INSIGHTS INTO THE ROLE OF ACETYLATION IN SALMONELLA ENTERICA: REGULATION OF REVERSIBLE LYSINE ACETYLATION AND THE DETOXIFICATION OF OXIDIZED METHIONINE DERIVATIVES

by

KRISTY LYNN HENTCHEL

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

ABSTRACT

Protein and small molecule acetylation are widespread in nature. Acetylation can affect the function, activity, or stability of proteins, or alter the structure of small molecules. Acetylation provides an advantageous mechanism for altering activity of the target substrate, allowing for rapid alteration of physiology in response to environmental cues. Many of the enzymes catalyzing acetylation reactions belong to the Gcn5 N-acetyltransferase (GNAT) family. The majority of organisms encode around ~25 GNATs and the majority of these enzymes do not have known functions.

In *S. enterica*, reversible lysine acetylation (RLA) of proteins is controlled by a GNAT (Pat) and a NAD$^+$-dependent sirtuin deacetylase (CobB). RLA control is needed to modify the activity of the AMP-forming CoA ligases, such as the acetyl-CoA synthetase (Acs) enzyme that is required for growth in conditions containing acetate. The first portion of this work describes the identification of a transcriptional regulator, IolR, which activates expression of this system, and integrates the expression with that of the target substrate, Acs. This provided the first example of
the acetylation / deacetylation system and its target being under the control the same transcriptional regulator.

Only one protein acetyltransferase has been identified in *S. enterica* (Pat), and the majority of the characterized GNATs of *S. enterica* acetylate small molecule targets. The second portion of this work characterizes a subset of GNAT enzymes annotated as phosphinothricin acetyltransferases. These enzymes display varying specificity for phosphinothricin, or the closely related analogues methionine sulfoximine and methionine sulfone. This work provided information on how *S. enterica* and organisms respond to and detoxify harmful compounds present in the environment.

INDEX WORDS: acetylation, Gcn5 N-acetyltransferase, bacteria, *Salmonella*
INSIGHTS INTO THE ROLE OF ACETYLATION IN SALMONELLA ENTERICA: REGULATION OF REVERSIBLE LYSINE ACETYLATION AND THE DETOXIFICATION OF OXIDIZED METHIONINE DERIVATIVES

by

KIRSTY LYNN HENTCHEL

BS, Indiana University, 2009

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

2015
INSIGHTS INTO THE ROLE OF ACETYLATION IN SALMONELLA ENTERICA: REGULATION OF REVERSIBLE LYSINE ACETYLATION AND THE DETOXIFICATION OF OXIDIZED METHIONINE DERIVATIVES

by

KRISTY LYNN HENTCHEL

Major Professor: Jorge C. Escalante-Semerena
Committee: Diana M. Downs
Anna C. Glasgow Karls
Timothy R. Hoover

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2015
ACKNOWLEDGEMENTS

I want to thank the faculty members at both the University of Wisconsin and the University of Georgia for their help, support, and valuable advice throughout my graduate studies. I would like to especially recognize my advisor, Dr. Jorge Escalante-Semerena, for being wonderfully supportive and optimistic about my science, and taking the time to help me grow and develop as a scientist. My thesis committee members at both the University of Wisconsin (Dr. Diana Downs, Dr. Paul Weimer, Dr. Nancy Keller, and Dr. Rodney Welch) and the University of Georgia (Dr. Diana Downs, Dr. Anna Glasgow Karls, and Dr. Timothy Hoover) have each contributed to my scientific development throughout my graduate career.

I would also like to acknowledge my undergraduate advisor, Dr. Malcolm Winkler, for providing me with a solid scientific foundation for my graduate work, as well as continuing to support me throughout my graduate career. Under the mentorship of both he and Dr. Smirla Ramos-Montanez I first learned how to approach a scientific question.

Members of the Escalante lab (Dr. Alex Tucker, Dr. Heidi Crosby, Dr. Chi Ho Chan, Dr. Sandy Thao, Ted Moore, Norbert Tavares, and Chelsey VanDrisse) have provided help with technical support, scientific feedback, experimental design, troubleshooting, and data analysis. These individuals have been an invaluable resource and have helped me tremendously. Thank you to all of my colleagues, classmates, friends, and family who have provided encouragement throughout my graduate career.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGEMENTS</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
</tbody>
</table>

CHAPTER

1 Introduction ................................................................................................................. 1
   Overview ................................................................................................................ 1
   GCN5 N-Acetyltransferases ..................................................................................... 1
   Protein Acetylation ............................................................................................... 3
   Global Approaches to Identify Acetylated Proteins ............................................... 4
   Small Molecule Acetylation .................................................................................... 6
   Conclusions ............................................................................................................. 6
   Dissertation Outline ............................................................................................... 7
   References ............................................................................................................... 8

2 Literature Review: Acylation of biomolecules in prokaryotes: A widespread strategy
for the control of biological function and metabolic stress ........................................ 16
   Summary ................................................................................................................ 17
   Introduction .......................................................................................................... 17
   Lysine Acetyltransferases ..................................................................................... 21
   Lysine Deacetylases ............................................................................................ 33
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Throughput Identification of Acetylation Targets</td>
<td>39</td>
</tr>
<tr>
<td>Validated Reversible Lysine Acetylation Targets in Bacteria and Archaea</td>
<td>46</td>
</tr>
<tr>
<td>GNAT Structure and Substrate Specificity</td>
<td>58</td>
</tr>
<tr>
<td>Role of RLA in Maintaining Metabolic Homeostasis</td>
<td>67</td>
</tr>
<tr>
<td>Transcriptional Regulation of Genes Encoding RLA Enzymes</td>
<td>74</td>
</tr>
<tr>
<td>Conclusions</td>
<td>75</td>
</tr>
<tr>
<td>Footnotes</td>
<td>76</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>76</td>
</tr>
<tr>
<td>References</td>
<td>76</td>
</tr>
</tbody>
</table>

3 Deciphering the regulatory circuitry that controls reversible lysine acetylation in *Salmonella enterica* ............................................. 108

Abstract and Importance ....................................................................... 109
Introduction ....................................................................................... 110
Results ............................................................................................... 112
Discussion ............................................................................................ 138
Materials and Methods ......................................................................... 140
Acknowledgements ............................................................................... 149
References ........................................................................................... 149

4 In *Salmonella enterica*, the Gcn5-related acetyltransferase MddA (formerly YncA) acetylates methionine sulfone and methionine sulfoximine, blocking their toxic effects........................................................................................................... 157

Abstract............................................................................................... 158
Introduction ....................................................................................... 158
Materials and Methods ................................................................. 162
Results .......................................................................................... 171
Discussion ..................................................................................... 194
Acknowledgements ....................................................................... 196
References ..................................................................................... 197

5 Probing substrate specificity of phosphinothricin acetyltransferase homologues in

\textit{Deinococcus radiodurans} and \textit{Geobacillus kaustophilus} .................. 205

Abstract ....................................................................................... 206
Introduction ................................................................................... 207
Materials and Methods .................................................................. 209
Results ........................................................................................... 215
Discussion ..................................................................................... 226
Acknowledgements ....................................................................... 226
References ..................................................................................... 227

6 Conclusions and Future Directions ................................................. 231

Summary and Conclusions ........................................................... 231
Future Directions ........................................................................... 234
References ..................................................................................... 236
LIST OF TABLES

Table 2.1: Prevalence of prokaryotic RLA components .................................................28
Table 2.2: Roles of E. coli Gcn5 N-acetyltransferases .....................................................32
Table 2.3: Validated substrates of prokaryotic lysine acetyltransferases .........................47
Table 3.1: Strains and plasmids used in this study ............................................................114
Table 3.2: Primers used in this study ..............................................................................142
Table 4.1: Strains and plasmids used in this study ............................................................163
Table 4.2: Primers used in this study ..............................................................................164
Table 4.3: Kinetic parameters for S. enterica MddA$^{WT}$ ..................................................184
Table 5.1: Strains and plasmids used in this study ............................................................210
Table 5.2: Primers used in this study ..............................................................................211
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Gcn5-related N-acetyltransferase structure and mechanism</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Reversible lysine acetylation (RLA) system in <em>S. enterica</em></td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>Schematic of N(^e) and N(^a) acetylation</td>
<td>19</td>
</tr>
<tr>
<td>2.2</td>
<td>Reversible lysine acetylation (RLA) schematic</td>
<td>20</td>
</tr>
<tr>
<td>2.3</td>
<td>Acyltransferase nomenclature and classification</td>
<td>22</td>
</tr>
<tr>
<td>2.4</td>
<td>Acetylation mechanism of GNATs</td>
<td>24</td>
</tr>
<tr>
<td>2.5</td>
<td>GNAT and sirtuin structural overview</td>
<td>27</td>
</tr>
<tr>
<td>2.6</td>
<td>Diversity in the domain organization of prokaryotic protein acetyltransferases</td>
<td>29</td>
</tr>
<tr>
<td>2.7</td>
<td>Deacetylation mechanisms of HDACs and Sirtuins</td>
<td>34</td>
</tr>
<tr>
<td>2.8</td>
<td>Methods for the analysis of acetylomes</td>
<td>40</td>
</tr>
<tr>
<td>2.9</td>
<td>Comprehensive overview of bacterial acetylome studies</td>
<td>43</td>
</tr>
<tr>
<td>2.10</td>
<td>Synthesis of acyl-CoAs by AMP- and ADP-forming acyl-CoA synthetases</td>
<td>49</td>
</tr>
<tr>
<td>2.11</td>
<td>Binding of cAMP induces a 40 Å structural change in <em>M. tuberculosis</em> PatA</td>
<td>60</td>
</tr>
<tr>
<td>2.12</td>
<td>Interactions between the <em>T. thermophile</em> Gcn5 protein and a peptide substrate</td>
<td>61</td>
</tr>
<tr>
<td>2.13</td>
<td>Molecular interactions of <em>S. lividans</em> Pat(^{GNAT}) and <em>S. enterica</em> Acs(^{CTD})</td>
<td>63</td>
</tr>
<tr>
<td>2.14</td>
<td>Acetylation determinants outside the motif containing the acetylation site</td>
<td>64</td>
</tr>
<tr>
<td>2.15</td>
<td>CoA homeostasis</td>
<td>69</td>
</tr>
<tr>
<td>3.1</td>
<td>Reversible Lysine Acetylation (RLA) in <em>S. enterica</em></td>
<td>111</td>
</tr>
<tr>
<td>3.2</td>
<td>IolR activates <em>pat</em> expression <em>in vivo</em></td>
<td>116</td>
</tr>
</tbody>
</table>
Figure 3.3: Expression of *pat* on acetate and *myo*-inositol........................................117
Figure 3.4: *Pat* does not acetylate IolR or Crp .................................................................119
Figure 3.5: Growth study controls.........................................................................................120
Figure 3.6: IolR is a tetramer.................................................................................................121
Figure 3.7: IolR binds directly and specifically to the *pat* promoter region .......................123
Figure 3.8: IolR protein binds *pat* promoter at position -112 to -70 ....................................125
Figure 3.9: IolR binds to the P<sub>pat45</sub>-nt probe................................................................126
Figure 3.10: An *iolR* strain has a growth defect on 10 mM acetate ........................................128
Figure 3.11: Induction of *acs* expression restores growth of an *iolR* strain on 10 mM acetate ..130
Figure 3.12: Activity of acetyl-CoA synthetase (Acs)............................................................132
Figure 3.13: IolR controls expression of *acs* and *cobB* .........................................................133
Figure 3.14: Glucose differentially controls expression of *pat*, *cobB*, and *acs* .......................135
Figure 3.15: Crp activates *pat* expression ..............................................................................137
Figure 4.1: Chemical structures of methionine analogs ..........................................................161
Figure 4.2: MSX and MSO inhibit growth of an *mddA::cat<sup>+</sup>* strain .................................172
Figure 4.3: High levels of MddA<sup>E82Q</sup> variant allow an *mddA1::cat<sup>+</sup>* strain to grow in the presence of MSX .................................................................174
Figure 4.4: The effect of increasing MSX concentration on *mddA<sup>+</sup>* and *mddA::cat<sup>+</sup>* strains is detrimental to growth................................................................................175
Figure 4.5: SeMddA does not prevent growth inhibition by PHO ............................................177
Figure 4.6: Glutamine and methionine counteract the deleterious effects of MSX and MSO on growth in the absence of MddA ..............................................................................178
Figure 4.7: Addition of glutamine fully restores growth of an mddA strain exposed to MSX and MSO in rich medium ................................................................. 179

Figure 4.8: SeMddA is a dimer in solution .................................................................................. 181

Figure 4.9: SeMddA\textsuperscript{WT} acetylates methionine derivatives ........................................ 183

Figure 4.10: Neither MSX nor acetyl-MSX permits growth of a methionine auxotroph in S. \textit{enterica} ............................................................................................................. 186

Figure 4.11: An mddA\textsuperscript{+} strain exhibits biphasic growth in minimal medium containing glutamine and MSX ................................................................. 187

Figure 4.12: Growth of the mddA::cat\textsuperscript{+} glnP::Tn10d strain in the presence of MSX .......... 189

Figure 4.13: Deletion of two amino acid transporters (GlnHPQ, MetNIQ) restores growth of a \Delta mddA strain exposed to MSX and MSO .................................................................. 191

Figure 5.1: Chemical structures of PPT analogues .................................................................. 208

Figure 5.2: Growth of the \textit{S. enterica} mddA\textsubscript{1}::cat\textsuperscript{+} strain in conditions containing MSX, MSO, or PPT ............................................................................................................. 216

Figure 5.3: Complementation of annotated PPT acetyltransferases from \textit{D. radiodurans} and \textit{G. kaustophilus} in a heterologous host .................................................................................. 217

Figure 5.4: Overexpression provides resistance to higher levels of PPT, MSX, and MSO in a heterologous host ........................................................................................................ 219

Figure 5.5: Substrate specificity of annotated PPT acetyltransferases ..................................... 221

Figure 5.6: Bioinformatic analyses of annotated PPT acetyltransferases ................................. 224
CHAPTER 1
INTRODUCTION

OVERVIEW
Acetylation of biomolecules is a conserved mechanism to rapidly modify cellular components in response to environmental cues and metabolic stress [reviewed in (1)]. Acetyltransferases transfer the acetyl moiety, typically from acetyl-Coenzyme A, to a primary amine group of a protein or small molecule. This modification may be removable through the action of a deacetylase. RLA is known to affect the function of diverse cellular processes including regulation of gene expression (2), metabolism (3-6), and cell structure (7). RLA exerts its effects by modulating protein-protein interactions, enzyme activity, DNA and substrate binding, as well as protein stability (8, 9).

Work in this thesis (i.) characterizes the transcriptional regulation of the reversible lysine acetylation system in Salmonella enterica, and (ii.) examines detoxification of harmful amino acid derivatives via acetylation.

GCN5 N-ACETYTLTRANSFERASES
Gcn5 N-acetyltransferases (GNATs) are the only family of acetyltransferases that are conserved throughout conserved in archaea, bacteria, and eukaryotes (10-12), and all of the protein acetyltransferases identified in prokaryotes to date belong to the GNAT family (1) (Fig. 1.1A). Although these enzymes have very low sequence homology, they share a high structural
Figure 1.1. Gcn5-related N-acetyltransferase structure and mechanism. (A) Structure of *Tetrahymena thermophila* tGcn5 enzyme bound to Coenzyme A (CoA, black sticks) and a peptide substrate (purple sticks). (B) GNAT reaction mechanism.
homology with the presence of a conserved catalytic fold (10). The core domain is composed of a central β-sheet (six antiparallel strands) which is comprised of four motifs, named A-D (13). The acetylation reaction proceeds via deprotonation of the active site lysine residue of the target substrate by a catalytic glutamate residue of the GNAT, which allows for a direct nucleophilic attack of lysine on the carbonyl carbon of acetyl-CoA (10, 14-16) (Fig. 1.1B).

On average a given organism encodes ~25 GNATs, and many of these enzymes have no known role. Even in well-characterized organisms like S. enterica, only half of GNATs have a known function. GNAT enzymes of S. enterica have been shown to target primary amines (17, 18), including the N-termini of proteins (19, 20), aminoglycoside antibiotics (21), polyamines (22), a nucleotide sugar (23), glutamate (24), toxic aminoacyl nucleotides (25), toxic methionine derivatives (26), and transfer RNAs (27). The only known protein lysine acetyltransferase in S. enterica is Pat, (discussed in more detail below) (12, 28).

**PROTEIN ACETYLATION**

*Control of protein function by reversible lysine acetylation.* Reversible lysine acetylation (RLA) is an important posttranslational modification that allows an organism to quickly and reversibly regulate the activity of proteins, including those involved in transcription, translation, carbon utilization, and stress responses (12, 29-31) by modulating the acetylation state of epsilon amino group of lysyl residues critical for function [reviewed in (1)]. RLA has shown to be a functional system in diverse prokaryotes, including S. enterica, Rhodopseudomonas palustris, Streptomyces lividans, Escherichia coli, and Bacillus subtilis (6, 29, 32-35).

The RLA system controls protein acetylation in S. enterica (12, 36, 37), and is comprised of a Gcn5 N-acetyltransferase (Pat), and its partner NAD⁺-consuming sirtuin deacetylase (CobB). Pat
and CobB work together to modify the activity acetyl-CoA synthetase (Acs), responsible for the conversion of acetate to acetyl-CoA (38) (Fig. 1.2). RLA control of Acs is a widely conserved mechanism and present in bacteria to humans (39). Control of Acs function via acetylation is critical as uncontrolled Acs activity leads to cell death via depletion of cellular ATP pools (33). Acetylation of the conserved active site lysine residue inactivates Acs, while removal of the acetyl moiety by the CobB reactivates the enzyme.

GLOBAL APPROACHES TO IDENTIFY ACETYLATED PROTEINS

Studies examining the total acetylated protein population (the “acetylome”) in prokaryotes have been accelerated by recent advances in mass spectrometry-based approaches to enrich for and detect acetylated lysine residues in total protein populations (8). Bacterial acetylomes have been characterized in *E. coli* (40, 41), *S. enterica* (42), *B. subtilis* (34), *Erwinia amylovora* (35), *R. palustris* (3), *Staphylococcus aureus* (43), *Geobacillus kaustophilus* (44), *Vibrio parahaemolyticus* (45), *Thermus thermophilus* (46), *Mycobacterium tuberculosis* (47), and *Mycoplasma pneumonia* (48). In these studies the number of putatively acetylated proteins per organism has ranged from 62-667, providing evidence that acetylation of proteins is prevalent modification in these organisms. The majority of the identified acetylated proteins are involved in central metabolism (40, 43, 45). While these studies provide an important platform for studying the role of protein acetylation, these data have not been experimentally validated to determine if the function of the identified proteins is in fact controlled by acetylation. Validation of these data obtained through global ‘omics’ approaches is key to improving our understanding of the role of acetylation in prokaryotic cellular physiology.
Reversible lysine acetylation (RLA) system in *S. enterica*. Activity of the acetyl-CoA synthetase (Acs) enzyme of *S. enterica* is controlled by acetylation. Acs is acetylated (inactivated) by the Gcn5 N-acetyltransferase (Pat) and deacetylated (activated) by the NAD$^+$-dependent sirtuin deacetylase (CobB).

**Figure 1.2** Reversible lysine acetylation (RLA) system in *S. enterica*. Activity of the acetyl-CoA synthetase (Acs) enzyme of *S. enterica* is controlled by acetylation. Acs is acetylated (inactivated) by the Gcn5 N-acetyltransferase (Pat) and deacetylated (activated) by the NAD$^+$-dependent sirtuin deacetylase (CobB).
SMALL MOLECULE ACETYLATION

GNATs protect against many types of external and internal stresses (49-51). The majority of characterized GNATs acetylate small molecules including aminoglycoside antibiotics (49), glutamate (24), polyamines (22), aminoalkylphosphonic acid (52), dTDP-fucosamine (23), transfer RNAs (27), toxic aminoacyl nucleotides (25), phosphinothricin (53), and methionine sulfoximine (26, 51). Identification of small molecule substrates for GNATs has been an ongoing challenge in the field.

It could be hypothesized that the diversity of stresses controlled by GNATs correlates with the environment encountered by the organism. Therefore, the relevance of GNAT function to physiology varies between organisms, with respect to not only cellular stresses but also metabolic pathways. For example, *S. enterica* contains ~26 GNATs, *Myxococcus xanthus* encodes ~46 GNATs, and *S. lividans* encodes up to ~70 GNATs. This suggests that perhaps *M. xanthus* and *S. lividans* occupy a more diverse environment than *S. enterica*, while also maintaining a more complex metabolism.

