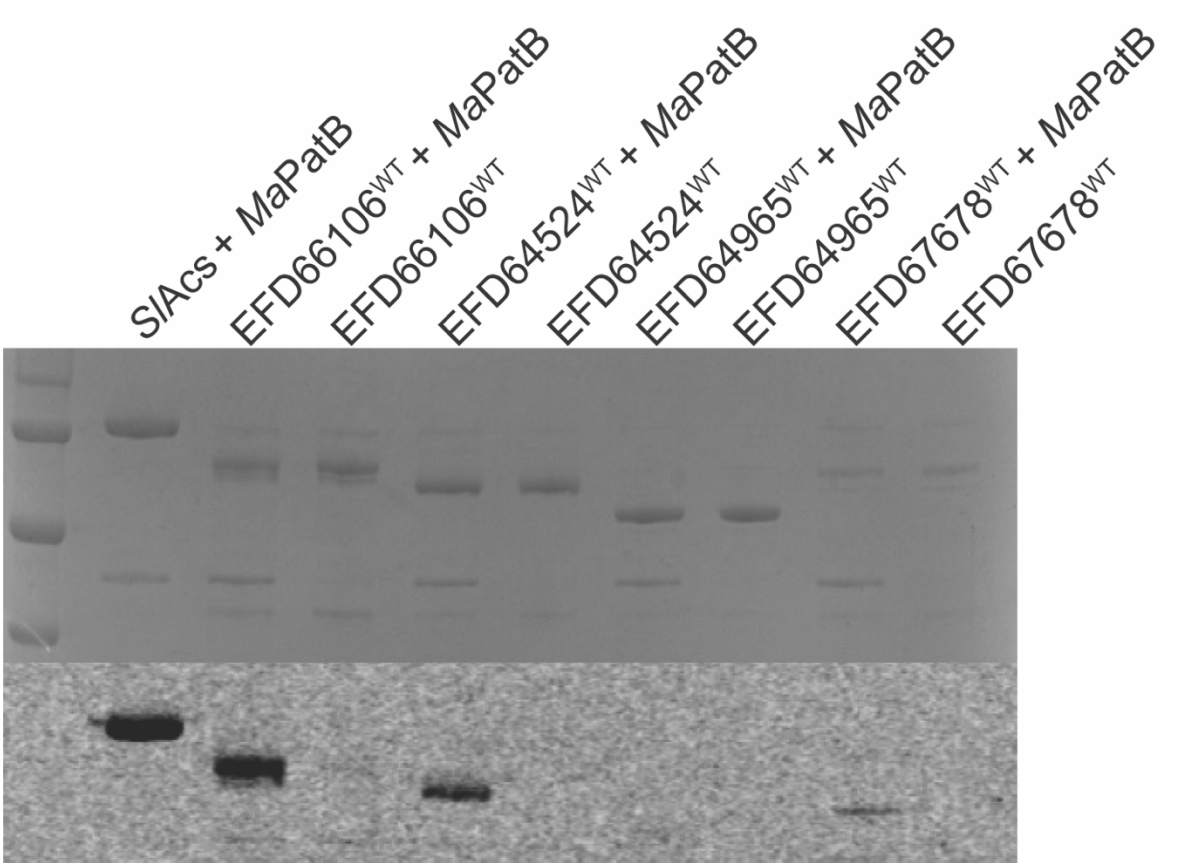


Figure 8.10. Acetylation of EFD66106, EFD64524, EFD64965, and EFD67678 by *Ma*PatB. CoA ligases (3 μM) were incubated with *Micromonospora* PatB (*Ma*PatB, 1 μM) and [1-^{14}C]-acetyl-CoA. Proteins were separated by SDS-PAGE and stained with Coomassie Brilliant Blue R to visualize proteins. Precision Plus Protein^{TM} (BioRad) standard was used as a molecular marker (MM) with relevant molecular masses indicated in kDa. Acetylation was visualized by phosphor imaging. Postitive control of *Sl*Acs is 70 kDa, and *Ma*PatB is 37 kDa.

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* PatB promotes efficient acetylation of AMP-forming CoA ligases *in vivo*. Previous work with *S. lividans* Acs identified a serine residue within the acetylation motif that was acetylated by an unknown mechanism when purified from *S. lividans* (19). It wasn't until this serine was acetylated that *Ma*PatB could acetylated and inactivate Acs beyond a 20% decrease. We see this same 20% decrease of EFD65795 and EFD68037 upon acetylation by *Ma*PatB. It would be interesting in the future to purify these CoA ligases from *S. lividans* and characterize their posttranslational modifications, to determine if these similarities amongst Acs exist.

We also identified other putative CoA synthetases from *S. lividans* via bioinformatic analyses. A few of these CoA ligases (EFD66106 and EFD64524) were acetylated by *Ma*PatB *in vitro*. It would be interesting to determine if the CoA ligases that were not modified by *Ma*PatB (EFD64965 or EFD67678) were instead modified by *Sl*PatA. These studies could provide some interesting insight into the specificity of *S. lividans* acetyltransferases and how CoA ligases are regulated in the cell.

Is activity of the acetylation targets regulated by reversible lysine acetylation?