CONCLUSIONS

Acetylation is an emerging field in prokaryotic biology and is advancing by tremendous strides through the use of high-throughput and detailed mechanistic studies diverse organisms. GNATs have been identified in bacteria, archaea, and eukaryotes, and are also highly abundant. Although the physiological role of the majority of GNATs remains unknown, and the elucidation of their function is been an ongoing challenge to physiologists, advances in this area of research will provide valuable insights into the strategies used by cells to cope with metabolic stress.
DISserTATION OUTLINE

This work focuses on the role of acetylation in the γ-Proteobacterium *Salmonella enterica* and furthers our knowledge of the role of acetylation in cellular physiology. I identified a transcriptional regulator of the reversible lysine acetylation (RLA) system in *S. enterica* and describe how defects in this regulatory system affect C2 metabolism. I characterized the function of a previously unknown *S. enterica* GNAT, MddA, which is responsible for the detoxification of harmful amino acid derivatives. MddA was previously incorrectly annotated as a phosphinothricin acetyltransferase. This lead to the investigation of various putative phosphinothricin acetyltransferases from *Deinococcus radiodurans*, *Geobacillus kaustophilus*, *Burkholderia xenovorans*, and *Bacillus subtilis*, which demonstrated these enzymes have varying substrate specificities for phosphinothricin analogues.

Chapter 2 provides an in-depth literature review on the topic of reversible lysine acetylation. RLA control of proteins affects many cellular processes and is an important mechanism to rapidly and reversibly modify protein function. This literature review presents a close examination of our current understanding of RLA in prokaryotes.

In chapter 3 we report that IolR, a repressor of *myo*-inositol catabolism, activates the expression of genes encoding components of the RLA system (Pat, CobB) in *S. enterica*, and integrates this expression with the target of RLA, the acetyl-CoA synthetase (Acs). Evidence from DNA binding assays and DNA footprinting studies demonstrated that the IolR protein directly regulates *pat* expression. The absence of IolR caused an imbalance in the ratios of Pat and CobB, which in turn affected the level of acetylated (inactive) / unacetylated (active) Acs. These results suggest that transcriptional control of the RLA system is critical and slight changes in protein concentration leads to an imbalance that is detrimental to the cell.
Chapter 4 describes the role of the MddA GNAT in the detoxification of oxidized forms of methionine. Methionine sulfoximine and methionine sulfone inhibited growth of an \textit{S. enterica} \(\Delta mddA\) strain unless glutamine or methionine was present in the medium. Deletion of two amino acid transporters (GlnHPQ and MetNIQ) in a \(\Delta mddA\) strain restored growth in the presence of MSX. A spectrophotometric assay and mass spectrometry were used to show that MddA was responsible for the acetylation of methionine sulfoximine and methionine sulfone. These results show MddA is the mechanism used by \textit{S. enterica} to respond to oxidized forms of methionine, which MddA detoxifies by acetyl-CoA-dependent acetylation.

Chapter 5 outlines the characterization of putative phosphinothricin acetyltransferases using \textit{S. enterica} as a heterologous host to assay for function. Using \textit{in vivo} complementation studies I examined the specificities of annotated phosphinothricin acetyltransferases from \textit{D. radiodurans} and \textit{G. kaustophilus}, and corroborated these results with enzyme studies \textit{in vitro}. Using bioinformatic approaches the specificity of putative phosphinothricin acetyltransferases from \textit{Burkholderia xenovorans} and \textit{Bacillus subtilis} were predicted, and tested \textit{in vivo}. These studies provide method for the rapid characterization of putative phosphinothricin acetyltransferases using \textit{S. enterica}.

REFERENCES


33. **Chan CH, Garrity J, Crosby HA, Escalante-Semerena JC.** 2011. In *Salmonella enterica*, the sirtuin-dependent protein acylation/deacylation system (SDPADS)


CHAPTER 2

LITERATURE REVIEW - ACETYULATION OF BIOMOLECULES: A WIDESPREAD STRATEGY FOR THE CONTROL OF BIOLOGICAL FUNCTION AND METABOLIC STRESS

---

SUMMARY

Acylation of biomolecules (e.g., proteins, small molecules) is a process that occurs in cells of all domains of life, and has emerged as a critical mechanism for the control of many aspects of cellular physiology, including chromatin maintenance, transcriptional regulation, primary metabolism, cell structure, and likely other cellular processes. Although this review focuses on the use of acetyl moieties to modify a protein or small molecule, it is clear that cells can use many weak organic acids (e.g., short-, medium-, long-chain mono and dicarboxylic aliphatics and aromatics) to modify a large suite of targets. Acetylation of biomolecules has been studied for decades within the context of histone-dependent regulation of gene expression and antibiotic resistance. It was not until the early 2000s that the connection between metabolism and physiology and protein acetylation was reported. This was the first instance of a metabolic enzyme (acetyl-CoA synthetase) whose activity was controlled by acetylation via a regulatory system responsive to physiological cues. The alluded system was comprised of an acyltransferase and a partner deacylase. Given the reversibility of the acylation process this system is also referred to as reversible lysine acylation (RLA). A wealth of information has been obtained since the discovery of RLA in prokaryotes, and we are just beginning to visualize the extent of the impact that this regulatory system has on cell function.

INTRODUCTION

Post-translational Modifications

Post-translational modifications (PTMs) are important for the regulation of protein structure and function (1). These modifications allow organisms to rapidly respond and adapt to changing environmental conditions, forgoing the need to transcribe and translate new proteins by simply
modifying the function of existing proteins. Examples of PTMs include acetylation (2), glycosylation (3), lipidation (4), methylation (5), S-nitrosylation (6), phosphorylation (7, 8), succinylation (9), ubiquitylation (10), adenylylation and phosphocholinylation (11), ADP-ribosylation (12), serine/threonine O-acetylation (13), proteolysis (14-16), and others.

Protein acylation is a broadly distributed PTM in prokaryotes. This review focuses on reversible acetyl-CoA-dependent, enzyme-driven protein acetylation, which is broadly distributed in all domains of life. Protein acetylation can occur at the N-terminus on the alpha amino group (N\textsuperscript{\alpha} acetylation) or the epsilon amino moiety of lysyl side chains (N\textsuperscript{\varepsilon} acetylation) (Fig. 2.1). Notably, N\textsuperscript{\alpha}-acetylation can occur co- or post-translationally, alters protein stability, and is typically irreversible (17, 18). In contrast, N\textsuperscript{\varepsilon}-acetylation can modify protein structure and function, and typically is reversible by a deacetylase (Fig. 2.2).

**Discovery of Protein Acetylation**

Lysine acetylation was first reported in the 1960’s as a modification of the lysine-rich N-terminal tails of eukaryotic histones (2). Histones are protein components of chromatin, the compact DNA structure in eukaryotes, and acetylation of these proteins is tightly controlled. Histone acetylation is carried out by histone acetyltransferases (HATs), and is associated with decreased DNA binding due to the loss of interactions between the epsilon amino (N\textsuperscript{\varepsilon}) group of lysines and the phosphate anions of the DNA strands. Acetylation of the N\textsuperscript{\varepsilon} group of lysines effectively neutralizes the positive charge of lysine, relaxing histone / DNA interactions, hence providing an opportunity for the transcriptional machinery to decode genes that otherwise would be unavailable (19, 20). In eukaryotes, the tumor suppressor protein p53 was the first mammalian
Figure 2.1. Schematic of $N^\varepsilon$ and $N^\alpha$ acetylation. Protein acetylation can occur via two methods, (A) acetylation of the $\varepsilon$-amino group of internal lysine residues ($N^\varepsilon$ acetylation, red) or (B) acetylation of the $N$-terminal $\alpha$-amino group ($N^\alpha$ acetylation, blue). $N^\varepsilon$-acetylation occurs post-translationally, can be reversible, and can alter protein structure and function. $N^\alpha$-acetylation occurs co- or post-translationally, is typically not reversible, and alters protein stability.
Figure 2.2. **Reversible lysine acylation (RLA) schematic.** A protein substrate (form 1) is modified by a protein lysine acetyltransferase, Pat (of Gcn5-like N-acetyltransferase family, or GNAT), resulting in the acetylated protein (form 2). This modification is reversible, either by a NAD⁺-consuming class III sirtuin deacetylase, CobB, or a Zn(II)-dependent protein deacetylase. The sirtuin deacetylase uses NAD⁺ as a substrate, not as a coenzyme. Sirtuins modify the carboxyl group of the ribose of the nicotinamide mononucleotide (NMN) moiety of NAD⁺, simultaneously releasing nicotinamide (Nm). The resulting by-product is O-acetyl-ADP-ribose (O-AADPR).
transcription factor shown to be regulated by acetylation (21). Now, there are more than 100 transcription factors have been identified as acetylation targets (22).

As mentioned above, reversible lysine acylation (RLA) was discovered in prokaryotes in the early 2000s (23, 24), and this discovery led to a rapid expansion of the role of acylation, specifically acetylation, in prokaryotic cell physiology. Subsequently, RLA has been observed in bacteria, archaea, and eukaryotes (25), and it is now clear that many non-histone proteins are also post-translationally regulated by RLA.

At present, RLA is known to affect the function of diverse cellular processes including chromatin maintenance (26), regulation of gene expression (27), metabolism (28-31), and cell structure (32). RLA exerts its effects by modulating DNA binding, protein-protein interactions, enzyme activity, substrate binding, and protein stability (33, 34).

Studies have shown that a variety of proteins are regulated by RLA, including the metabolic enzymes acetyl-CoA synthetase (23), phosphoenolpyruvate carboxykinase (35), the M2 isoform of pyruvate kinase (36), phosphoglycerate mutase-1 (37), and the structural protein α-tubulin, a subunit of microtubules (38). Acetylation has been suggested to rival phosphorylation in both its prevalence and diversity of target substrates (39).

LYSINE ACETYLTRANSFERASES

Diversity

There are three classes of lysine (Lys, K) acetyltransferases (LATs, a.k.a. KATs) that catalyze the transfer of the acetyl moiety from acetyl-CoA to the ε-amino group of lysine side chains (Fig. 2.3). These classes of LATs comprise a large and diverse set of enzymes named after their founding member(s), including (i) the Gcn5-related N-acetyltransferase family (GNATs, [named
**Figure 2.3. Acyltransferase nomenclature and classification.** Abbreviations: LAT/KAT, Lysine (K) acetyltransferase; GNAT, Gcn5 N-acetyltransferase; HAT, histone acetyltransferase; CAT, chloramphenicol acetyltransferase; aaAT, arylamine acetyltransferase; ssAT, spermine / spermidine acetyltransferase; Other, unclassified acetyltransferase; $N^\epsilon$, acetylation of the epsilon amino group of a lysine; $N^\alpha$, acetylation of the alpha amino group of any N-terminal amino acid.
after yeast Gcn5 protein (Pfam 00583)), (ii) the MYST family [named after human MOZ, yeast Ybf2/Sas3, yeast Sas2, and human Tip60 (Pfam 01853)], and (iii) the p300/CBP family [named after human hp300 and hCBP (Pfam 06466)]. The above-mentioned enzyme families differ in sequence similarity, domain organization, substrate specificity, and catalytic mechanism (40-48). The MYST and p300/CBP families are only present in eukaryotes, while the GNAT family is present in all domains of life.

Conservation of the acetyl-CoA binding domain

Regardless of their mechanistic differences, acetyltransferases contain a conserved core domain which binds to acetyl-CoA through interactions with the pyrophosphate and pantothenate moieties (41, 47, 49-52) [reviewed in (42, 53)]. Neither the adenine base nor the acetyl moiety significantly contributes to binding the core domain. Due to this binding mode, acetyltransferases bind acetyl-CoA with high affinity but low specificity (54, 55), allowing for recognition of various acyl-CoA thioesters as well as free coenzyme A (CoASH). Their specificity is then due to their structurally divergent N- and C-terminal domains outside of the core domain (49). In some cases, other acyl-CoA thioesters, such as propionyl-CoA and succinyl-CoA, are physiologically relevant substrates (43, 56, 57), and are discussed in more detail below.

Mechanisms of acetylation. The mechanism of transfer of acyl moieties used by the GNAT and MYST families involves a catalytic glutamate residue that acts as a general base, facilitating a water-mediated proton abstraction from the side chain of the substrate lysine (45, 49, 53, 58) (Fig. 2.4). The lysine amine group initiates a nucleophilic attack on the carbonyl carbon of the acetyl moiety of CoA, allowing the direct transfer of the acyl group to the substrate lysine. Members of the MYST family employ either a catalytic mechanism similar to GNATs, or a ping-
Figure 2.4. Acetylation mechanism of GNATs. The GNAT acetylation mechanism involves a catalytic glutamate that acts a general base, facilitating a water-mediated proton abstraction from the side chain of the substrate lysine. The ε-amino group of lysine performs a nucleophilic attack on the carbonyl carbon of the acetyl moiety of CoA, allowing direct transfer of the acetyl group to the lysine side chain.
pong mechanism (aka double-displacement mechanism) involving an acetylated enzyme intermediate (46). Members of the p300/CBP family are structurally distinct and do not use a catalytic base to initiate the transfer of the acyl moiety (47). Instead, it appears that the p300/CBP family uses a Theorell-Chance mechanism, a sequential mechanism that does not form a stable ternary complex (47).

**Bacterial Gcn5-related N-Acyltransferases (GNATs)**

The GNAT family is conserved among archaea, bacteria, and eukaryotes (24, 53, 59). All known bacterial lysine acetyltransferases identified to date belong to the GNAT superfamily (60). Despite having low sequence homology, GNATs share a conserved catalytic fold, and can acetylate both protein and small molecule substrates (53). The first bacterial GNATs were characterized as aminoglycoside N-acetyltransferases from *Enterococcus faecium* (61) and *Serratia marcescens* (62), demonstrating that GNATs acetylate diverse substrates, ranging from histones to antibiotics (2, 62, 63).

**Utilization of alternative acyl-CoA substrates.** Modifications of the $N^\epsilon$ amino group of lysine by propionyl, malonyl, succinyl, and butyryl moieties have also been demonstrated for metabolic enzymes, transcription factors, and histones in both bacteria and eukaryotes (56-58, 64, 65). For example, in *Salmonella enterica*, the activity of propionyl-CoA synthetase (PrpE) is controlled by propionylation (56). Both lysine propionylation and butyrylation have also been identified as reversible modifications that occur on histones (57, 66), expanding the range of GNAT-mediated regulation through their ability to utilize various acyl-CoA substrates.

**Overview of GNAT structures.** GNATs comprise one of the largest enzyme superfamilies identified thus far (>10,000 members). Dozens of GNATs structures have been resolved and are
available from the RCSB PDB (http://www.rcsb.org/pdb/home/home.do). Despite having low to moderate primary sequence homology, GNATs contain a core catalytic domain that is structurally well conserved (Fig. 2.5A, B). The GNAT domain contains a central β-sheet (six antiparallel strands) composed of four distinct motifs, A (β4, α3), B (β5, α4), C (β1, α1-2), and D (β2-3), which were originally identified by sequence similarity (67). Motif A has the highest conservation and is important for acetyl-CoA binding and catalysis [reviewed in (53, 54)].

**Abundance and distribution of GNATs.** The number of GNATs present in a given organism varies, with the majority of commonly studied organisms encoding approximately 20-25 GNATs. As an example of the diversity and abundance of GNATs: the Gram-positive, intracellular pathogen *Listeria monocytogenes* encodes ~14 GNATs, the purple non-sulfur α-proteobacterium *Rhodopseudomonas palustris* encodes ~26, and the Gram-positive actinomycete *Streptomyces lividans* encodes ~72 putative GNATs. The majority of these GNATs are uncharacterized, with no known function. Refer to Table 2.1 for an overview of the number of GNATs present in commonly studied prokaryotes. The sheer prevalence of GNATs raises many questions regarding the physiological role and substrate specificity of these enzymes. It could be speculated that the range of GNATs an organism encodes may be driven by the diversity of the environments inhabited.

**Diversity of domain organization of GNAT protein acetyltransferases.** Protein acetyltransferases of the GNAT family exhibit diverse domain architecture; four of which have been studied and will be discussed (Fig. 2.6).

Type I. The *S. enterica* Pat (SePat) enzyme, the first enzyme of this type to be discovered (24), has homologues in *E. coli* (EcPka), *R. palustris* (RpPat), *Vibrio* species, and cyanobacteria, among others. These SePat homologues are comprised of two distinct domains, a large (~700
Figure 2.5. GNAT and Sirtuin structural overview. GNAT domains are comprised of a central β-sheet and contain four motifs (A, gold; B, blue; C, red; and D, green). Two structures are shown. (A). *Tetrahymena thermophila* *TtGcn5* is shown with an H3 11-mer peptide substrate (purple sticks). (B). *Salmonella enterica* *SeAAC(6’)* is shown in complex with its substrate kanamycin (purple sticks), and CoA shown in black sticks. Examples of sirtuis structures are shown in panels C and D). These enzymes contain a Rossmann fold domain (blue) and a variable Zn(II)-binding domain (green, with Zn(II) shown as gray sphere). The binding sites for NAD⁺ (C, gold) and the acetyllysine substrate (C, red sticks; lysine in black sticks) are located in a cleft between the two domains. The products of the reaction, nicotinamide (red sticks) and ADP-ribose (gold sticks) are shown in panel D. *Thermotoga maritima* *TmSir2* (panel C) is shown with a peptide substrate (red sticks). The archaeon *Archaeoglobus fulgidus* *AfSir2* is shown in panel D. (A) PDB 1QSN; (B) PDB 2QIR; (C) PDB 2H4F; (D) PDB 1YC2.
<table>
<thead>
<tr>
<th>Organism</th>
<th>GNAT</th>
<th>Sirtuin</th>
<th>HDAC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ARCHAEAE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanococcus maripaludis</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pyrococcus furiosus</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sulfolobus solfataricus</td>
<td>8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td><strong>BACTERIA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mycobacterium tuberculosis</td>
<td>19</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Streptomyces coelicolor</td>
<td>72</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Bacteroidetes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteroides thetaiotaomicron</td>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Deinococcus-Thermus</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermus thermophilus</td>
<td>9</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td><strong>Firmicutes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus subtilis</td>
<td>28</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Clostridium difficile</td>
<td>24</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Geobacillus kaustophilus</td>
<td>9</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Lactobacillus casei</td>
<td>32</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Listeria monocytogenes</td>
<td>14</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>29</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Streptococcus pneumoniae</td>
<td>22</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Proteobacteria</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caulobacter crescentus</td>
<td>26</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Rhodopseudomonas palustris</td>
<td>24</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ruegeria sp TM1040</td>
<td>25</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Beta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bordetella pertussis</td>
<td>12</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Neisseria meningitidis</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Delta / Epsilon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacter jejuni</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Myxococcus xanthus</td>
<td>47</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Gamma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erwinia amylovora</td>
<td>19</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>26</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Francisella tularensis</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Haemophilus influenzae</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Legionella pneumophilia</td>
<td>18</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Proteus mirabilis</td>
<td>13</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa</td>
<td>22</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Salmonella enterica</td>
<td>26</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Shigella flexneri</td>
<td>21</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Vibrio cholera</td>
<td>31</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Yersinia pestis</td>
<td>14</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Spirochaetes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia burgdorferi</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 2.6. Diversity in the domain organization of prokaryotic protein acetyltransferases. GNAT protein acetyltransferases characterized to date exhibit different domain organization: (Type I) an N-terminal domain of unknown function homologous to an ADP-forming acyl-CoA synthetase domain fused to a C-terminal GNAT domain; (Type II) an N-terminal GNAT domain fused to a C-terminal domain of unknown function homologous to an ADP-forming acyl-CoA synthetase domain, which contains a GPS motif, a 37-aa long, degenerate proline-rich domain typically found in collagen (250); (Type III) a GNAT domain fused to an N-terminal cAMP-binding domain; (Type IV) a single GNAT domain.
residues) \(N\)-terminal domain, and the catalytic GNAT domain (~200 residues) at the \(C\)-terminus (68, 69). Although the function of the large \(N\)-terminal domain of Pat enzymes remains largely unclear, insights into its relevance to Pat function have been reported (70, 71). Briefly, results obtained from \textit{in vitro} and \textit{in vivo} analyses of single-amino acid \(Se\)Pat variants showed that such variants had low enzymatic activity (70). Furthermore, results from recently reported structural work aimed at understanding the substrate specificity of Pat enzymes suggested that in the absence of the large domain of Pat, the catalytic domain of the enzyme inefficiently interacts with its protein substrate (71). Whatever the role of the \(N\)-terminal domain may be it is likely to also play a role in sensing acetyl-CoA. This conclusion was drawn on the basis of results from isothermal calorimetry experiments, which showed that \(Se\)Pat binds two molecules of acetyl CoA, one binds to the \(N\)-terminal domain, the other one to the catalytic domain (70, 72-77).

Type II. \textit{Streptomyces lividans} encodes a protein acetyltransferase (\(S/l\)PatA) in which the domain order is reversed relative to the domain order observed in \(Se\)Pat, \(Ec\)Pka, and \(Rp\)PatA. That is, in \(S/l\)PatA the GNAT catalytic domain is located at the \(N\)-terminus and the large domain of unknown function is located at the \(C\)-terminus (31). Other members of the actinomycetes, as well as the archaeon \textit{Archaeoglobus fulgidus}, exhibit the same domain organization. Notably, the \(S/l\)PatA large domain contains a proline-rich region that includes a degenerate collagen-like GPS motif. The role of the degenerate collagen-like GPS has not been established, and the presence of this additional feature of \(S/l\)PatA may suggest that regulation of this enzyme is more complicated than the regulation other Pat homologues.

The large domains of Type I and Type II Pat enzymes share homology with ADP-forming acyl-CoA synthetases (Pfam 13380) that catalyze the reaction: free acid + ATP + CoA \(\leftrightarrow\) acyl-CoA + ADP + P. ADP-forming acyl-CoA synthetases have been characterized in archaea.
and protists (75-77). In spite of this homology, no catalytic activity has been attributed to the N-terminal domain of Type I or Type II GNAT acetyltransferases (73, 74).

Type III: *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* each encode a two-domain protein acetyltransferase (*MtPatA* and *MsPatA*, respectively) in which the C-terminal GNAT domain is fused to an N-terminal cAMP-binding regulatory domain (discussed in more detail below).

Type IV: Several prokaryotes encoding single-domain GNAT protein acetyltransferases have been characterized. These enzymes are substantially smaller (~200 residues) compared to the large two-domain Pat homologues (~800-1100 residues), yet they appear to perform similar functions. Single-domain GNATs have been identified and characterized in *B. subtilis* (*BsAcuA*) (30, 78), *R. palustris* (*RpKatA*) (28), and the archaeon *Sulfolobus solfataricus* (*SsPat*) (79). The range of domain architectures and organization in the bacterial and archaeal protein GNATs reveals that lysine acetylation is most likely regulated by diverse signals within these organisms.

**A word of caution about nomenclature of Pat enzymes.** The reader should be cautioned about the misuse of the ‘Pat’ abbreviation being used to name acetyltransferases that do not belong to the Type I, Type II, or Type III protein acetyltransferases. Case in point. The ‘Pat’ enzyme from *Sulfolobus solfataricus* (*SsPat*; PDB 3F8K) is only 160-residues long (73, 74). Such a length is substantially shorter than the typical length of Type I Pat enzymes, which are between 850 and 1100 residues long. The ‘SsPat’ enzyme actually is a Type IV lysine acetyltransferase.

**Known GNAT functions in Escherichia coli and Salmonella enterica.** The model organisms *E. coli* and *S. enterica* encode ~26 GNAT homologues, only half of which have known or predicted functions (Table 2.2). These GNATs target primary amines (80, 81),
<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aat</td>
<td>Leucyl, phenylalanyl-tRNA-protein transferase</td>
<td>(93)</td>
</tr>
<tr>
<td>ArgA</td>
<td>Glutamate acetyltransferase</td>
<td>(86)</td>
</tr>
<tr>
<td>PanM/PanZ</td>
<td>PanD maturation factor</td>
<td>(14)</td>
</tr>
<tr>
<td>PhnO</td>
<td>Aminoalkylphosphonic acid acetyltransferase</td>
<td>(251)</td>
</tr>
<tr>
<td>Pka/YfiQ/Pat</td>
<td>N-Lysine protein acetyltransferase</td>
<td>(24)</td>
</tr>
<tr>
<td>RimI</td>
<td>S18 ribosomal protein acetyltransferase</td>
<td>(89)</td>
</tr>
<tr>
<td>RimJ</td>
<td>S5 ribosomal protein acetyltransferase</td>
<td>(89)</td>
</tr>
<tr>
<td>RimL</td>
<td>L12 ribosomal protein acetyltransferase</td>
<td>(90)</td>
</tr>
<tr>
<td>SpeG</td>
<td>Spermidine acetyltransferase</td>
<td>(84)</td>
</tr>
<tr>
<td>TmcA</td>
<td>tRNA^{Met} cytidine acetyltransferase</td>
<td>(88)</td>
</tr>
<tr>
<td>WecD</td>
<td>dTDP-fucosamine acetyltransferase</td>
<td>(85)</td>
</tr>
<tr>
<td>YhhY</td>
<td>Aminoacyl nucleotide acetyltransferase</td>
<td>(87)</td>
</tr>
<tr>
<td>YncA/MntA</td>
<td>Putative methionine sulfoximine/sulfone acetyltransferase</td>
<td>(91, 252)</td>
</tr>
<tr>
<td>YafP</td>
<td>Possible nitroaromatic compound acetyltransferase</td>
<td>(253)</td>
</tr>
</tbody>
</table>
including the \( N \)-termini of proteins (82, 83), aminoglycoside antibiotics (63), polyamines (84), a nucleotide sugar (85), glutamate (86), toxic aminoacyl nucleotides (87), and transfer RNAs (88). Three GNAT enzymes, RimI, RimJ, and RimL, acetylate the \( \alpha \)-amine group at the \( N \)-terminus of ribosomal proteins S18, S5, and L12, respectively (89, 90). The \( Ec \)YncA homologue of \( S. enterica \) (SeMddA) was recently shown to acetylate and detoxify oxidized methionine derivatives (e.g. methionine sulfoximine, methionine sulfone) (91). Pka is the only identified protein lysine acetyltransferase (Pat in \( S. enterica \)) (24, 92) (discussed in more detail below). Two other \( E. coli \) GNATs appear to modify proteins through unusual mechanisms. Aat is a leucyl, phenylalanyl-tRNA-protein transferase that modifies proteins targeted for degradation through the \( N \)-end rule degradation pathway by the attachment of a leucine or phenylalanine to lysyl and arginyl residues. (93, 94). In \( S. enterica \), a GNAT homologue known as PanM (formerly YhhK) promotes cleavage and maturation of the \( L \)-aspartate-\( \alpha \)-decarboxylase zymogen (pro-PanD), an enzyme of the coenzyme A biosynthetic pathway (14, 15). PanZ, the PanM homologue in \( E. coli \) performs the same function in this bacterium (95, 96).

**LYSINE DEACYTLASES**

\( N \)-Lysine protein acetylation is reversible by protein deacyetylases (referred to as histone deacetylases, or HDACs (Pfam 08295)). Members of four HDAC families catalyze \( N \)-lysine deacylation. The related classes I, II, and IV HDACs do not require cofactors and catalyze lysine deacylation via hydrolysis of the acyl group (Pfam 08295) (55, 97) (Fig. 2.7A). In contrast, class III HDACs, commonly referred to as sirtuins (Pfam 02416), are a mechanistically distinct family of \( \text{NAD}^+ \)-dependent deacyetylases (97, 98) (Fig. 2.7B). Bacteria and archaea typically encode one to two sirtuin homologues, whereas eukaryotes encode several (59, 99) (Table 2.1).
Figure 2.7. Deacetylation mechanisms of HDACs and Sirtuins. (A) HDAC-mediated catalysis is mediated by a histidinyl residue, which acts as a general base and in conjunction with the Zn(II) ion, activates a metal bound water molecule for nucleophilic attack of the substrate carbonyl. The products of the HDAC reaction are deacetylated protein and acetate. (B) Sirtuin-catalyzed deacetylation is initiated by binding of NAD\(^+\) to the catalytic site. The formation of an imidate intermediate occurs through a one-step ADP-ribosylation and inversion of configuration. The products of the sirtuin reaction are the deacylated protein, nicotinamide, and an O-acyl-ADP-ribose product that is derived by mono-ADP-ribosylation of the removed acyl group.
**Zinc-dependent Histone (Lysine) Deacetylases (HDACs)**

HDACs belonging to class I, II, and IV families also deacetylate non-histone protein substrates and therefore are referred to as lysine deacetylases [reviewed in (100, 101)]. The core feature of HDAC structure is an α/β deacetylase fold comprised of 8-stranded parallel β-sheet in which conserved residues coordinate a Zn(II) ion required for catalysis. Catalysis is mediated by a histidinyl residue, which acts as a general base and in conjunction with the Zn(II) ion activates a metal bound water molecule that triggers the nucleophilic attack on a carbonyl group of the substrate. HDACs are unique in their ability to catalyze deacetylation of both protein and small molecule substrates (Fig. 2.7). *R. palustris* (LdaA) and *B. subtilis* (AcuC) proteins are examples of bacterial HDACs that deacetylate proteins, and are discussed below.

**NAD⁺-dependent Sirtuin Deacetylases**

Though described as deacetylases, some sirtuins have depropionylase activity (56, 57, 102). In eukaryotes, certain sirtuins have also been shown to have protein desuccinylation (103, 104) and demalonylation (105) activity, and are thought to play a critical role in mitochondrial metabolism [reviewed in (106, 107)]. The ability of sirtuins to deacylate a variety of modifications correlates with the ability of their partner GNAT acetyltransferases to utilize alternative acyl-CoA donors. This provides a mechanism with a wider range of modifications available to the cell to maintain homeostasis under diverse physiological conditions in response to environmental changes.

**Sirtuins hydrolyze NAD⁺.** Due to their requirement for NAD⁺, sirtuin deacetylases have garnered a great deal of interest for their ability to ‘sense’ and respond to NAD⁺ levels, which in turn reflect on the cellular energy status (108, 109). Deacetylation is an energetically favorable process and can be catalyzed by the Class I, II and IV deacetylases without cofactors. It is
therefore of interest that sirtuins (Class III) hydrolyze NAD$^+$ in the course of the deacetylation reaction, an essential metabolic cofactor (110, 111). The resynthesis of NAD$^+$ from the hydrolysis products requires \( \sim 8.2 \text{ kcal/mol} \) of energy and is energetically expensive (110, 112). It stands to reason that there is a compelling reason to tie NAD$^+$ levels to sirtuin activity (i.e. protein acetylation state) (109).

**Sirtuin reaction mechanism.** Sirtuin deacetylases have an unusual catalytic mechanism that uses NAD$^+$ not as a cofactor, but as a co-substrate that is cleaved during the deacylation reaction, yielding \( O\text{-acyl-ADP-ribose (O-AADPR)} \) (97, 113-116). Sirtuin-catalyzed deacetylation is initiated by binding of NAD$^+$ to the catalytic site. The formation of an imidate intermediate occurs through a one-step ADP-ribosylation and inversion of configuration [reviewed in (117)] (Fig. 2.7). The inversion of configuration was predicted by the original discovery that sirtuins were enzymes that used pyridine nucleotides as substrates (98). The reaction of the coenzyme B$_{12}$ biosynthetic pathway in *S. enterica* that is also performed by the CobB sirtuin proceeds via a nucleophilic attack inverting the configuration of the \( N\)-glycosidic bond between the base and the ribosyl moiety of the pyridine nucleotide co-substrate (117). The products formed from the deacetylation reaction are (i.) the deacetylated protein, (ii.) nicotinamide, and (iii.) \( O\text{-acyl-ADP-ribose} \). The \( O\text{-acyl-ADP-ribose} \) by-product is generated through the mono-ADP-ribosylation of the removed acyl group (102, 116, 118, 119).

**Physiological importance of \( O\text{-acetyl-ADP-ribose} \).** The physiological role of \( O\text{-AADPR} \) is unknown in prokaryotes, but some information about its metabolism has been reported in eukaryotic systems. In eukaryotes, \( O\text{-AADPR} \) may act as a signaling molecule and may regulate gene silencing, ion channel gating and redox regulation (118, 120). There are several eukaryotic enzymes identified to utilize \( O\text{-AADPR} \) as a substrate including two NUDIX (Nucleoside
Diphosphate linked to X hydrolases (Ysa1 from yeast; NudT5 from mouse) (121), ADP-ribosyl hydrolase (ARH3, human) (122), and two uncharacterized enzyme activities including an esterase and nuclear acetyltransferase (from yeast and human) (121). These enzymes utilize O-AADPR in various ways, generating O-acetyl-ribose-phosphate and AMP (Ysa1 and NudT5), acetate and ADP-ribose (ARH3, esterase), or an unknown acetylated product and ADP-ribose (nuclear acetyltransferase). More work is needed to elucidate the role of O-AADPR in prokaryotic physiology.

**Effect of the NAD⁺:NADH ratio on sirtuin function.** Both NADH and nicotinamide have been reported to inhibit sirtuin function, and may play a role in regulation of sirtuin activity (123-125). Nicotinamide condenses with an ADPr-like intermediate formed during the deacetylation reaction, which prevents the reaction from moving forward (109), and inhibits sirtuin activity non-competitively at concentrations consistent with physiological nicotinamide levels (30–200 µM) (113, 126-130).

In contrast, NADH competitively inhibits sirtuin activity with reported $K_i$ values in the sub-millimolar to millimolar range, consistent with physiological NADH levels (97, 108, 123, 131, 132). Sirtuin activity varies with the NAD⁺:NADH ratio. A high NAD⁺:NADH ratio indicates that the cell is efficiently oxidizing NADH back to NAD⁺, implying among other things, that a strong proton motive force is being generated under such conditions. A strong proton motive force results in high ATP levels and the concomitant increase in the demand for acetyl-CoA for anabolic purposes. If under such conditions acetate is present in the environment, the cell can activate it to acetyl-CoA using the acetyl-CoA synthetase (Acs), an enzyme known to be under RLA control (discussed further below) (23). Acetylated Acs is inactive, hence, the ratio of acetylated (inactive):deacetylated (active) Acs would be expected to be low when the
NAD⁺:NADH ratio is high, since NAD⁺ would be available for sirtuin to deacetylate (i.e., activate) acetylated Acs (132). The NAD⁺:NADH ratio is known to fluctuate with metabolism, thus changes in this pool of free NAD⁺ likely regulate protein acylation state, with deacylation occurring when the NAD⁺:NADH is high, and acetylation being favored when the NAD⁺:NADH ratio is low.

**Overview of sirtuin structures.** Three-dimensional crystal structures of sirtuins from all domains of life are available (e.g., *Archaeoglobus fulgidus*, PDB 1YC2; *Thermatoga maritima*, PDB 4BUZ; *E. coli*, PDB 1S5P; *Saccharomyces cerevisiae*, PDB 2HJH; and human SIRT1, 4KXQ; SIRT2, 3ZGO; SIRT3, 4BN4; SIRT5, 3RIY; SIRT6, 3PKI) (Fig. 2.5C, D). Sirtuins are comprised of a catalytic domain that contains a Rossmann fold domain and a variable Zn(II)-binding domain, with divergent N- and C- terminal regions [reviewed in (133, 134)]. The Zn(II) ion in sirtuins is structurally important and does not contribute to catalysis (135). The binding sites for the nicotinamide and ribose moieties of NAD⁺ and the acetyllysine substrate are located in the cleft between the large (Rossmann fold) and small (Zn(II)-binding) domains (Fig. 2.5C, D). This binding cleft allows for substrate selectivity among different sirtuins. Importantly, NAD⁺ is oriented opposite to the typical orientation seen with Rossmann fold-containing enzymes, in which the nicotinamide moiety binds to the N-terminal half of the β-sheet and the adenine binds to the C-terminal half. This orientation reversal ensures the elimination of nicotinamide.
HIGH-THROUGHPUT IDENTIFICATION OF ACETYLATED PROTEINS

Global Approaches for the Identification of Total Acetylated Protein Population

Characterization of the total acetylated protein population (the “acetylome”) in a given organism has been accelerated by the development of sensitive mass spectrometry-based methods that detect the precise location of acetylated lysine residues within any given protein pool (33). Such approaches have putatively identified a large number of acetylation targets, many of which appear to be acetylated at several sites. Early studies combined the use of two-dimensional separation of proteins followed by detection of acetyllysine residues by immunoblotting (136). Subsequent studies characterizing the acetylomes of eukaryotes used immunoprecipitation as a way to enrich for acetyllysine peptides present in tryptic digests of a protein pool (137-139). The combined use of anti-acetyllysine antibodies and mass spectrometry are the foundation for the current proteomic methodology used to identify acetylated peptides and proteins (Fig. 2.8A). These global studies have been extended to archaea (N-terminal acetylation) (140), bacteria (141-143), plants (144), parasites (145, 146), and humans (147), helping cement the contributions of acetylation to cellular physiology. Such global studies suggest that acetylation controls diverse processes including metabolism, transcription, translation, and cell structure.

Alternative global approaches for the analysis of acetylomes. Proteome microarrays provide an alternate method for identifying the putative targets of acetyltransferases as well as other modifying enzymes. Protein chip technology was first reported in yeast for the analysis of protein kinases (148). Lin et al used a similar approach to identify 13 substrates of the NuA4 acetyltransferase, including the phosphoenolpyruvate carboxykinase (Pck1p) enzyme (35). This same experimental method was used in another study to construct a microarray comprised of
Figure 2.8. Methods for the analysis of acetylomes. (A) A representative workflow of the methodology typically used to determine total acetylated protein from an organism. (B) A representative workflow of a recently described method to determine the level of acetylation of identified acetylated target proteins.
E. coli proteins. These microarrays were probed with the S. enterica bacterial protein acetyltransferase (SePat) enzyme, resulting in the identification of seven substrates, including several transcription factors (149).

Other studies have taken advantage of labeling strategies like Stable Isotope Labeling with Amino acids in cell Culture (SILAC) quantitative mass spectrometry (150). SILAC allows for detection of differences in protein abundance using in vivo incorporation of non-radioactive isotope labels. Benefits of this analysis are the identification and quantification of relative differential changes in complex protein samples.

Recent approaches have begun to tackle the problem of the stoichiometry of the level acetylation at individual sites in order to understand the biological significance of acetylation events (151, 152). The first method for determining site-specific stoichiometry of acetylated peptides using no immunoenrichment was used to investigate the acetylome of E. coli (151). This was achieved by chemically acetylating unmodified lysine residues using acetic anhydride labeled with stable isotopes to generate an acetyllysine pair and analyzed by mass spectrometry (Fig. 2.8B). In this scenario, proteins endogenously acetylated contain ‘light’ acetyllysine residues, while proteins modified chemically contain ‘heavy’ acetyllysine residues. After resolution by mass spectrometry, the stoichiometry is determined by examining the ratio of light to heavy peak areas. By this method the authors identified proteins that use or generate acetyl-CoA, and those involved in transcription and translation, are the most highly acetylated (151).

Bacterial Acetylome Studies

Bacterial acetylomes have been characterized in E. coli (141, 142), S. enterica (153), Bacillus subtilis (154), Erwinia amylovora (155), R. palustris (28), Staphylococcus aureus (156),
Geobacillus kaustophilus (157), Vibrio parahaemolyticus (158), Thermus thermophilus (159), and M. tuberculosis (143), Mycoplasma pneumonia (160), and Streptomyces roseosporus (161). These studies have identified a range of 62-667 putatively acetylated proteins per organism, with the majority of acetylated proteins involved in central metabolism and translation (141, 156, 158) (Fig. 2.9).

The large number of putative acetylation targets detected by mass spectrometry has raised important questions as to the significance of the detected acetylation events, and how they contribute to cellular physiology. Specifically, it is important to know how these acetylation events occur, at what frequency, and whether they affect protein function or stability. Two independent studies of acetylated proteins in E. coli identified 85 and 91 putative acetylation targets (141, 142). Surprisingly, only six proteins acetylated at the same lysine residue were identified in both studies (60). Since then a third acetylome study has been performed in E. coli identified 349 acetylated proteins (162). In the related enterobacterium S. enterica, Wang et al found 191 putatively acetylated proteins, many of which were metabolic enzymes (153). Although experimentally validating some of the results in vitro, studies by others groups were not able to reproduce the findings (28).