Deacetylation by sirtuin (NAD^+-dependent) deacetylases or histone deacetylases (NAD^+-independent) reactivates members of the AMP-forming CoA ligase family. *S. lividans* encodes 2 sirtuins and 1 histone deacetylase. 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. Kouzarides T. 2007. Chromatin modifications and their function. Cell 128:693-705.
2. Liang W, Deutscher MP. 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 285:29054-290548.
3. Starai VJ, Celic I, Cole RN, Boeke JD, Escalante-Semerena JC. 2002. Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science 298:2390-2392.
4. Soppa J. 2010. Protein acetylation in archaea, bacteria, and eukaryotes. Archaea 2010:pii: 820681.
5. Kim GW, Yang XJ. 2011. Comprehensive lysine acetylomes emerging from bacteria to humans. *Trends Biochem Sci* 36:211-220.
6. Weinert BT, Iesmantavicius V, Wagner SA, Scholz C, Gummesson B, Beli P, Nystrom T, Choudhary C. 2013. Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. Mol Cell 51:265-272.
7. Berndsen CE, Denu JM. 2008. Catalysis and substrate selection by histone/protein lysine acetyltransferases. Curr Opin Struct Biol 18:682-689.
8. Crosby HA, Pelletier DA, Hurst GB, Escalante-Semerena JC. 2012. System-wide studies of *N*-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. J Biol Chem 287:15590-15601.
9. Crosby HA, Rank KC, Rayment I, Escalante-Semerena JC. 2012. Structure-guided expansion of the substrate range of methylmalonyl-CoA synthetase (MatB) of *Rhodopseudomonas palustris*. Appl Environ Microbiol 78:6619-6629.

10. Kuo YM, Andrews AJ. 2013. Quantitating the specificity and selectivity of Gcn5-mediated acetylation of histone H3. PLoS One 8:e54896.

11. Hentchel KL, Escalante-Semerena JC. 2015. Acylation of biomolecules in prokaryotes: a widespread strategy for the control of biological function and metabolic stress. Microbiol Mol Biol Rev 79:321-346.

12. Gulick AM. 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. Crosby HA, Heiniger EK, Harwood CS, Escalante-Semerena JC. 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.

14. Gardner JG, Grundy FJ, Henkin TM, Escalante-Semerena JC. 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD^+ involvement in *Bacillus subtilis*. J Bacteriol 188:5460-5468.

15. Hallows WC, Lee S, Denu JM. 2006. Sirtuins deacetylate and activate mammalian acetyl-CoA synthetases. Proc Natl Acad Sci U S A 103:10230-10235.

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

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

18. Tucker AC, Escalante-Semerena JC. 2013. Acetoacetyl-CoA synthetase activity is controlled by a protein acetyltransferase with unique domain organization in *Streptomyces lividans*. Mol Microbiol 87:152-167.

19. VanDrisse CM, Escalante-Semerena JC. 2018. In *Streptomyces lividans*, acetyl-CoA synthetase activity is controlled by O-serine and N(epsilon)-lysine acetylation. MolMicrobiol 107:577-594.

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

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

22. Chater KF. 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.

23. Courtois S, Cappellano CM, Ball M, Francou FX, Normand P, Helynck G, Martinez A, Kolvek SJ, Hopke J, Osburne MS, August PR, Nalin R, Guerineau M, Jeannin P, Simonet P, Pernodet JL. 2003. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. Appl Environ Microbiol 69:49-55.

24. McMahon MD, Guan C, Handelsman J, Thomas MG. 2012. Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. Appl Environ Microbiol 78:3622-3629.

25. Seow KT, Meurer G, Gerlitz M, Wendt-Pienkowski E, Hutchinson CR, Davies J. 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.

26. Shirling EB, Gottlieb D. 1966. Methods for characterization of *Streptomyces* species. Int J Syst Bacteriol 16:313-340.

27. Kieser T, Bibb MJ, Buttner MJ, Chater K, Hopwood DA. 2000. Media, buffers, and suppliers, p 405-420, Practical *Streptomyces* Genetics. John Innes Foundation, Norwich, England.

28. Kieser T, Bibb MJ, Buttner MJ, Chater K, Hopwood DA. 2000. Growth and preservation of *Streptomyces*, p 43-62, Practical *Streptomyces* Genetics. John Innes Foundation, Norwich, England.

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

30. Talmi-Frank D, Jaffe CL, Baneth G. 2009. Quantitative computerized western blotting. Methods Mol Biol 536:103-113.

31. O'Farrell PH. 1975. High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007-4021.

32. Elion EA, Marina P, Yu L. 2007. Constructing recombinant DNA molecules by PCR, p Unit 3.17.1-3.17.12. *In* Ausubel FM, R. Brent RE, Kingston DD, Moore JG, Seidman JA, Smith a, Struhl K (ed), Curr Protoc Mol Biol. Greene Publishing Associates & Wiley Interscience, New York, N.Y.

33. Blommel PG, Fox BG. 2007. A combined approach to improving large-scale production of tobacco etch virus protease. Protein Expr Purif 55:53-68.

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

CHAPTER 9

CONCLUDING REMARKS

SUMMARY AND CONCLUSIONS.

Each chapter in this dissertation outlines data related to prokaryotic acetyltransferases. For more in depth conclusions with regards to each project please refer to the conclusions section within each chapter. Overall, this dissertation highlights the potential within the field of acetylation. Chapters 3, 4, 5, 6, and 7 characterize acetyltransferases all with different substrates. Some of these acetyltransferase substrates are small molecule (Chapters 3, 6, and 7), while others are protein substrates (Chapters 4, 5, and 8). When tackling the characterization of putative acetyltransferases, we must analyze protentional small molecule or protein substrates and think globally about the physiological consequences. While this dissertation has made advances in identifying an abundance of processes affected by acetylation, this only means that there are infinite other possibilities. We have encountered the vast unknown and can confidently say, acetylation could be involved in any process within the cell. Hopefully the graduate students that follow in my footsteps can make some waves and cause major breakthroughs with regards to prokaryotic acetylation and how it relates to cell physiology. Thanks for reading.