A combination of mass-spectrometry analysis, in vivo genetics analyses, and in vitro validation for the Gram-negative photosynthetic bacterium R. palustris has yielded the most comprehensive list of bona fide acetylation targets to-date (28). In this study, acetylated proteins were identified by tandem mass spectrometry by comparing the acetylome of a wild-type R. palustris strain to the acetylomes of strains in which one or more acetyltransferases were absent. Stringent cut-offs were applied to reduce noise by using two different algorithms. Proteins that
### Functional Annotation of Acetylated Proteins

<table>
<thead>
<tr>
<th>Organism</th>
<th>Protein(^{Ac})</th>
<th>% Total(^{Ac})</th>
<th>Validation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bacillus subtilis</em></td>
<td>185</td>
<td>4.5%</td>
<td>N</td>
<td>(154)</td>
</tr>
<tr>
<td><em>Erwinia amylovora</em></td>
<td>78*</td>
<td>2.2%</td>
<td>N</td>
<td>(155)</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>349</td>
<td>8.5%</td>
<td>N</td>
<td>(142)</td>
</tr>
<tr>
<td><em>Geobacillus kaustophilus</em></td>
<td>114</td>
<td>3.1%</td>
<td>N</td>
<td>(157)</td>
</tr>
<tr>
<td><em>Mycobacterium tuberculosis</em></td>
<td>137</td>
<td>3.4%</td>
<td>N</td>
<td>(143)</td>
</tr>
<tr>
<td><em>Mycoplasma pneumoniae</em></td>
<td>218</td>
<td>33.6%</td>
<td>N</td>
<td>(160)</td>
</tr>
<tr>
<td><em>Rhodoseudomonas palustris</em></td>
<td>62</td>
<td>1.3%</td>
<td>Y</td>
<td>(28)</td>
</tr>
<tr>
<td><em>Salmonella enterica</em></td>
<td>191</td>
<td>4.2%</td>
<td>Y</td>
<td>(153)</td>
</tr>
<tr>
<td><em>Staphylococcus aureus</em></td>
<td>412*</td>
<td>9.7%</td>
<td>N</td>
<td>(156)</td>
</tr>
<tr>
<td><em>Streptomyces roseosporus</em></td>
<td>614*</td>
<td>9.1%</td>
<td>N</td>
<td>(161)</td>
</tr>
<tr>
<td><em>Thermus thermophilus</em></td>
<td>128</td>
<td>5.7%</td>
<td>N</td>
<td>(159)</td>
</tr>
<tr>
<td><em>Vibrio parahemolyticus</em></td>
<td>656</td>
<td>13.6%</td>
<td>N</td>
<td>(158)</td>
</tr>
</tbody>
</table>

*Table legend:*
- Metabolic pathways
- Unknown function
- Transport & cell division
- Translation
- Replication & repair
- Transcription
- Stress & adaptation

**Figure 2.9. Comprehensive overview of bacterial acetylome studies.** Functional annotation of identified acetylated proteins from bacterial acetylome studies. Protein\(^{Ac}\): number of identified acetylated proteins; % Total\(^{Ac}\): the percent of acetylated proteins of the entire proteome; Validation: N, no; Y, yes. *Values are different than previously reported; however, the values listed reflect the available data were obtained from supplementary information from the cited references. The percentage scale at the top of the figure should be used to estimate the percentage of acetylated proteins in each of the categories within any given microorganism in the figure.
were identified in this comparison were validated by both \textit{in vitro} and \textit{in vivo} methods. The authors also confirmed that acetylation altered the activity of each of target protein.

\textbf{Current noteworthy issues in the field.} Bioinformatics analyses reveal that protein acetyltransferases are conserved in nearly all reported genomes. This suggests that acetylation is widespread in prokaryotes and eukaryotes, which in turn implies that protein acetylation is not limited to the regulation of proteins involved in DNA maintenance or transcription. The acetylome studies provide a framework for the identification of putative targets of acetylation, but detailed mechanistic studies are needed to validate these proteomics-based results and to demonstrate the findings are biologically relevant. There are several issues that need to be taken into consideration, and are discussed below.

\textit{(i) Global, non-enzymatic acetylation by acetyl-phosphate (AcP).} Recent work in \textit{E. coli} suggests that global, non-enzymatic, low-level lysine acetylation is also mediated by the reactive, high-energy metabolite acetyl-phosphate (AcP), and that this acetylation event is globally regulated by growth phase and metabolism (81). Weinert \textit{et al} compared the level of protein acetylation from \textit{E. coli} cells at different growth phases using SILAC quantitative mass spectrometry (81). The analysis revealed that the bulk level of protein acetylation was dramatically increased in stationary phase. An increase in protein acetylation was also strongly correlated with an increase in AcP levels. Taken together, these results suggest that AcP may be directly involved in widespread, growth phase-dependent chemical (non-enzymatic) acetylation of \textit{E. coli} proteins. The physiological significance of this phenomenon is not understood (81). Further support for the role of AcP as a non-specific donor of acetyl groups in lysine acetylation in \textit{E. coli} was recently reported (152). Direct acetylation by acetyl-CoA has also been suggested as a mechanism of non-enzymatic acetylation in eukaryotes (163).
(ii) Validation of proteomic approaches. Because of its broad distribution in nature, RLA elicited a great deal of interest amongst biologists eager to define the role of RLA in cell physiology. However, several issues need to be addressed before we understand whether or not the function of specific proteins is under RLA control. Issues that need clarification are:

(a) Non-enzymatic protein acetylation. Chemical, non-enzymatic lysine acetylation can occur when pH is $\geq 8.0$ (151, 164, 165). The autoacetylation activity of some proteins (166) further compounds this problem, leading to reporting of false positive enzyme-driven acetylation. Therefore, multiple controls must be used to distinguish between autoacetylation and enzyme-driven acetylation, namely the reaction substrates with the addition of inactive variants of the modifying acetyltransferase enzyme.

(b) Physiological relevance of multiple acetylation sites. Many of the reported acetylomes identify proteins with multiple acetylated lysine residues (~5-10 in some cases). This information contrasts sharply with observations for validated RLA targets such as acetyl-CoA synthetase (Acs) and other members of the acyl-CoA synthetases (discussed further below), in which the acetylation of a single lysine residue is necessary and sufficient to alter enzyme function (23, 28, 167). Recently, the first validated example of an Acs homologue from *Saccharopolyspora erythraea* that undergoes multiple acetylation events was reported, and the observations were validated *in vitro* and *in vivo* (168). However, of the four original acetylation events, only two were shown to have an effect on enzyme activity, highlighting the importance of *in vivo* and *in vitro* validation. It is imperative to determine the effect that each acetylated lysine may have on substrate binding, catalytic activity, or resistance to proteolytic activity to gain insights into the physiological importance of RLA.
(c) Possible over-representation of central metabolic enzymes. The fact that the majority of acetylated peptides identified are involved in central metabolism, specifically glycolysis, has been perhaps overemphasized. Notably, membrane-associated proteins, proteins present at low levels, or proteins not expressed under the culture conditions tested may represent important groups of targets under the control of RLA which are missed by the current methodology.

(d) Validation. Very limited, and in some cases no validation of results obtained by mass spectrometry analyses has been reported in studies of bacterial acetylomes. Specifically, the function of the proteins identified as acetylation targets have not been analyzed in vitro to identify the modifying enzymes, or whether or not acetylation alters protein function or stability. At present, large-scale mass spectrometry results of bacterial acetylomes provide only putative targets until further validation is performed.

VALIDATED REVERSIBLE LYSINE ACYLATION (RLA) TARGETS IN BACTERIA AND ARCHAEA

Discovery of RLA in Prokaryotes

The role of RLA in bacteria, specifically in relation to acetylation of metabolic enzymes, was discovered in *S. enterica*. Refer to Table 2.3 for a to-date comprehensive list of validated RLA targets in prokaryotes. The protein acetyltransferase of *S. enterica*, SePat, was first identified as the enzyme responsible for the acetylation and inactivation of acetyl-CoA synthetase (Acs), the enzyme that activates acetate when present in the environment at low concentrations ($\leq 10$ mM) (24). Acs belongs to the acyl-CoA synthetase family (Pfam 00501), which convert acetate to acetyl-CoA (24). Acyl-CoA synthetases are ubiquitous across all domains of life (169-171).
Table 2.3. Validated substrates of prokaryotic lysine acetyltransferases

<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bacillus subtilis</em> AcuA</td>
<td>AMP-forming acetyl-CoA synthetase</td>
<td>(30)</td>
</tr>
<tr>
<td>AcsA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em> Pka</td>
<td>AMP-forming acetyl-CoA synthetase</td>
<td>(177)</td>
</tr>
<tr>
<td>Acs</td>
<td>AMP-forming acetyl-CoA synthetase</td>
<td></td>
</tr>
<tr>
<td>RNase R</td>
<td>Stable RNA exoribonuclease</td>
<td>(68, 92)</td>
</tr>
<tr>
<td>ResB</td>
<td>Response regulator for capsule synthesis</td>
<td>(149)</td>
</tr>
<tr>
<td><em>Mycobacterium smegmatis</em> PatA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSMEG_4207</td>
<td>Universal stress protein</td>
<td>(191)</td>
</tr>
<tr>
<td>Acs</td>
<td>AMP-forming acetyl-CoA synthetase</td>
<td></td>
</tr>
<tr>
<td>FadD2, FadD4, FadD5, FadD10, FadD12, FadD13, FadD22, FadD35</td>
<td>AMP-forming acyl-CoA synthetases</td>
<td>(178)</td>
</tr>
<tr>
<td><em>Rhodopseudomonas palustris</em> KatA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acs</td>
<td>AMP-forming acetyl-CoA synthetase</td>
<td>(28)</td>
</tr>
<tr>
<td>PrpE</td>
<td>AMP-forming propionyl-CoA synthetase</td>
<td>(28)</td>
</tr>
<tr>
<td><em>Rhodopseudomonas palustris</em> Pat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acs</td>
<td>AMP-forming acetyl-CoA synthetase</td>
<td>(69)</td>
</tr>
<tr>
<td>PrpE</td>
<td>AMP-forming propionyl-CoA synthetase</td>
<td>(28)</td>
</tr>
<tr>
<td>BadA, HbaA, AliA</td>
<td>AMP-forming aromatic and alicyclic acyl-CoA synthetases</td>
<td>(69)</td>
</tr>
<tr>
<td>PimA</td>
<td>AMP-forming pimeloyl-CoA synthetase</td>
<td>(28)</td>
</tr>
<tr>
<td><em>Salmonella enterica</em> Pat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acs</td>
<td>AMP-forming acetyl-CoA synthetase</td>
<td>(23, 169)</td>
</tr>
<tr>
<td>PrpE</td>
<td>AMP-forming propionyl-CoA synthetase</td>
<td>(56)</td>
</tr>
<tr>
<td><em>Streptomyces lividans</em> PatA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AacS</td>
<td>AMP-forming acetoacetyl-CoA synthetase</td>
<td>(31)</td>
</tr>
<tr>
<td><em>Sulfolobus solfataricus</em> Pat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alba</td>
<td>Chromatin protein</td>
<td>(59)</td>
</tr>
</tbody>
</table>
This family of enzymes converts weak organic acids to their CoA thioesters through two half reactions via an adenylated intermediate (see reaction below) (172) (Fig. 2.10A).

\[
\text{Acetate} + \text{ATP} + \text{Acs} \rightarrow \text{Acetyl-AMP} + \text{PP}_i + \text{Acs} \quad \text{Eq. [1]}
\]

\[
\text{Acetyl-AMP} + \text{CoA} + \text{Acs} \rightarrow \text{Acetyl-CoA} + \text{AMP} + \text{Acs} \quad \text{Eq. [2]}
\]

As mentioned above, \( \text{SeAcs} \) is required during growth at low concentrations of acetate (\( \leq 10 \) mM). When the concentration of acetate in the environment is >10 mM, assimilation of acetate occurs via the phosphotransacetylase / acetate kinase (Pta / Ack) pathway (173).

\( \text{SeAcs} \) activity is regulated by reversible acetylation of an active site lysine residue, Lys609 (23). This discovery opened the doors to numerous studies on regulation of metabolic enzymes by acetylation and shifted the focus from regulation of histones by acetylation to acetylation as a means to regulate metabolism and physiology.

\( \text{SeAcs} \) activity was first predicted to be under the control of acetylation was when it was observed that the sirtuin deacetylase (\( \text{SeCobB} \)) was needed in growth conditions in which acetate (10 mM) or propionate (30 mM) was the sole source of carbon and energy (23, 29). In support of this idea, the \( \text{SeCobB} \) protein was shown to deacetylase Acs \( \text{in vitro} \) (23), resulting in reactivation of the Acs enzyme. These findings revealed for the first time that acetylation was a means to modulate activity of a metabolic enzyme (23). The modifying protein acetyltransferase of this bacterium, \( \text{SePat} \), was subsequently discovered by selecting for derivatives of a \( \Delta \text{cobB} \) strain that grew on 10 mM acetate (24). Deletion of the protein acetyltransferase (\( \text{SePat} \)) in a strain lacking \( \text{SeCobB} \) restored growth on both acetate (10 mM) and propionate (30 mM). \( \text{In vitro} \) studies demonstrated that \( \text{SePat} \) could both acetylate \( \text{SeAcs} \) and propionylate the propionyl-CoA synthetase (PrpE) (24, 56).
Figure 2.10. Synthesis of acyl-CoAs by AMP- and ADP-forming acyl-CoA synthetases. The AMP-forming and ADP-forming acyl-CoA synthetases convert organic acids to their CoA thioesters. AMP-forming acyl-CoA synthetases perform this reaction through two half reactions via an adenylated intermediate (A), while the ADP-forming acyl-CoA synthetase reaction is driven by the energy of hydrolysis of the γ-phosphate of ATP (B).
The acetylated lysine residue (Lys609 in SeAcs), is required for adenylylation of organic acids, and is universally conserved in the AMP-forming acyl-CoA synthetases. Acetylation of this lysyl residue prevents adenylylation of the acid substrate and blocks enzyme activity. Thus, SePat acetylation of the SeAcs active site lysine blocks the conversion of acetate to acetyl-AMP (23), the first half-reaction catalyzed by SeAcs. It is proposed that the lysyl residue (i) aids in the orientation of the carboxylate moiety of the acid and phosphoryl groups of ATP for the in-line attack, and (ii) stabilizes the transition state through the positive charge interactions (174, 175).

**Acetylation of acetyl-CoA synthetase (Acs) is a conserved regulatory mechanism.** Since its discovery, it has been shown that RLA controls Acs and other members of the acyl-CoA synthetases via acetylation, and that this regulatory mechanism of metabolism is present in prokaryotes and eukaryotes (29, 176). Acetylation of Acs homologues by SePat homologues has also been demonstrated in *E. coli* (141, 177), *R. palustris* (69), *B. subtilis* (30), *Streptomyces* spp. (31), and *Mycobacterium* spp. (178). Insight into the need for post-translational control of Acs activity was provided in studies that demonstrated dysregulation of this enzyme caused a severe imbalance in the energy charge of the cell, leading to growth arrest (179). An expanded explanation of these effects is described below.

**RLA in Gram-negative Bacteria**

**Non-acyl-CoA synthetase targets of the Salmonella enterica protein acetyltransferase.** It has been suggested that SePat can acetylate a variety of metabolic enzymes, including glyceraldehyde-3-phosphate dehydrogenase (GapA) and isocitrate lyase (AceA), and the glyoxylate shunt regulator isocitrate dehydrogenase kinase/phosphatase (AceK) (153). However, difficulties in reproducing these results by others have been reported (28). These discrepancies
need to be clarified before any conclusions about the involvement of RLA in the modulation of such key enzymes can be validated.

In *Escherichia coli*, the protein acetyltransferase (Pka) alters the fate of RNase R. Results from early studies suggested that the Pat homologue from *E. coli*, EcPka, acetylated the RNA polymerase alpha subunit; however, results from subsequent analysis failed to confirm this claim (180, 181). A *bona fide* substrate of EcPka is exoribonuclease RNase R, shown to degrade highly structured mRNA. RNase R is acetylated during exponential phase, destabilizing the protein, making it prone to proteolytic degradation (68, 92). This is the first example of acetylation by a *SePat* homologue affecting protein stability and thus enzyme activity. Notably, RNase R cannot be deacetylated by the *E. coli* sirtuin CobB, the only known protein deacetylase in this bacterium, and therefore is not a reversible modification (68).

In *Rhodopseudomonas palustris*, RLA controls the activity of many AMP-forming acyl-CoA synthetases. The role of RLA in the physiology of the purple non-sulfur photosynthetic *α*-proteobacterium *R. palustris* has been investigated (28). In this bacterium, RLA modulates the activity of enzymes involved in the anaerobic catabolism of aromatic organic acids and other fatty acids. Results from proteomics global analyses indicate that *RpPat* acetylates many acyl-CoA synthetases (AMP-forming) in addition to acetyl-CoA synthetase (Acs) and propionyl-CoA synthetase (PrpE), including those involved in activating short-, medium- and long-chain fatty acids and aromatic acids (*e.g.*, BadA, HbaA, AliA, PimA, HcsA, FadD, FcsA, LcsA, and IbuA) (28). It is significant that all of the verified *RpPat* substrates from this study were acyl-CoA synthetases. This data strongly suggests that *RpPat* may specifically recognize and regulate this class of enzymes via a negative feedback mechanism.
When an RpPat substrate, benzoyl-CoA synthetase (BadA) was acetylated in vivo in the absence of RpPat, Crosby et al predicted the existence of a second R. palustris protein acetyltransferase (69). A single-domain GNAT, RpKatA (for K (Lys) acetyltransferase A), was identified based on its limited sequence homology (33% identical over 64 residues within the GNAT domain) to other known protein acetyltransferases. RpKatA also acetylated the conserved catalytic lysine of acyl-CoA synthetases whose substrates included short-, medium-, long-, and branched-chain fatty acids in addition to aromatic organic acids (28, 69). It is noteworthy that although RpKatA has similar enzymatic capabilities as RpPat (Type I GNAT), it is a much smaller protein comprised only of a catalytic domain (Type IV GNAT).

Insights into the role of RpKatA were obtained by performing mass spectrometry-based proteomics analysis. Briefly, four acyl-CoA synthetases (BadA, AliA, HbaA, and PrpE) were acetylated in a pat deletion strain, but no acetylation of these proteins was seen in a pat katA double mutant (28). From in vitro studies, the authors learned that RpPat and RpKatA had different substrate specificities for the acyl-CoA synthetases of R. palustris. For example, there are two acyl-CoA synthetases, hexanoyl-CoA synthetase A (HcsA) and long-chain acyl-CoA synthetase A (LcsA), which were acetylated by RpPat but were not acetylated by RpKatA (28). It is not known how RpKatA and RpPat recognize their substrates or why there is overlapping activity between the enzymes.

In addition to having two protein acetyltransferases, R. palustris also encodes two protein deacetylases, a sirtuin deacetylase (SrtN) and a Zn(II)-dependent lysine deacetylase (LdaA). Genetic evidence suggests that both deacetylases play a role in regulating acyl-CoA synthetases in R. palustris (28).
RLA in Gram-positive Bacteria

**RLA controls acetyl-CoA synthetase activity in Bacillus subtilis.** In *B. subtilis*, the GNAT *BsAcuA* (Type IV protein acetyltransferase) is comprised of only a single GNAT domain, and has no significant sequence homology to *SePat* (Type I protein acetyltransferase). However, a structure of a *BsAcuA* homologue resolved from *Exiguobacterium sibiricum* demonstrated that *BsAcuA* does contain the conserved GNAT domain (PDB 2Q04) (78). *BsAcuA* acetylates and inactivates acetyl-CoA synthetase A (AcsA) at the conserved catalytic lysine (Lys549) (30). Deacetylation of *BsAcsA*Ac can occur by either of the two protein deacetylases of *B. subtilis*, the Zn(II)-dependent *BsAcuC* deacetylase and / or the sirtuin *BsSrtN* (30, 182).

The genes encoding *BsAcuA* (GNAT) and *BsAcuC* (deacetylase) are located within the *acuABC* operon (183), and are divergently transcribed from *acsA* (target). It was initially thought that the *acuABC* operon was involved in acetoin utilization, as deletion of *acuA* caused a growth defect when cells were grown on acetoin (183). However, it has since been shown that the acetoin utilization pathway is encoded by *acoABCLR* in *B. subtilis* (184). At present it is unclear if or how *BsAcuB* is involved in the *BsAcuA*- and *BsAcuC*-dependent regulation of *BsAcsA*, or if acetoin utilization is either directly or indirectly regulated by acetylation.

**Acs from Streptomyces lividans is the exception to the Acs acetylation paradigm.** Metabolic regulation in actinomycetes, like *Streptomyces*, is of interest because of the diverse natural products they produce (185-188). *S. lividans* encodes a protein acetyltransferase, *SlPatA* (Type II), which also acetylates AMP-forming acyl-CoA synthetases, including *S. lividans* acetoacetyl-CoA synthetase (*SlAacS*) and *S. enterica* Acs (31). *SlPatA* only weakly modified the *S. lividans* Acs homologue. However, it efficiently acetylated the related enzyme, *SlAacS*, both *in vitro* and *in vivo* (31). Aacs is present in all domains of life, and this work provided the first...
example of regulation of its activity by acetylation. Recently, the structure of \( SL\)AacS was reported and the structure provided for the first time an ordered view of the 30-residue extension of the C-terminus of this type of enzyme, and it was suggested that such an extension may interact with catalytic residues of the \( N\)-terminal domain (189). A comparison of the \( SL\)Acs and \( SL\)AacS would provide valuable insights into the determinants that make AMP-forming acyl-CoA synthetases good substrates for the \( SL\)PatA enzyme.

\( SL\)PatA is the first characterized Pat homologue that does not efficiently acetylate its cognate Acs enzyme in vitro (31), suggesting that \( SL\)PatA may not be the enzyme responsible for Acs acetylation in \( S.\) lividans. Alternatively, acetylation of \( SL\)Acs by \( SL\)PatA may require additional factors that are not required by \( SC\)Pat and \( RP\)Pat for acetylation of Acs orthologues from those organisms, or simply \( SL\)Acs is not enzymatically acetylated.

\( S.\) lividans encodes two sirtuin deacetylases, CobB1 and CobB2, and a Zn(II)-dependent AcuC-type deacetylase. Work in the closely related \( Streptomyces\) coelicolor demonstrated that Acs was acetylated and that CobB1 deacetylated Acs in vitro (190). However, the acetyltransferase responsible for acetylation of \( S.\) coelicolor Acs was not identified.

Recently, studies showed that in another actinomycete, \( Saccharopolyspora\) erythraea, a homologue of the GNAT-related AcuA enzyme from \( B.\) subtilis, acetylates the \( S.\) erythraea Acs enzyme at four different positions, and that in a mutant of \( S.\) erythraea lacking the \( Sa\)AcuA enzyme, \( Sa\)Acs is not acetylated at all (168). The \( Sa\)AcuA enzyme is the first of its class to be experimentally shown to target several residues of an Acs homologue, as described above. Whether AcuA homologues in other actinomycetes are responsible for the single or multiple acetylation of Acs in this class of microorganisms remains to be determined.
In *Mycobacterium* spp. the universal stress protein USP is under RLA control. *M. tuberculosis* and *M. smegmatis* encode unique protein lysine acetyltransferases (*MtPatA* and *MsPatA*, respectively). In these organisms, the GNAT domain is attached to a cyclic AMP (cAMP) binding domain (Type III) (178, 191). cAMP allosterically activates *MtPatA* and *MsPatA*, enhancing their activity >2-fold (178, 191-194).

*MsPatA* acetylates a universal stress protein (USP, MSMEG_4207) at a single lysine residue, and acetylation increases in the presence of cAMP (191). The in vivo significance of USP acetylation was not tested, likely because the function of most USPs is unclear, but there is evidence suggesting that USPs provide resistance to various stressors [reviewed in (195)].

*MsPatA* and *MtPatA* share 57% identity, thus they may have similar substrates, except for USP, which is not conserved in mycobacteria that encode homologues of *MsPatA*. Based on these findings, Xu et al used *MsPatA* to acetylate whole-cell lysates with or without the acetyl-CoA analogue chloroacetyl-CoA. The authors identified *MsAcs* as an acetylated protein target of *MsPatA* (178). Because *MsAcs* could not be overproduced in *E. coli*, in vitro experiments aimed at showing acetylation of *MsAcs* by *MsPatA* were not performed (178). Using an alternative approach, the authors purified *MtAcs*, (76% identical to *MsAcs*), and demonstrated *MsPatA*-dependent acetylation of *MtAcs* at the expected catalytic lysine residue, which abolished *MtAcs* activity (178). Although the degree of identity between *MtPatA* and *MsPatA* is high, there is precedent in the literature of Acs enzymes that are very poorly acetylated by Pat enzymes (e.g., *SlAcs*). Therefore, in the absence of experimental evidence, it is premature to conclude that *MsAcs* is a substrate of *MsPatA*.

An independent study identified eight additional acyl-CoA synthetases as substrates of *MsPatA* (192). The single protein deacetylase in *M. tuberculosis*, an NAD⁺-dependent sirtuin
homologue (MRA_1161, from H37Ra), deacetylated MtAcs and all eight acyl-CoA synthetases in vitro (178, 192), suggesting that that this likely constitutes a regulatory system comparable to the RLA systems found in S. enterica and R. palustris (196). The authors were able to demonstrate that acetylation in M. tuberculosis was dependent upon intracellular cAMP levels by examining the acetylation level of known targets in conditions with varying cAMP concentrations. Acetylation of the acyl-CoA synthetase targets was seen only in conditions with higher levels of cAMP (178, 192).

**RLA in Archaea**

The first studies examining protein acetylation and deacetylation in archaea were performed in *Halobacterium salinarum* (previously *H. halobium*) in which a 2Fe-2S ferredoxin protein was identified to be acetylated at a specific lysine residue (Lys118) (197). While this was the first example of protein acetylation in archaea, no further studies of acetylated proteins involved in processes other than gene expression have been reported.

**Acetylation of the Sulfolobus solfataricus chromatin protein (ALBA).** As mentioned above, little is known about acetylation in archaea, even though many archaeal species have acetyltransferase and deacetylase homologues. Some archaea encode histone proteins similar to eukaryotes. However, the archaeal histones differ in that they do not contain the flexible N-terminal tails and are not post-translationally modified (198, 199). In addition to histones, archaea have another chromatin protein known as Sso10b or Alba (for acetylation lowers binding affinity). When bound to DNA, Alba inhibits transcription (200, 201).

Investigators discovered that Alba homologues purified from *Sulfolobus* spp. were 84-Da larger in mass than predicted. Because an acetyl group adds 42 Da to the protein mass, it was
suggested that the Alba proteins were acetylated at two sites (200, 202). A Type IV GNAT-family acetyltransferase, SsPat (single GNAT domain), was identified by homology to SePat, and was shown to acetylate Alba in vitro at Lys16 (59). Subsequence studies showed acetylation of Lys16 decreased the ability of Alba to bind DNA by ~3-fold (59). However, the authors determined that Alba is a relatively poor substrate for SsPat and that other substrates may exist in S. solfataricus (74). Lys16 of Alba was shown to be deacetylated by the sirtuin deacetylase in S. solfataricus (Sir2) (200), which increased the affinity of Alba for DNA (200). The control of Alba by acetylation and deacetylation seems to mirror histone regulation in eukaryotes, providing an example of a conserved regulatory process.

**RLA Targets Whose Modifying Acetyltransferases are Not Known**

A large number of proteins have been reported to be acetylated, but the identity of the modifying GNAT has not been discovered. Two of these examples include the transcriptional regulator RcsB and the chemotaxis response regulator CheY.

In addition to Acs, an E. coli proteome array incubated with SePat (92% identical to EcPka) and radiolabeled [1-¹⁴C] acetyl-CoA suggested that SePat acetylated several proteins, including the bacterial transcription factor, RcsB (149). RcsB is involved in regulating the expression of genes that affect flagellar and capsule synthesis, as well as cell division (203, 204). Acetylation of RcsB decreases its ability to bind DNA, an effect that is reversed by incubation of RcsB^Ac with the sirtuin EcCobB and NAD$^+$ (149). More recently reported work did not find direct evidence that EcPka acetylates RcsB in E. coli, leaving the identity of the acetyltransferase that modifies RcsB in these bacteria unanswered (205).
The response regulator CheY, involved in bacterial chemotaxis, is predicted to be under the control of RLA (166, 206). While the phosphorylation of CheY has been extensively studied, the effect of acetylation on the activity of the protein is still poorly understood (207-209). Acetylation of CheY inhibits binding to three of its interacting partners, the CheA kinase, the CheZ phosphatase, and the flagellar motor switch FliM protein (166). CheY is acetylated at multiple sites \textit{in vivo} and the majority of the acetylated residues are grouped on the surface of the protein near the C-terminus, the region that binds to the protein targets (210).

While there is \textit{in vivo} and \textit{in vitro} evidence that the CobB sirtuin deacetylates CheY, the identity of the acetyltransferase is unknown (209). Previously, acetylation of CheY was hypothesized to occur via either autoacetylation, or acetyl-CoA synthetase-dependent acetylation by some unknown mechanism (206, 211). To date, studies of CheY have been performed by chemical acetylation of the protein using acetic anhydride (208). However, this method does not acetylate CheY to the same extent as what is seen for the protein \textit{in vivo} (166, 206). It is currently hypothesized that there must be a GNAT acetyltransferase responsible for the acetylation of CheY and is yet to be discovered.

**GNAT STRUCTURE AND SUBSTRATE SPECIFICITY**

\textbf{Structural divergence of ADP-forming acyl-CoA synthetases.} As mentioned earlier, the large domain of Type I and Type II GNATs are homologous to ADP-forming acyl-CoA synthetases. This catalytic mechanism involves a transfer of a phosphate group to a conserved histidinyl residue located within a flexible loop in subdomain 2 (212) (Fig. 2.10B). \textit{EcPka} and \textit{SePat} have divergent sequences in the flexible loop region and lack the catalytic residue, suggesting that their ADP-forming acyl-CoA synthetase domains may lack enzymatic activity.
Other protein acetyltransferases, like RpPat, encode the catalytic histidine. Whether or not RpPat maintains acyl-CoA synthetase activity has not yet been investigated.

*Mycobacterium tuberculosis* PatA is a sensor of carbon quality. The crystal structure of *Mt*PatA has been resolved in the presence and absence of cAMP, and revealed an intricate regulatory mechanism (PDB 4AVB) (194) (Fig. 2.11). Structural studies demonstrated that *Mt*PatA can exist in either an active or auto-inhibited state (194). In the absence of cAMP *Mt*PatA adopts the auto-inhibited state, in which the C-terminal helix (lid) blocks entrance of the protein substrate into the active site of the GNAT domain (PDB 4AVA). In the active state, the cAMP-binding domain is rotated 40° relative to the GNAT domain, causing the inhibitory lid to refold and swing away, exposing the active site cleft (194) (Fig. 2.11). Binding of cAMP stabilizes the active state of *Mt*PatA, enabling acetylation of the target. Acetyl-CoA co-purified with the *Mt*PatA and was present in both crystal structures, demonstrating that it binds tightly to the enzyme and that *Mt*PatA is poised to respond to cAMP levels (194).

**Determinants Needed for Recognition and Acetylation of Protein Targets by GNATs**

The ternary structure of the *Tetrahymena thermophila* GNAT, tGcn5, in complex with an 11-residue peptide from histone H3 and CoA (PDB 1QSN) (Fig. 2.12A) revealed a constellation of interactions between GNATs and their protein substrates (49) (Fig. 2.12B). Notably, CoA binding to tGcn5 triggers structural changes that facilitate its interactions with the protein substrate. From the above-mentioned structure one can see that the role of the catalytic residue (Glu122) is to abstract a proton from the ε-amino group of Lys14 via an ordered water molecule bridging the two residues. Further positioning of Lys14 is afforded by hydrophobic interactions between residues in tGcn5 and methylene groups and the ε-amino of the lysyl side chain.
Figure 2.11. Binding of cAMP induces a 40-Å structural change in *M. tuberculosis* PatA. In the absence of cAMP, *MtPatA* adopts an auto-inhibited state, where a ‘lid’ (blue) blocks the entrance of the substrate to the active site of the GNAT domain (red). In the presence of cAMP, the cAMP-binding domain (gold) rotates 40° relative to the GNAT domain. This causes the lid to swing away from the GNAT domain, exposing the active site cleft. Also shown, acetyl-CoA (black sticks), cAMP (gray spheres), C-terminal helix (green). *MtPatA* structure (PDB 4AVA); *MtPatA* structure with cAMP (PDB 4AVB).
Figure 2.12. Interactions between the *T. thermophila* Gcn5 protein and a peptide substrate. (A) The structure of rGcn5 and a histone H3 11-peptide residue (PDB 1QSN) demonstrated the presence of CoA (not shown) causes structural changes that may facilitate interactions with its protein substrate (interacting residues shown in blue). (B) Molecular interactions of rGcn5 (blue) with the peptide substrate (green) are shown. Modified from figure published in *Nature*; reproduced with permission.
The glycyl side chain next to Lys14 interacts with τGcn5 through van der Waals forces, most likely introducing flexibility into the substrate protein. Binding of the protein substrate to τGcn5 positions the G-K-X-P motif of the former in close proximity to acetyl-CoA in the active site of τGcn5. Such position is maintained via interactions between the prolyl side chain and CoA.

**Structure of a GNAT:protein substrate complex.** Recently, the first structure of a GNAT family member in complex with a protein substrate with tertiary structure was reported (71, 213). The GNAT domain from *S. lividans* SlPatA (SlPatA\textsuperscript{GNAT}) was crystallized in complex with the C-terminal domain of *S. enterica* Acs (SeAcs\textsuperscript{CTD}) (71) (Fig. 2.13A). A comparison of the SlPatA\textsuperscript{GNAT}-SeAcs\textsuperscript{CTD} and τGcn5-peptide structures revealed (i) a glycine residue (G608) preceding the target lysine (K609) was important for positioning the lysine for interaction with SlPatA\textsuperscript{GNAT} and (ii) a hydrophobic pocket in SlPatA\textsuperscript{GNAT} positioned the lysine side chain (K609) near the catalytic glutamate (E123) of the GNAT (Fig. 2.13B). The interaction surface of the SlPatA\textsuperscript{GNAT}-SeAcs\textsuperscript{CTD} acetylation complex was more extensive than the surface observed for τGcn5-peptide interactions, indicating that SlPatA recognizes substrate sequences outside of the flexible loop containing the target lysine (K609). These interactions included complementary ionic interactions of positively charged side chains in SeAcs\textsuperscript{CTD} with negatively charged side-chains in SlPatA\textsuperscript{GNAT} (Fig. 2.13C, D). Reversing the charges in either SeAcs\textsuperscript{CTD} or SlPatA\textsuperscript{GNAT} significantly decreased interactions between these proteins (71).

**Diversity of determinants in the motif containing the acetylation site.** All bona fide substrates of *R. palustris* RpPat are AMP-forming acyl-CoA synthetases and display a high degree of conservation surrounding the acetylation site. A consensus sequence at the site of acetylation can be approximated by the motif PX\textsubscript{4}GK (23, 169) (Fig. 2.14A). The Gly residue preceding the target lysine is conserved in RpPat substrates, as seen with the τGcn5 substrate
Figure 2.13. Molecular interactions of *S. lividans* PatA\(^{GNAT}\) and *S. enterica* Acs\(^{CTD}\). (A) Crystal structure of the interactions between *S. lividans* PatA\(^{GNAT}\) and *S. enterica* Acs\(^{CTD}\) (PDB 4U5Y). The S/PatA\(^{GNAT}\) catalytic residue (E123) is shown in red sticks and the acetylated lysine of SeAcs\(^{CTD}\) (K609) is shown in blue sticks. (B) Interactions between S/PatA\(^{GNAT}\) (surface) and SeAcs\(^{CTD}\) (sticks). (C, D) Electrostatic potential of the S/PatA\(^{GNAT}\)-SeAcs\(^{CTD}\) surface interface with negatively charged regions in red, positively charged regions in blue, and neutral residues in white. This research was originally published in the Journal of Biological Chemistry. Alex C. Tucker, Keenan C. Taylor, Katherine C. Rank, Ivan Rayment, and Jorge C. Escalante-Semerena. Insights into the Specificity of Lysine Acetyltransferases. *J. Biol. Chem.* 2014; 289:36249-36262. © the American Society for Biochemistry and Molecular Biology.
Figure 2.14. Acetylation determinants outside the motif containing the acetylation site. (A) Consensus motif containing the acetylation site (indicated by the arrow) generated from the alignments of acyl-CoA synthetases acetylated by *R. palustris* Pat (RpPat). The letter height corresponds to the frequency of a particular amino acid residue in that position. (B) The electrostatic potential of the *RpMatB* (methylmalonyl CoA synthetase), a protein that is not acetylated by RpPat. The illustration shows negatively charged regions in red and positively charged regions in blue. (C) Electrostatic potential of the *RpMatB* and *B. xenovorans* BclM (benzoate:CoA synthetase) chimera protein (*RpMatB*-BxBclM chimera, B3), a protein that is acetylated by RpPat. The illustration shows negatively charged regions in red and positively charged regions in blue. (D) Overlay of the C-terminal domain of the *RpMatB*-BxBclM chimeras (B1, PDB 4GXQ; B3, PDB 4GXR) aligned with the C-terminal domains of *RpMatB* (PDB 4FUQ), with the BxBclM-derived residues of the B1 chimera in yellow, BxBclM-derived...
residues of the B3 chimera in orange, the wild-type RpMatB residues in cyan. The consensus motif containing the acetylation site (PXaGK) is shown in red in the active site loop, with the acetylated lysine residue (K488) shown as red sticks. This research was originally published in the Journal of Biological Chemistry. Heidi A. Crosby, Katherine C. Rank, Ivan Rayment, and Jorge C. Escalante-Semerena. Structural Insights into the Substrate Specificity of the Rhodopseudomonas palustris Protein Acetyltransferase RpPat: Identification of a Loop Critical for Recognition by RpPat. *J. Biol. Chem.* 2012; 287:41392-41404. © the American Society for Biochemistry and Molecular Biology.
mentioned previously, and may be a common feature of GNAT substrates (49). However, recent results support the conclusion that this motif is not sufficient for acetylation to occur. RpMatB, an acyl-CoA synthetase which activates the dicarboxylic acid methylmalonate to methylmalonyl-CoA (214), contains the PX₄GK motif but is not acetylated by RpPat (215) (Fig. 2.14B).

The term ‘acetylation motif’ is misleading because it oversimplifies what is encoded in this motif. While it is true that the PX₄GK motif identified for AMP-forming acyl-CoA synthetase acetylation targets is necessary for acetylation to occur, it is not sufficient. Alanine scanning of 14 residues surrounding the acetylation site of RpPimA (a bona fide substrate of RpPat) demonstrated that nearly half of them were important for acetylation by RpPat, whereas only two were required for RpPimA enzymatic activity (23, 28, 167). Interestingly, one acyl-CoA synthetase was identified in R. palustris that appeared to evade acetylation through the presence of a leucine residue two positions upstream of the conserved lysine. Changing the leucine to a valine residue restored recognition of the substrate and its acetylation. Collectively, these results emphasize the fact that the presence of the ‘acetylation motif’ in AMP-forming acyl-CoA synthetases is not a good predictor of a protein being under RLA control.

Using Protein Chimeras to Probe GNAT Substrate Specificities

To investigate R. palustris RpPat specificity, a series of chimeric proteins in which portions of RpMatB (methylmalonyl-CoA synthetase, not acetylatable) were replaced with the corresponding sequences from known RpPat substrates (Fig. 2.14C). The RpPat chimeras were constructed with pimelate-CoA synthetase (RpPimA) or benzoate-CoA synthetase from Burkholderia xenovorans, BxBclM (28, 216). Introduction of residues from RpPimA or BxBclM into RpMatB allowed the chimeras, RpPimA-RpMatB and BxBclM-RpMatB to be recognized.
and acetylated by *Rp*Pat. Significantly, *RpPimA-RpMatB* chimeras with *RpPimA* residues located ~20Å away from the target lysine (K488) allowed *RpPat* to recognize and acetylate *RpMatB*. These data indicated that *RpPat* recognizes additional structural elements in protein substrates in addition to the residues immediately surrounding the target lysine, as seen with *SpPatA*. This information may help account for the substrate specificity of protein acetyltransferases for their structurally diverse substrates.

The three-dimensional crystal structure of the *BxBclM-RpMatB* chimera identified a loop, (named chimera loop), that is important for recognition by *RpPat* (215) (Fig. 2.14D). It seems that the shape and electrostatic potential of the chimera loop play important roles, as minor changes in the loop allows an acyl-CoA synthetase to ‘escape’ acetylation by *RpPat* (214). This indicates that although ‘acetylation motifs’ may suggest a protein is controlled by RLA, each substrate should be validated experimentally, as structural elements outside of the motif can affect the ability of the acetyltransferase to recognize and acetylate the target.

**ROLE OF RLA IN MAINTAINING METABOLIC HOMEOSTASIS**

**Acetyl-CoA, Energy Charge, NAD⁺, and cAMP Link RLA to Central Metabolism**

Both components of the RLA system (acylation and deacylation) involve the essential coenzymes CoA and NAD⁺. Acylation requires acyl-CoA thioesters, connecting this process to CoA homeostasis, carbon load, and energy charge, while the sirtuin-catalyzed deacylation reaction requires NAD⁺, an indicator of high energy levels in the cell. Acylation is further regulated by metabolic cofactors in acetyltransferases such as the *Mycobacterium* Pat proteins, which respond to cAMP levels, an indicator of the quality of carbon source available in the cell. As a result, RLA modifies proteins in response to the metabolic state of the cell.
Acetylation and acetyl-CoA levels. Bacterial protein acetyltransferases use acetyl-CoA as a substrate, linking protein acetylation to acetyl-CoA levels, which are regulated by the metabolic activity of the cell. The two-domain Pat homologues SePat, EcPka, RpPat, and SlPatA have large regulatory domains that bind acetyl-CoA to allosterically regulate acetyltransferase activity (72). Acetyl-CoA is a metabolite linked to many pathways including carbon utilization (e.g. glycolysis), the tricarboxylic acid (TCA) cycle, the acetate kinase/phosphotransacetylase pathway, and fatty acid biosynthesis / degradation (Fig. 2.15).

Although some insights into the reasons why AMP-forming acyl-CoA synthetases are regulated by RLA have been reported (see below), there may be a number of other reasons why cells control this class of enzymes so carefully. Not all acyl-CoA synthetases are regulated by acetylation. In prokaryotes, data have been reported about the propionylation of propionyl-CoA synthetase (PrpE), a modification that affects the same lysine as acetylation does, also abolishing the activity of the enzyme (56). The use of RLA to control acyl-CoA synthetases by acetylation or propionylation is not unique to Gamma-proteobacteria, since there is abundant evidence of the same type of control happening in the Alpha-proteobacterium R. palustris. The common theme here is that RLA helps maintain a balance in the intracellular acetyl-CoA (or propionyl-CoA) pools, while an acyl-CoA synthetase (MatB) that contributes to the succinyl-CoA pool is not under RLA control (28, 214). It is possible that acetylation of these AMP-forming acyl-CoA synthetases may control CoA homeostasis by preventing depletion of CoA or build-up of acetyl-CoA or propionyl-CoA.
Figure 2.15. CoA homeostasis. Schematic of the contributions of CoA and acetyl-CoA to cellular metabolism. CoASH, Coenzyme A; Ac-CoA, acetyl-CoA, O-AADPR, O-acetyl-ADP-ribose; Ac-P, acetyl-phosphate; Ac-AMP, acetyl-AMP, PPI, pyrophosphate; Pr-CoA, propionyl-CoA; α-KG, alpha-ketoglutarate; OAA, oxaloacetate; Suc-CoA, succinyl-CoA; dP-CoA, dephospho-Coenzyme A.
**Possible effects on CoA homeostasis.** Coenzyme A (CoA) is an essential metabolic cofactor and CoA homeostasis is important for cell survival. As an acyl carrier group, CoA activates the carbonyl groups of carboxylic acids, including fatty acids and amino acids. The resulting thioester bond increases the electrophilicity of the carbonyl carbon, facilitating nucleophilic attacks, thus making the carbonyl carbon more prone to react with thiolates, hydroxyl, and amino groups (217, 218). Reactive acyl-CoA thioesters are used by ~ 4% of all known enzymes, which catalyze over 100 reactions involved in diverse cellular processes, including the TCA cycle, fatty-acid degradation, and fatty-acid, amino-acid, and secondary-metabolite biosynthesis (219, 220) (Fig. 2.15). Due to both limiting substrate availability and allosteric regulation of central metabolic enzymes, CoASH and acyl-CoA control metabolic flux through glycolysis and the TCA cycle (221-227).

CoASH, acetyl-CoA, succinyl-CoA, and malonyl-CoA comprise the bulk of the CoA pool (228). Availability of nutrients, phase of growth, and environmental conditions all affect the CoA pool and can cause the balance of acyl-CoA species to alter by more than an order of magnitude in a matter of minutes (220, 229-231). For example, acetyl-CoA is the major species during exponential growth on glucose (300 µM), while CoASH is the predominant species during growth on acetate (100 µM) (228).

Because CoASH and acyl-CoA availability affect many cellular processes, an imbalance in CoA homeostasis results in profound consequences on cellular metabolism. For example, depletion of CoA stalls protein synthesis and reduces the supply of acyl carrier protein (ACP). Protein synthesis is stalled by depletion of CoA due to the lack of available acetyl-CoA as well as inhibition of the TCA cycle and production of amino-acid precursors (232). Reduced levels of ACP limits fatty-acid biosynthesis, ultimately resulting in reduced phospholipid synthesis (232-
Due to the deleterious effects on the cell caused by an imbalance in CoA homeostasis, there must be tight control over both the total CoA pool, as well as the relative concentrations of the various CoA species, in response to the metabolic status of the cell. Given that the RLA system recycles acylated CoAs, this could be a mechanism for the maintenance of CoA homeostasis.

**Effect of RLA on energy charge.** Insights into why acyl-CoA synthetases are under RLA control have been reported in *S. enterica* (179). From this work, the authors learned that, in this bacterium, an imbalance in the protein acetyltransferase (*SePat*):sirtuin deacetylase (*SeCobB*) ratio has a profound effect on cell growth under conditions that depend on the activity of acetyl-CoA synthetase (*SeAcs*), i.e. \( \leq 10 \text{ mM acetate as the sole source of carbon and energy.} \) Results from experiments where *SePat* was ectopically synthesized under the control of an inducible promoter showed, that incremental levels of *SePat* eventually arrested growth because the energy charge of the cell was lowered to a level (0.17) that could not support growth. The depletion of ATP and the concomitant production of AMP was determined to be the reason for the drop in energy charge. That is the cell did not have enough ATP to convert AMP to ADP so the ATPase could synthesize more ATP, thus restoring the energy charge of the cell.

**Deacetylation and NAD\(^+\) levels.** Sirtuins require NAD\(^+\) as a co-substrate, linking deacetylation to the availability of NAD\(^+\) in the cell. NAD\(^+\) is the oxidized form of NADH, an important electron donor to the electron transport system, which generates the proton motive force that drives the synthesis of ATP by the membrane-bound ATPase. When bacteria are grown in conditions with differing nutritional and oxygen availability, the NAD\(^+\) pools have greater variation compared to the NADH pools (130). This suggests that the NAD\(^+\) pool is (i) dynamic and (ii) an important reporter of cellular carbon and energy status. For example, growth
conditions that generate high NAD$^+$ levels, like aerobic respiration, could cause an increase in the deacylation of sirtuin targets. A recent report provided evidence of a new and unprecedented role for NAD$^+$ in cell physiology. Cahová et al presented experimental evidence supporting the idea that bacteria stabilize RNA molecules by capping their 5’ end with NAD$^+$ (235). Whether NAD$^+$-capped RNAs are substrates for sirtuins is an intriguing possibility that should be explored.

**Regulation of protein acetylation by cAMP.** In *E. coli*, cAMP is involved in catabolite repression, a process in which cells preferentially use glucose and only utilize other available carbon sources once glucose has been depleted. The global transcriptional regulator, Crp mediates this cAMP-dependent response (236). In mycobacteria, cAMP not only plays a role in basic physiology, it acts as a secondary messenger and is involved in re-routing host signaling during infection (237). Notably, *M. tuberculosis* has 15 adenylate cyclase enzymes for production of cAMP, whereas *E. coli* only has one (CyaA) [reviewed in (238)]. The adenylate cyclases of *M. tuberculosis* are allosterically activated by signals like low pH (239, 240), bicarbonate / CO$_2$ (241), and saturated fatty acids (242), which occur during the course of infection.

While cAMP concentrations do not fluctuate with addition of glucose to the medium, cAMP availability does increase ~50-fold during macrophage infection (237). For *M. tuberculosis* to enter a non-replicating persister or ‘quiescent’ state during chronic infection requires a significant downshift in metabolism, which is thought to be achieved by diverting acetyl-CoA away from the TCA cycle, toward synthesis of triacylglycerides (243). Mycobacteria accumulate and store triacylglycerides during stress conditions, although their role is not fully understood.
In *M. tuberculosis*, cAMP is integral to the control of the RLA system through the allosteric regulation of the acetyltransferase, *MtPatA*. It is possible that Mycobacteria may use cAMP availability to adjust flux through AMP-forming acyl-CoA synthetases, the targets of *MsPatA* (192). Taken together, it is also conceivable that *M. tuberculosis* would down-regulate Acs activity as a mechanism to reduce the available acetyl-CoA pool in order to slow down cellular metabolism during chronic infection.

**Cellular Stress and RLA**

Mammalian sirtuins have been reported to play a role in cellular protection by promoting positive effects on processes like DNA repair, cell survival, and stress resistance [reviewed in (244)]. Recent studies have shown that this effect is also conserved in bacteria.

Deletion of the *E. coli* sirtuin deacetylase (*EcCobB*) increases acetylation levels in the cell, which was shown to increase resistance to both heat and oxidative stress (245). This same study performed whole-transcriptome analysis of a $\Delta$cobB strain and found many stress related systems were repressed in this mutant, including genes related to heat shock, osmotic stress, acid resistance, cold shock, and carbon starvation (245). These data provided compelling evidence that RLA is a mechanism used by cells to respond to environmental stressors.

An acetylome study of *M. tuberculosis* identified 10% of the enzymes involved in fatty acid biosynthesis were acetylated (143). Fatty acid biosynthesis in this organism plays a role in colony morphology and biofilm formation. The authors demonstrated that deletion of the sirtuin deacetylase in this organism (MRA_1161) resulted in a more granular morphology, a decrease in biofilm formation, and increased resistance to heat stress (143). These studies further support the hypothesis that RLA is a mechanism for protection against environmental stresses.
TRANSCRIPTIONAL REGULATION OF GENES ENCODING RLA ENZYMES

While the field is beginning to identify acetyltransferases, partner deacetylases, and protein targets, as well as distinguish structural determinants needed for recognition and enzyme regulation, much remains unknown about the transcriptional regulation of the genes encoding the components of the RLA systems.

Regulation of RLA Genes in *Escherichia coli*

The *pka* gene, which encodes the protein acetyltransferase of *E. coli*, has a similar expression profile to that of *acs*, one of its target substrates (177). When *E. coli* grows in certain conditions (i.e. glucose minimal medium), the TCA cycle cannot process acetyl-CoA quickly enough and instead utilizes an overflow pathway to excrete acetate. This acetate pool is later assimilated via Acs and utilized by the cell when glucose is depleted (246).

This “acetate switch” is mediated in part by the expression of *acs*, which is low during exponential phase but increases dramatically in late exponential and early stationary phases (247). Expression of *acs* is controlled by several transcription factors including the nucleoid proteins Fis and IHF, as well as Crp (177, 248). Interestingly, the *E. coli* protein acetyltransferase *pka* has a similar expression profile to its target, *acs*, during growth on glucose (177). Transcription of *pka* is also activated by Crp in response to cAMP levels (177). The activation of *pka* and *acs* expression by Crp may represent a regulatory mimic of the direct cAMP-dependent activation of GNAT activity observed in mycobacteria. At present, it appears that the *cobB* gene encoding the sirtuin in *E. coli* is constitutively expressed and is not dependent on cAMP levels (177).
Multiple Isoforms of the CobB Deacetylase are Present in *Salmonella enterica*

In *S. enterica*, the CobB sirtuin exists as two isoforms, due to the presence of two independent start codons that are read in the same open reading frame. This results in the production of a long isoform (CobB\textsubscript{L}, 273 a.a.), and a short isoform (CobB\textsubscript{S}, 236 a.a.), which lacks 37 of the amino acids present in CobB\textsubscript{L} (249). Interestingly, both CobB\textsubscript{L} and CobB\textsubscript{S} isoforms are active, although CobB\textsubscript{S} is the dominant isoform *in vivo* and is produced at ~10-fold higher levels (249). The presence of the dual start codons is not limited to *S. enterica*, and is found throughout the enterobacteria. The physiological relevance of CobB\textsubscript{L} and CobB\textsubscript{S} and their contribution to RLA is currently unknown.

**CONCLUSIONS**

RLA is an emerging field in prokaryotes that is advancing by leaps and bounds through the use of high-throughput and detailed mechanistic studies in a variety of organisms. Validation of ideas obtained through global ‘omics’ approaches is key to improving our understanding of the role of RLA. GNAT protein acetyltransferases and their cognate protein deacetylases have been identified in bacteria, archaea, and eukaryotes. The abundance of GNATs in cells of all domains of life is a strong indicator of the relevance of these enzymes to life. Although the physiological role of the majority of these enzymes remains unknown, and the elucidation of their function is a challenge to cell physiologists, efforts to advance this research area will likely provide valuable insights into the strategies used by cells to cope with metabolic stress.
FOOTNOTES
RLA, reversible lysine acetylation; PTM, post-translational modification; HAT, histone acetyltransferase; Co-ASH, Coenzyme A; Ac-CoA, acetyl-CoA; GNAT, Gen5-related N-acetyltransferase; Pat, protein acetyltransferase; Acs, acetyl-CoA synthetase; CobB, NAD\(^+\)-dependent sirtuin deacetylase; O-AADPR, O-acyl-ADP-ribose; HDAC, histone deacetylase; NUDIX, nucleotide diphosphate linked to X; SILAC, stable isotope labeling with amino acids in cell culture; Ec, Escherichia coli; Se, Salmonella enterica; Sl, Streptomyces lividans; Bs, Bacillus subtilis; Ms, Mycobacterium smegmatis; Mt, Mycobacterium tuberculosis; Ss, Sulfolobus solfataricus; Rp, Rhodopseudomonas palustris, Bx, Burkholderia xenovorans.

ACKNOWLEDGEMENTS
This work was supported by USPHS grant R01 GM062203 to J.C.E.-S. We would like to thank Dr. Heidi Crosby, Dr. Alex Tucker, and Chelsey VanDrisse for critical reading and feedback during the drafting stages of this review.

REFERENCES


29. **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**:545-555.


93


104


CHAPTER 3

DECIPHERING THE REGULATORY CIRCUITRY THAT CONTROLS REVERSIBLE LYSINE ACETYLATION IN *SALMONELLA ENTERICA*²

ABSTRACT

In *Salmonella enterica*, the reversible lysine acetylation (RLA) system is comprised of the protein acetyltransferase (Pat) and sirtuin deacetylase (CobB). RLA controls the activities of many proteins, including the acetyl-CoA synthetase (Acs), by modulating the degree of Acs acetylation. We report that IolR, a *myo*-inositol catabolism repressor, activates the expression of genes encoding components of the RLA system. *In vitro* evidence shows that the IolR protein directly regulates *pat* expression. An *iolR* mutant strain displayed a growth defect in minimal medium containing 10 mM acetate, a condition in which RLA function is critical to control Acs activity. Increased levels of Pat, CobB, or Acs activity reversed the growth defect, suggesting the Pat:CobB ratio in an *iolR* strain is altered and that such a change affects the level of acetylated, inactive Acs. Results of *in vitro* assays showed an ~25% decrease in Acs activity in cell-free extracts of the mutant *iolR* strain relative a strain carrying the wild-type *iolR* allele. An *iolR* mutant strain displayed decreased expression of *pat, cobB*, and *acs*, and glucose differentially regulated expression of *pat, cobB*, and *acs*. The catabolite repressor protein (Crp) positively regulated expression of *pat* while having no effect on *cobB*.

IMPORTANCE

Reversible lysine acylation (RLA) is used by cells of all domains of life to modulate the function of proteins involved in diverse processes. Work herein begins to outline the regulatory circuitry that integrates the expression of genes encoding enzymes that control the activity of a central metabolic enzyme in C2 metabolism. Genetic analyses revealed subtle effects on RLA that greatly impacted the growth behavior of the cell. This work provides the first insights into the complexities of the system responsible for controlling RLA at the transcriptional level.
INTRODUCTION

Reversible lysine acetylation (RLA) is a posttranslational regulatory mechanism present in all domains of life (1). RLA allows an organism to rapidly and reversibly modulate the biological activity of proteins involved in carbon utilization, transcription, translation, and stress responses (2-5) by modulating the acetylation state of epsilon amino group of lysyl residues critical for function [reviewed in (6)]. In the last decade, studies have provided insights into how the RLA system works in diverse prokaryotes (3, 7-10). In *S. enterica* the RLA system is comprised of a protein acetyltransferase (Pat) of the Gcn5 \( N \)-acetyltransferase (GNAT) family, and a NAD\(^+\)-consuming sirtuin deacetylase (CobB) (2) (Fig. 3.1). Relevant to this work is the RLA control of acetyl-CoA synthetase (Acs), an AMP-forming CoA ligase involved in acetate utilization (11). Pat is responsible for the acetylation and inactivation of Acs (2), while removal of the acetyl moiety of Acs\(^{Ac}\) by the CobB deacetylase reactivates Acs (12). RLA-dependent regulation of Acs is imperative, as uncontrolled Acs results in growth arrest by depletion of ATP pools (8).

In addition to post-translational regulation, expression of *acs* is controlled by several transcriptional regulators (13). While the regulatory region of *cobB* in *S. enterica* has been examined to some extent (14), the transcriptional regulation of genes encoding the enzymes of the RLA system (*pat, cobB*) has not been investigated. It has been shown that the catabolite repressor protein (Crp) regulates the *E. coli* pat homologue (*pka*) (15), although a role for Crp regulation of *pat* in *S. enterica* has not been reported.

In this work a genetic approach was used to identify *S. enterica* genes whose products affected the *pat* promoter (\( P_{pat} \)). We show that inactivation of *iolR* (*stm4417*), encoding an RpiR-like transcriptional repressor, decreased *pat* expression (16). RpiR-like regulators are involved in sugar catabolism and can function as activators and repressors (17, 18). In *S. enterica, Bacillus*
Figure 3.1. Reversible Lysine Acetylation (RLA) in *S. enterica*. Activity of the AMP-forming acetyl-CoA synthetase (Acs) is post-translationally modified by the protein acetyltransferase Pat. This modification is reversible by the activity of the NAD⁺-consuming class III sirtuin deacetylase, CobB. O-AADPR, O-acetyl ADP ribose; Nm, nicotinamide.
subtilis, Corynebacterium glutamicum, and Sinorhizobium meliloti, IolR negatively regulates expression of the myo-inositol utilization operon (16, 19). Myo-Inositol (cyclohexane-1,2,3,4,5,6-hexol) is an abundant cyclic polyol in soil and its utilization as a sole carbon source depends on the presence of a large number of genes organized as a genomic island (16), which is present in γ-proteobacteria, α-proteobacteria, and Gram-positive bacteria (20-26).

Here we present in vivo evidence that IolR activates pat expression in S. enterica, and that IolR binds to the pat promoter in vitro. We also report that acs and cobB are transcriptionally activated by IolR, which places the reversible lysine acetylation system of S. enterica under IolR control. Significantly, an iolR mutant strain displayed a growth defect in minimal medium containing 10 mM acetate, which we suggest is due to an imbalance of active (non-acetylated) / inactive (acetylated) Acs ratio caused by changes in pat and cobB expression in the absence of IolR. Lastly, we show that Crp, a global regulator of carbon metabolism, regulates pat and acs expression in S. enterica. To our knowledge, this is the first report of global, integrative transcriptional control of genes encoding the enzymes of the RLA system in S. enterica and its effect on carbon metabolism.

RESULTS

IolR regulates pat expression. A genetic screen was used to identify genes whose functions affected the expression of pat, the gene encoding the protein acetyltransferase in S. enterica. Changes in pat expression were monitored in strain JE7449, which carried a chromosomal pat::MudJ (lacZ⁺ kan⁺) reporter (hereafter pat-lacZ⁺; Table 3.1). This strain was transduced to tetracycline resistance (TcR) via a P22 lysate grown on a pool of ~100,000 strains, each of which assumed to contain one Tn10d(tet⁺) element randomly inserted in the genome. TcR derivatives of
the pat-lacZ\textsuperscript{+} reporter strain (~20,000) were screened for changes in β-galactosidase activity, leading to the identification of two colonies that were less blue than the parental strain. The DNA sequence flanking the Tn10\textsuperscript{d}(tet\textsuperscript{+}) elements located both insertions within iolR (stm4417), the gene encoding the repressor of the myo-inositol utilization (iol) genes. No other insertions affecting pat expression were identified in the screen. To confirm that the iolR::Tn10\textsuperscript{d}(tet\textsuperscript{+}) element was responsible for the reduced expression of the pat-lacZ\textsuperscript{+} reporter, phage P22 grown on the original iolR::Tn10\textsuperscript{d}(tet\textsuperscript{+}) pat-lacZ\textsuperscript{+} strain was used to transduce strain JE7449 (pat-lacZ\textsuperscript{+}) to Tc\textsuperscript{R}. The reconstructed iolR::Tn10\textsuperscript{d}(tet\textsuperscript{+}) pat-lacZ\textsuperscript{+} strain (JE10535) displayed the same reduction in pat-lacZ\textsuperscript{+} expression as measured in the original mutant strain (data not shown).

To independently confirm the effect of IolR on pat expression, an iolR::cat\textsuperscript{+} mutation was introduced into strain JE7449 (pat-lacZ\textsuperscript{+}). Measurements of β-galactosidase activity of the pat-lacZ\textsuperscript{+} iolR::cat\textsuperscript{+} strain (JE10714) during growth in nutrient broth (NB) showed a reproducible ~2-fold decrease in pat promoter (P_{pat}) activity relative to that in the pat-lacZ\textsuperscript{+} iolR\textsuperscript{+} strain (Fig. 3.2A). Complementation analysis with an iolR copy provided in trans restored pat expression to wild type level (Fig. 3.2B). The effect of IolR on pat expression was confirmed using qRT-PCR, and showed a 5-fold down-regulation of the pat transcript in an iolR::cat\textsuperscript{+} strain compared to wild type (Fig. 3.2C). From these data we concluded that the decrease in pat expression was due to the absence of IolR. The effect of IolR on pat expression was tested on acetate (10 mM) and myo-inositol (55 mM). In the absence of IolR, pat-lacZ\textsuperscript{+} expression decreased 1.4-fold on acetate (Fig. 3.3A) and 1.3-fold on myo-inositol (Fig. 3.3B), compared to the iolR\textsuperscript{+} strain.
### Table 3.1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference/source</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR6583</td>
<td>metE205 ara-9</td>
<td>K. Sanderson via J. Roth</td>
</tr>
<tr>
<td><strong>Derivatives of TR6583</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JE2845</td>
<td>cobB1206::MudJ</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>JE4637</td>
<td>pACS3</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>JE7449</td>
<td>pat10::MudJ</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>JE7758</td>
<td>Δacs2</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>JE9912</td>
<td>Δacs2 / pACS7</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>JE10040</td>
<td>pBAD30&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>JE10535</td>
<td>pat10::MudJ iolR52::Tn10d(tet)&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>JE10713</td>
<td>iolR51::cat</td>
<td></td>
</tr>
<tr>
<td>JE10714</td>
<td>pat10::MudJ iolR51::cat</td>
<td></td>
</tr>
<tr>
<td>JE10727</td>
<td>pat10::MudJ iolR51::cat / pBAD30</td>
<td></td>
</tr>
<tr>
<td>JE10728</td>
<td>pat10::MudJ iolR51::cat / pIOLR1</td>
<td></td>
</tr>
<tr>
<td>JE14413</td>
<td>DiolR53</td>
<td></td>
</tr>
<tr>
<td>JE14962</td>
<td>iolR::cat / pACS3</td>
<td></td>
</tr>
<tr>
<td>JE14972</td>
<td>cobB1206::MudJ iolR51::cat</td>
<td></td>
</tr>
<tr>
<td>JE16466</td>
<td>crp892::cat</td>
<td></td>
</tr>
<tr>
<td>JE16596</td>
<td>Δacs2 iolR51::cat / pACS7</td>
<td></td>
</tr>
<tr>
<td>JE16743</td>
<td>pat10::MudJ crp892::cat</td>
<td></td>
</tr>
<tr>
<td>JE16744</td>
<td>pat10::MudJ DiolR51 crp892::cat</td>
<td></td>
</tr>
<tr>
<td>JE16771</td>
<td>pat10::MudJ crp892::cat / pBAD30</td>
<td></td>
</tr>
<tr>
<td>JE16843</td>
<td>cobB1206::MudJ iolR51::cat / pBAD30</td>
<td></td>
</tr>
<tr>
<td>JE16844</td>
<td>cobB1206::MudJ iolR51::cat / pIOLR1</td>
<td></td>
</tr>
<tr>
<td>JE16934</td>
<td>iolR51::cat / pBAD30</td>
<td></td>
</tr>
<tr>
<td>JE16935</td>
<td>iolR51::cat / pIOLR1</td>
<td></td>
</tr>
<tr>
<td>JE17322</td>
<td>pat10::MudJ crp892::cat / pCRP2</td>
<td></td>
</tr>
<tr>
<td>JE18891</td>
<td>iolR51::cat / pCOBB8</td>
<td></td>
</tr>
<tr>
<td>JE18927</td>
<td>iolR51::cat / pPAT27</td>
<td></td>
</tr>
<tr>
<td>JE19811</td>
<td>pat10::MudJ iolR51::cat / pPAT27</td>
<td></td>
</tr>
</tbody>
</table>

**E. coli strains**

*E. coli* C41(IDE3) with *ompT hsdS (r<sup>B</sup> m<sup>B</sup>) gal (λDE3)* including at least one non-characterized mutation

{Miroux, 1996 #20626; Rocco, 2008 #11866}

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Genotype</th>
<th>Reference/source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pIOLR1</td>
<td><em>iolR</em> cloned into pBAD30&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pIOLR3</td>
<td><em>iolR</em> cloned into pTEV6&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pCRP2</td>
<td><em>crp</em> cloned into pBAD30&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pCRP3</td>
<td><em>crp</em> cloned into pTEV16&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pACS3</td>
<td>P&lt;sub&gt;acs&lt;/sub&gt; cloned into pFZY1&lt;sup&gt;g&lt;/sup&gt;</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>pACS7</td>
<td><em>acs</em> cloned into pBAD30&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>pACS33</td>
<td><em>acs</em> cloned into pTEV5&lt;sup&gt;q&lt;/sup&gt;</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>pCOBB8</td>
<td><em>cobB</em> cloned into pBAD30&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>pPAT27</td>
<td><em>pat</em> cloned into pBAD30&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Laboratory collection</td>
</tr>
<tr>
<td>pPAT8</td>
<td><em>pat</em> cloned into pTEV6&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Laboratory collection</td>
</tr>
</tbody>
</table>

<sup>a</sup> All strains and plasmids were constructed during the course of this work, unless otherwise stated.
<sup>b</sup> MudJ is an abbreviation of MudI1734 (*lacZ<sup>+</sup> kan<sup>+</sup>) {Castilho, 1984 #11544}
pBAD30 refers to a cloning vector described in {Guzman, 1995 #11361}

Tn10d(tet) is an abbreviation of Tn10D16D17 {Way, 1984 #21272}

pTEV6 is an overexpression vector described in {Rocco, 2008 #11866}

pTEV16 is a modified version of pTEV5 with BspQ1 restriction sites (C. M. VanDrisse and J. C. Escalante-Semerena, unpublished data)

pFZY1 is a vector containing a promoterless lacZ fusion {Koop, 1987 #3238}

pTEV5 is an overexpression vector described in {Rocco, 2008 #11866}
Figure 3.2. IolR activates pat expression in vivo. Activity of a pat-lacZ+ chromosomal operon fusion was assessed in the presence (JE7449) or absence (JE10714) of iolR to measure pat promoter activity in vivo (A, B). Cell cultures were grown at 37°C in NB medium. The data presented are the average of two independent experiments from individual cultures performed in triplicate. Error bars represent standard deviation. Unpaired t test gave a P value of 0.0004 (A). C. qRT-PCR showed a 5-fold down-regulation in pat activity in an iolR strain relative to the iolR+ strain. Wild-type transcript level is set at 1, indicated by the dashed line. Error bars represent standard deviation.
Figure 3.3. Expression of pat on acetate and myo-inositol. Cultures were grown in NCE minimal medium with acetate (10 mM; panel A), or myo-inositol (55 mM; panel B). β-Galactosidase activity was measured at mid-log phase. Assays were determined in duplicate from three biological replicates.
**Pat does not acetylate IolR, and is not required during growth on myo-inositol.** IolR represses the expression of genes encoding myo-inositol degrading enzymes (16, 33), and it was surprising to find that IolR may also play a role in the activation of genes comprising the RLA system. We considered the possibility of a regulatory system in which Pat would control the DNA binding activity of IolR via acetylation, as reported for Pat and the *E. coli* transcription factor RcsB (3). However, we did not obtain any experimental evidence of Pat-dependent regulation of IolR function under conditions in which acetyl-CoA synthetase (Acs), a bona fide Pat substrate (12), was acetylated (Fig. 3.4).

Consistent with the above-mentioned observation, we determined no difference in growth rate of the *pat-lacZ*<sup>+</sup> *iolR::cat*<sup>+</sup> to a strain carrying the wild-type *pat* and *iolR* alleles when grown on myo-inositol (Fig. 3.5). However, the *pat-lacZ*<sup>+</sup> *iolR*<sup>+</sup> strain consistently showed a slight but reproducible delay in the onset of growth. As reported by others (16), we observed that the onset of growth of the *iolR::cat*<sup>+</sup> strain occurred substantially earlier than that of a strain carrying the wild-type *iolR* allele, an observation consistent with the lack of repression of the myo-inositol genomic island in a strain devoid of the IolR repressor (Fig. 3.5).

**IolR is a tetramer.** To study the role of IolR regulation of *pat*, the IolR protein was isolated to 96% homogeneity (Fig. 3.6). The oligomeric state of IolR was determined using FPLC gel filtration analysis (Fig. 3.6). Under the conditions tested, IolR eluted ~24 min after injection, a retention time consistent with the behavior of a protein whose mass was approximately ~134 kDa when compared to elution times of molecular mass standards. Since the predicted molecular mass of IolR was 31 kDa, it was concluded that IolR was either a dimer of dimers or a tetramer.
Figure 3.4. Pat does not acetylate IolR or Crp. An in vitro acetylation assay using [1-\textsuperscript{14}C] Ac-CoA was used to test Pat-dependent acetylation of IolR or Crp. Acs was used as a positive control. Reactions were analyzed SDS-PAGE and stained with Coomassie Blue to visualize proteins (A). Transfer of the acetyl group was visualized by phosphor-imaging (B). +Ctrl, positive control.
Figure 3.5. Growth study controls. Growth of *S. enterica* strains in NCE minimal medium with (A) *myo*-inositol (50 mM), (B) acetate (50 mM), and (C) glycerol (22 mM). Growth curves were performed using a Powerwave XS2 microplate reader (Bio-Tek Instruments) at 37°C with shaking in triplicate in three independent experiments. Strains analyzed: *iolR*<sup>+</sup> *pat*<sup>+</sup> (JE6583), *iolR::cat*<sup>+</sup> (JE10713), *iolR::cat*<sup>+</sup> / pVOC (JE16934), *iolR::cat*<sup>+</sup> / *pioIR*<sup>+</sup> (JE16935), and *pat::MudJ* (JE7449). Plasmids were induced with 100 µM *L*-(-)-arabinose. Error bars represent standard deviation. pVOC, vector-only control. Symbols shown in panel C apply to all panels.
Figure 3.6. IolR is a tetramer. (A) *S. enterica* IolR protein (30.8 kDa) was purified using nickel affinity purification. An SDS-PAGE gel shows molecular size standards (kDa) (lane 1) and purified IolR protein (lane 2), purified to 96% homogeneity. (B) The molecular mass of IolR (closed circle) in solution was estimated by gel filtration. Molecular mass standards (open circles) are thyroglobulin (bovine; 670 kDa), γ-globulin (bovine; 158 kDa), ovalbumin (chicken; 44 kDa), myoglobin (horse; 17 kDa) and vitamin B\textsubscript{12} (1.35 kDa).
**IolR binds to the pat promoter (P_{pat}) in vitro.** Electrophoretic mobility shift assays (EMSAs) were performed to determine if the effect of the IolR regulator on pat expression was the result of direct binding of IolR to P_{pat}. Experimental promoter analysis data was used to identify the transcription start site (TSS, www.imib-wuerzburg.de/research/salmonella) (34). To probe for the specificity of the interaction between IolR and P_{pat}, we added a non-specific competitor DNA P_{argS}, previously used to study IolR binding (16). Increasing concentrations of IolR shifted the P_{pat} probe, but not the P_{argS} probe, a result that supported the conclusion that IolR directly and specifically interacted with the P_{pat}, (Fig. 3.7A). Increasing amounts of IolR protein titrated against a fixed amount of the P_{pat} DNA probe without the presence of P_{argS} also yielded increased amounts of IolR/P_{pat} complex (Fig. 3.7B). IolR did not shift the mobility of P_{argS} until a molar excess of 50x protein was reached, indicative of non-specific binding of IolR to P_{argS} (Fig. 3.7C).

Previous studies by others showed that IolR negatively regulates the transcription of its own promoter, P_{iolR} (16). Using P_{iolR} (175 nt) as a positive control with the presence of the competitor probe P_{argS}, we confirmed the reported specificity of IolR for its promoter (16) (Fig. 3.7D). The P_{pat} and P_{iolR} probes each shifted at similar molar excess concentrations of IolR, supporting the conclusion that IolR directly and specifically interacted with P_{pat}.

**Region of the pat promoter recognized by IolR.** We performed DNA-footprinting analysis to identify the region within the pat promoter recognized by IolR. A 6FAM-5’-labeled 382-nucleotide probe containing the P_{pat} was incubated with varying concentrations of IolR protein or bovine serum albumin (negative control). After incubation and subsequent DNase digestion and purification of the DNA, samples were analyzed as described under Materials and Methods. Electropherogram overlays comparing IolR and BSA (negative control) and putative binding
Figure 3.7. IolR binds to the pat promoter region. Binding of IolR to the 6FAM-5' labeled pat promoter (P\textsubscript{pat}, 150 nt, 51 nM) was analyzed by electrophoretic mobility shift assays in the presence of increasing concentrations of IolR. (A) P\textsubscript{pat} and competitor DNA, (P\textsubscript{argS}, 196 nt), were incubated together to show binding specificity of IolR to P\textsubscript{pat}. (B) The P\textsubscript{pat} probe alone was incubated at varying concentrations of IolR. (C) Competitor DNA P\textsubscript{argS} was incubated with increasing concentrations of IolR to determine at what point non-specific binding interactions occur. (D) The interaction between IolR and P\textsubscript{iolR} (175 nt) was performed as a known binding control and incubated in the presence of competitor DNA, P\textsubscript{argS}. Protein concentration shown is in molar excess to probe (pmol). EMSAs were performed in triplicate.
sites were analyzed by aligning the sequenced probe data (data not shown). Data presented in figure 3.8 show a region of protection of $P_{pat}$ from nucleotides -112 to -70 relative to the predicted transcription start site (Fig. 3.8A). Experimental promoter analysis data from Kroger et al was used to identify the transcription start site (www.imib-wuerzburg.de/research/salmonella) (34) (Fig. 3.8B). A region of hypersensitivity was seen at position -112, an indicator of DNA bending as the result of the binding of a transcriptional regulator, causing an exposed site susceptible to increased cleavage by DNase. A control was performed in which the amount of IolR was doubled in the reaction (10 µg). With this increase in protein concentration we expected an increase in signal intensity, as seen in figure 3.8A.

Previous studies aimed at examining the regulation of the $iol$ genomic island in $S. enterica$ by IolR repression did not identify a conserved binding site (16, 35). The intent of the DNA-footprinting analysis was to compare the binding region of IolR within $P_{pat}$ to promoter regions regulated by IolR, with the idea of determining a consensus IOLR-binding region. While unable to determine a consensus site, the data confirmed the direct interaction between IolR and $P_{pat}$.

**IolR binds to the region of $P_{pat}$ identified by DNA footprinting.** A 45-nt probe corresponding to the protected region of $P_{pat}$ identified as the IolR-binding region was used to validate the DNA-footprinting experiments. A 6FAM-5’-P$_{pat}$45 45-nt probe was generated by annealing complementary primers and the binding of IolR to this region was examined. Data presented show that IolR binds to the 45-nt probe, confirming that the IolR-binding site is located within this region (Fig. 3.9). The reason for the presence of signals of higher molecular mass complexes is unclear. Possible explanations include the absence of a ligand sensed by IolR, or the formation of higher order IolR multimers.
Figure 3.8. IolR protein binds *pat* promoter at position -112 to -70. (A) DNA-footprinting analysis by capillary electrophoresis was used to define the IolR binding region on the *pat* promoter (Pₚₑₐ). On the graph, negative values represent an area of protection, bars indicate concentration of IolR protein (5 µg dark gray, 10 µg light gray), and bar heights represent the area difference between the IolR sample and the negative control (BSA). DNA-footprinting was performed and analyzed in two independent experiments. (B) Work by Kroger *et al.* (34) was used to identify the *pat* transcriptional start site (TSS).
Figure 3.9. **IolR protein binds to the P<sub>pat</sub> 45-nt probe.** Electrophoretic mobility shift assays were used to validate the binding region identified by footprinting experiments. A 6FAM-5’-labeled 45-nt probe of the <i>pat</i> promoter region, P<sub>pat</sub> (166 nM), corresponding to the identified binding region was incubated with increasing concentrations of IolR protein. Protein concentration shown is in molar excess to probe. EMSAs were performed in triplicate.
The absence of IolR impairs growth on 10 mM acetate. Growth of an iolR::cat\(^+\) strain was inhibited on 10 mM acetate, with a growth rate three times slower (doubling time = 36 h; Fig. 3.10A, solid triangles) than that of strain carrying the wild-type iolR allele (doubling time = 11 h; Fig. 3.10A, solid squares). Growth of the iolR::cat\(^+\) strain was restored when iolR was expressed ectopically (Fig. 3.10A, open triangles), indicating the growth defect was due to the absence of IolR. A similar growth defect was reported for an iolR mutant strain of C. glutamicum, but this observation was not investigated (33). No growth differences were observed for the iolR\(^+\) pat\(^+\), iolR::cat\(^+\), or pat::MudJ strains when grown on 50 mM acetate or glycerol (Fig. 3.5C, D).

Because the observed phenotype on 10 mM acetate correlated with lower levels of pat expression, we hypothesized that increases in the expression of pat under the control of an IolR-independent promoter would restore growth of the iolR::cat\(^+\) strain on 10 mM acetate. Indeed, the iolR::cat\(^+\) strain grew almost as well as the iolR\(^+\) strain when pat was expressed in trans (Fig. 3.10B, inverted triangles). An increase in the level of inducer (10-100 \(\mu\)M) compromised growth of the iolR::cat\(^+\) / ppat\(^+\) strain. The negative effect of higher pat expression was not surprising since increased Pat levels are known to increase the level of acetylated, inactive Acs (8).

The phenotype of the iolR strain is caused by an imbalance in the Pat:CobB ratio, which affects Acs activity. In S. enterica, Pat and CobB control Acs activity (6). Given that pat expression decreased in the iolR::cat\(^+\) strain, we hypothesized that the absence of IolR created an imbalance in the Pat:CobB ratio that favored Pat activity, thus a decrease in Acs activity due to Acs acetylation. We reasoned that such loss of Acs activity could be counteracted in several ways. Firstly, inactivation of pat in the iolR strain would block Acs acetylation, and should restore growth. Indeed, the poor growth of the iolR strain on 10 mM acetate (Fig. 3.10C, solid
Figure 3.10. An iolR strain has a growth defect on 10 mM acetate. Growth of *S. enterica* strains was examined in NCE minimal medium containing 10 mM acetate. Expression of iolR was induced using 100 µM L-(+)-arabinose, cobB expression was induced with 10 µM L-(+)-arabinose, and pat expression was induced with various concentrations of inducer, as indicated. Growth curves were performed using a Powerwave XS2 microplate reader (Bio-Tek Instruments) at 37°C with shaking in triplicate in three independent experiments. Strains analyzed: *iolR*⁺ (JE6583), *iolR::cat*⁺ (JE10713), *iolR::cat*⁺ / pVOC (JE16934), *iolR::cat*⁺ / piolR⁺ (JE16935), *pat*:MudJ (JE7449), *pat*:MudJ *iolR::cat*⁺ (JE10714), *cobB*:MudJ (JE2845), *cobB::MudJ iolR::cat*⁺ (JE14972), *iolR::cat*⁺ / *ppat*⁺ (JE18927), and *iolR::cat*⁺ / *pcobB*⁺ (JE18891). Error bars represent standard deviation. pVOC, vector-only control.
triangles) was reversed when *pat* was inactivated (Fig. 3.10C, open triangles). As expected, a *cobB* strain failed to grow on 10 mM acetate because acetylated Acs could not be reactivated by deacetylation (Fig. 3.10C, solid circles). Although inactivation of *iolR* presumably reduced Pat levels in the *cobB* strain (by lowering the expression of *pat*), the reduced level was apparently sufficient to keep Acs acetylated (i.e., inactive), thus growth was not restored (Fig. 3.10C, open squares). Additionally, inactivation of *pat* in an otherwise wild-type background had minimal effect on growth likely caused by an excess of Acs activity due to CobB deacetylation (2).

Secondly, if the net result of the change in Pat:CobB ratio in the *iolR::cat* strain was an increase in acetylated, inactive Acs, an increase the level of CobB sirtuin deacetylase in the *iolR::cat* strain would restore Acs to its active, deacetylated state, and consequently growth on 10 mM acetate would occur. This prediction was confirmed, as shown in figure 3.10D.

Thirdly, if the growth defect of the *iolR::cat* strain on 10 mM acetate was caused by a change in the level of Acs activity, it followed that overexpression of *acs* in the Δacs *iolR::cat* strain would restore growth on 10 mM acetate. Results obtained using control strains are shown and as expected, the Δacs strain failed to grow on 10 mM acetate (Fig. 3.11A, inverted triangles), and growth was restored by expression of *acs in trans* (Fig. 3.11A, circles). Shown in figure 3.11B is the effect of ectopic synthesis of Acs WT in the Δacs *iolR::cat* strain. Wild-type growth of the Δacs *iolR::cat* strain on 10 mM acetate was observed upon induction of *acs* expression (20 µM L-(-)-arabinose) (Fig. 3.11B, circles). Unsurprisingly, excessive levels of Acs (100 µM L-(-)-arabinose) (Fig. 3.5, open squares) had a deleterious effect on growth, as reported elsewhere (8).
**Figure 3.11. Induction of acs expression restores growth of an iolR strain on 10 mM acetate.** Growth of an iolR::cat+ strain containing acs expressed ectopically under the control of an L-(+)-arabinose-inducible promoter was examined in minimal medium containing acetate (10 mM). Control strains are shown in (A). The effects of acs induction are shown in (B). Growth curves were performed using a Powerwave XS2 microplate reader (Bio-Tek Instruments) at 37°C with shaking in triplicate in three independent experiments. Strains analyzed: iolR+ (JE6583), iolR::cat+ (JE10713), Δacs (JE7758), Δacs / pacs+ (JE9912), Δacs iolR::cat+ / pacs+ (JE16596). Expression of acs was induced with 0 µM (open triangles), 20 µM (open circles), or 100 µM (black circles (A), open squares (B)). Error bars represent standard deviation.
**Activity of Acs is decreased in an iolR::cat+ strain.** If the growth phenotype of the $iolR::cat^+$ strain on acetate was due to lower Acs activity, we should be able to detect differences in Acs activity in cell-free extracts. Indeed, a reproducible and statistically significant reduction (~25%) in Acs activity was found in cell-free extracts of the $iolR::cat^+$ strain relative to extracts of the $iolR^+$ strain (Fig. 3.12).

**IolR controls expression of acs and cobB.** Since ectopic expression of $acs$ restored growth of the $iolR::cat^+$ mutant, we surmised that $acs$ expression was lower in the mutant than in the wild-type strain. To address this possibility we used an $acs$-lacZ$^+$ reporter fusion to determine whether IolR was also involved in the regulation of $acs$ in *S. enterica*. Since alterations in $pat$ expression were likely to affect CobB levels, we also used a $cobB$-lacZ$^+$ fusion to assess the effect of the absence of IolR on $cobB$ expression. Data obtained from experiments with the above-mentioned transcriptional reporters support the idea that IolR somehow activated expression of both genes (Fig. 3.13). In the absence of IolR, expression of $acs$ (Fig. 3.13A) and $cobB$ (Fig. 3.13B) was reduced on average by 40% ($acs$) or 22% ($cobB$), respectively. The effect of IolR on $acs$ and $cobB$ expression was confirmed using qRT-PCR, and showed a >2-fold and >3-fold down-regulation, respectively, of the transcripts in an $iolR::cat^+$ strain compared to wild type (Fig. 3.13C). At present, it is unclear whether the effect of IolR on Acs and CobB levels is direct or indirect.

**Glucose differentially affects pat, cobB, and acs expression.** Due to the previously established role of the catabolite repressor protein (Crp) in the regulation of the *E. coli pat* homologue (*pka*) (15) we examined the effect of catabolite repression on genes encoding the RLA system, ± $iolR$. 


Figure 3.12. Activity of acetyl-CoA synthetase (Acs). The activity of Acs from whole cell extracts of an iolR+ or iolR::cat+ strain grown on acetate (10 mM) was measured using a coupled NADH-consuming spectrophotometric assay (36). The strains analyzed: iolR+ (JE6583) and iolR::cat+ (JE10713). Samples were analyzed in triplicate. Error bars represent standard deviation. Unpaired t test gave a P value of 0.002.
Figure 3.13. IolR controls expression of \textit{acs} and \textit{cobB}. Activity of \textit{acs-lacZ} and \textit{cobB-lacZ} reporters was assessed in backgrounds ± \textit{iolR} to measure \textit{P}_{\textit{acs}} and \textit{P}_{\textit{cobB}} activity. Cultures were grown at 37°C in NB medium (A, B). Optical density (650 nm) and β-galactosidase activity (420 nm) were measured hourly. The data were obtained from individual cultures performed in triplicate. Error bars represent standard deviation. Strains analyzed: \textit{cobB::MudJ (JE2845), cobB::MudJ \textit{iolR::cat} (JE14972), pACS3 \textit{P}_{\textit{acs}} (JE4637), and \textit{iolR::cat} / pACS3 \textit{P}_{\textit{acs}} (JE14962). C. qRT-PCR showed a 2.6-fold down-regulation in \textit{acs} activity, and a 3-fold decrease in \textit{cobB} activity in an \textit{iolR} strain relative to the \textit{iolR} strain. Wild-type transcript level is set at 1, indicated by the dashed line. Error bars represent standard deviation.
Expression of *pat* in cells grown in NB + glucose medium was reduced ~ 2-fold relative to the expression of *pat* in cells grown in NB lacking glucose (Fig. 3.14, compare black bars). This suggested that *pat* expression was subjected to catabolite repression, an idea that was further explored. Regardless of the presence of glucose in the medium, the absence of IolR reduced *pat-lacZ* expression 30-40% (Fig. 3.14A). The absence of IolR had a small but reproducible negative effect (~20%) on *cobB* expression in the absence of glucose, an effect that was magnified to ~40% when glucose was added (Fig. 3.14B). Significantly, in contrast to *pat* expression, the expression of *cobB* increased ~30% when glucose was added, suggesting that unlike *pat*, expression of *cobB* was not subject to catabolite repression (Fig. 3.14B, black bars).

In *E. coli*, Crp, controls the expression of *acs* (37). Consistent with the idea that in *S. enterica* *acs* expression is controlled by catabolite repression, transcription of *acs* was reduced 80% when glucose was present (Fig. 3.14C, compare black bars). IolR function also appeared to be important for the activation of *acs* in medium devoid of glucose, with ~50% reduction in *acs* expression in the *iolR::cat* strain relative to the wild-type (Fig 3.14C, NB medium). In the presence of glucose, expression of *acs* ± *iolR* was very similar (Fig. 3.14C, NB + glucose).

**Crp activates *pat* expression.** The effect of Crp on *pat* expression was examined in cultures grown in NB + ribose (10 mM). P<sub>pat</sub> activity in the ΔiolR crp::cat<sup>+</sup> strain was 2.5-fold lower than the P<sub>pat</sub> activity measured in a strain containing wild-type *iolR* and *crp* alleles (Fig. 3.15). This decrease in activity was similar to the one measured in a strain lacking *iolR*, and was restored when *crp* was provided in trans. P<sub>pat</sub> activity was slightly lower in the ΔiolR crp::cat<sup>+</sup> strain compared to strains lacking only *crp* or *iolR* (Fig. 3.15). Collectively, the data indicated that Crp was required for wild-type levels of *pat* expression in *S. enterica*. Pat-dependent acetylation of Crp was tested; however, the data indicate that Pat does not acetylate Crp (Fig. 3.4).
Figure 3.14. Glucose differentially controls expression of pat, cobB, and acs. Cultures were grown at 37°C in NB medium ± glucose (10 mM). Optical density (650 nm) and β-galactosidase activity (420 nm) were measured at mid-log phase (OD$_{650}$ ~0.7) to assay for pat and cobB promoter activity in a pat-lacZ$^+$ or cobB-lacZ$^+$ strain backgrounds ± iolR. The data are the average of two independent experiments from individual cultures performed in triplicate. Strains analyzed: pat::MudJ (JE7449), pat::MudJ iolR::cat$^+$ (JE10714), cobB::MudJ (JE2845), cobB::MudJ iolR::cat$^+$ (JE14972), pACS3 P$_{acs}$ (JE4637), and iolR::cat$^+$ / pACS3 P$_{acs}$ (JE14962). Error bars represent standard deviation.
Figure 3.15. Crp activates pat expression. Cultures were grown at 37°C in NB medium with ribose (10 mM). Optical density (650 nm) and β-galactosidase activity (420 nm) were measured at the peak of pat expression (OD$_{650}$ ≈ 0.7) to assay for pat promoter activity in a pat-lacZ$^+$ strain background. Plasmids were induced with L-(-)-arabinose (100 µM). The data presented are the average of two independent experiments from individual cultures performed in triplicate. Strains analyzed: pat::MudJ (JE7449), pat::MudJ iolR::cat$^+$ (JE10714), pat::MudJ crp::cat$^+$ (JE16743), pat::MudJ ΔiolR crp::cat$^+$ (JE16744), pat::MudJ iolR::cat$^+$ / pVOC (JE10727), pat::MudJ iolR::cat$^+$ / pIOLR1 (JE10728), pat::MudJ crp::cat$^+$ / pVOC (JE16771), and pat::MudJ crp::cat$^+$ / prcp$^+$ (JE17322). Error bars represent standard deviation. pVOC, vector-only control.
DISCUSSION

In S. enterica, IolR regulates the RLA system. The chief finding from the studies reported herein is that the IolR protein controls and integrates the expression of pat, cobB, and acs (Figs. 2, 8). Although we do not yet understand the molecular details of how IolR integrates the expression of the above-mentioned genes, collectively our in vivo genetic evidence supporting this claim is compelling (Figs. 5, 6). Furthermore, in vitro data obtained support the conclusion that IolR directly interacts with the pat promoter (Figs. 3, 4). Whether or not the effect of IolR on cobB and acs expression is direct remains to be determined.

IolR function is needed for growth on 10 mM acetate, which require RLA and Acs. IolR function is necessary for optimal growth on 10 mM acetate (Fig. 3.10) and the growth defect of an iolR strain suggests that IolR regulation of pat and cobB impacts the levels of Acs activity in the cell (Fig. 3.13). This conclusion is supported by data showing that the ectopic expression of acs complements growth of an iolR strain (Fig. 3.11). The subtle effects of the absence of IolR on pat, cobB, and acs expression make it difficult to determine the precise magnitude of the changes in Pat, CobB and Acs protein levels, and in the case of Acs there is also a need to distinguished between acetylated vs. non-acetylated protein. Our attempts to gain insights into these changes using Western blot analysis were unsuccessful due to the lack of required sensitivity to define the magnitude of the predicted changes (data not shown). However, the lower levels of Acs activity present in cell-extracts of the iolR strain (Fig. 3.12) support our conclusions.

The differentially repressive effect of glucose on pat and cobB expression is needed to ensure sufficient activation of acetate by Acs. The differential effect that glucose has on the expression of pat and cobB (Fig. 3.14) can be explained by considering the acetogenic nature of
glucose. During glucose catabolism, excess acetyl-CoA is diverted through the acetate kinase (AckA) / phosphotransacetylase (Pta) pathway yielding ATP via substrate-level phosphorylation and releasing CoA, which is needed by the pyruvate dehydrogenase to make more acetyl-CoA. In *S. enterica*, the assimilation of glucose-derived acetate is known to require the functions of AckA/Pta and Acs. Notably, during growth on glucose less Acs is made because Crp cannot fully activate *acs* expression when cAMP levels are low (38). One plausible way the cell can ensure that the limited amount of Acs made by the cell in the presence of glucose supplies enough acetyl-CoA to support growth is to lower the expression of *pat* while increasing *cobB* expression (Fig. 3.14). By so doing, less Acs becomes acetylated, and whatever Acs is acetylated is reactivated by the higher levels of CobB deacetylase made under these conditions. Given that *pat* expression is reduced whenever IolR or Crp are not made, it was surprising to find out that the absence of both regulators did not have an additive effect on *pat* expression (Fig. 3.15). This result could simply reflect the basal level of *pat* expression, or possibly, a more complex regulatory network in which Crp affects regulation of *iolR*. Such an idea would not be unprecedented, since in *B. subtilis*, the *iol* operon, including *iolR*, is under the control of catabolite repression (39).

**Why does IolR activate expression of the RLA system?** The dual regulatory role of IolR as a repressor of *iol* genes and an activator of *pat*, *cobB*, and *acs* is intriguing. The ability of IolR to function as an activator is not unprecedented. A report in *C. glutamicum* demonstrated that IolR activates expression of *pck* (encodes phosphoenolpyruvate carboxykinase) (33). IolR also regulates *srfJ*, a type-III secretion effector protein, in *Salmonella* (40), a result that is not surprising as *srfJ* lies within the *iol* genomic island, which is regulated by IolR.
Under conditions where acetate and myo-inositol are simultaneously present in the environment, the cell needs to integrate their metabolism. In thinking about this issue one must consider that acetate and myo-inositol are catabolized at substantially different rates, as acetate enters central metabolism as soon as it is converted into acetyl-CoA. In contrast, myo-inositol degradation requires the generation of a signal that upon binding to IolR lifts repression of the iol genes and synthesis of myo-inositol-degrading enzymes can occur. Generation of the signal needed to transcribe the iol genes is apparently an extensive process (Fig. 3.5A, comparing the differences in lag phases between the iolR+ and iolR strains). The use of IolR to integrate myo-inositol and acetate metabolism would be an efficient way to generate as much acetyl-CoA for anabolic purposes as possible while maintaining the capability of modulating the activity of the RLA system for the purpose of controlling the level of Acs activity.

Is myo-inositol utilization regulated by RLA? Recently, the total population of acetylated proteins (acetylome) of the myo-inositol utilizing bacteria B. subtilis and Erwinia amylovora were reported (9, 10). Notably, two enzymes involved in the degradation of myo-inositol, the malonate semialdehyde dehydrogenase (IolA) and carbohydrate kinase (IolC) were amongst the acetylated proteins identified. It is unclear whether the activity of IolA and/or IolC is controlled by RLA in either organism or in S. enterica. If any Iol proteins are under RLA control, it could provide a link between IolR regulation of RLA and RLA involvement in myo-inositol utilization.

MATERIALS AND METHODS

Culture media and chemicals. Nutrient broth (NB, Difco) containing NaCl (85 mM) was used as rich medium. The minimal medium used was no-carbon essential (NCE) minimal medium (41) containing MgSO_4 (1 mM), Wolfe’s trace minerals (1x) (42), and carbon source
[acetate (10 or 50 mM), glycerol (22 mM), or myo-inositol (50 mM)]. Antibiotics were added at following concentrations: tetracycline, 20 µg ml⁻¹; chloramphenicol 20 µg ml⁻¹; kanamycin, 50 µg ml⁻¹; and ampicillin, 100 µg ml⁻¹. When added, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) was present at 40 µg ml⁻¹ and the calcium chelator ethyleneglycol tetraacetic acid (EGTA) was present at 10 mM. Chemicals were purchased from Sigma-Aldrich.

**Bacterial strains.** All strains studied were derivatives of *S. enterica* serovar Typhimurium strain LT2 (unless noted, Table 3.1). Tn10d(tet⁺) refers to the transposase-defective mini-Tn10 (Tn10Δ16Δ17) (31); MudJ (refers to MudI1734) was used to generate a strain carrying the MudJ (kan⁺ lacZ⁺) element under the control of the pat promoter (29). The MudJ element inactivated the gene at the same time that it fused the expression of lacZ to the promoter of choice, either P_{pat} or P_{cobB}. All primers used in this study are listed in Table 3.2 (IDT, Coralville, Iowa).

Strains carrying a deletion of iolR or crp were constructed following described protocols (43). To construct the iolR and crp deletions in strain JE7449 (pat::MudJ), P22 phage was propagated using strain JE10713 (iolR::cat⁺), or strain JE16466 (crp::cat⁺) and used to transduce strain JE7449 to chloramphenicol resistance. Bacteriophage P22-mediated transductions were performed as described (44), using the high-frequency general transducing mutant of bacteriophage P22 HT105/1 int-210 (45). Phage-free, phage-sensitive transductants were isolated on non-selective green indicator plates as described (46).

**Plasmid construction.** Plasmids are listed in Table 3.1. To construct the arabinose-inducible iolR plasmid (pIOLR1), the 834-bp iolR allele from *S. enterica* and its ribosome-binding site (849 bp total) were amplified and inserted pBAD30 (30) using primers EcoRI and XbaI restrictions sites (pBAD30 iolR 5’ and pBAD30 iolR 3’). Primers were synthesized at the Biotechnology Center of the University of Wisconsin-Madison. To construct the plasmid for the
Table 3.2. Primers used in this study

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain construction</strong></td>
<td></td>
</tr>
<tr>
<td><strong>iolR</strong> – 5’DEL</td>
<td>5’-TGCTTAATATGGTTATTTACGAAATTTTCTATTAGAGTATCATGCATGGTGAGCTGGCTTTC-3’</td>
</tr>
<tr>
<td><strong>iolR</strong> – 3’DEL</td>
<td>5’-TTCACCAATGCGGATGATCGCTAAATACGATCATCGGCTTTGTTTTTACATATGAATATCCTCCTTAG-3’</td>
</tr>
<tr>
<td><strong>crp</strong> – 5’DEL</td>
<td>5’-GCTCTGGAGACAGCTTATAACAGAGGATAACCGCGCATG</td>
</tr>
<tr>
<td><strong>crp</strong> – 3’DEL</td>
<td>5’-AAATGGGCGCATGATAAAACCGGCCATTCTGACGGAATTACATATGAAATATCCTCCTTAG-3’</td>
</tr>
<tr>
<td><strong>Cloning</strong></td>
<td></td>
</tr>
<tr>
<td>pBAD30 <strong>iolR</strong> 5’</td>
<td>5’-AAAAAAGAATTCATTAGAGTATCATGCATGTCTAAACAT</td>
</tr>
<tr>
<td>pBAD30 <strong>iolR</strong> 3’</td>
<td>5’-AAAAAATCTAGATTACTCCGTCGCCAGCGCCAGTGAAA</td>
</tr>
<tr>
<td>pTEV6 <strong>iolR</strong> 5’</td>
<td>5’-AAAAAAGGTACCATGTCTAAACATCAAACTCAACT</td>
</tr>
<tr>
<td>pTEV6 <strong>iolR</strong> 3’</td>
<td>5’-AAAAAAAAGCTTTTACTCCGTCGCCAGCGCCTA-3’</td>
</tr>
<tr>
<td>pBAD30 <strong>crp</strong> 5’</td>
<td>5’-GAATTCGCTAGCCCAAAAAAACGG</td>
</tr>
<tr>
<td>pBAD30 <strong>crp</strong> 3’</td>
<td>5’-AAGCTTGGCTGTTTTGGCGGATGA</td>
</tr>
<tr>
<td>pTEV16 <strong>crp</strong> 5’</td>
<td>5’-NNGCTCTTCNAGCATGGTGCTTGGCAAACCGCAAACAG-3’</td>
</tr>
<tr>
<td>pTEV16 <strong>crp</strong> 3’</td>
<td>5’-NNGCTCTTCNNTTAATACGGGTGCCGTAGACGACGA-3’</td>
</tr>
<tr>
<td><strong>EMSA and DNA footprinting</strong></td>
<td></td>
</tr>
<tr>
<td>P <strong>pat</strong> 150 FAM</td>
<td>5’-CGAGGCAGGCAACCAGTTA-3’</td>
</tr>
<tr>
<td>P <strong>pat</strong> 150 Rev</td>
<td>5’-CATACCGGTATTCCCCCTACACA-3’</td>
</tr>
<tr>
<td>P <strong>iolR</strong> 175 FAM</td>
<td>5’-CTGATCCTGTGCAATTATG-3’</td>
</tr>
<tr>
<td>P <strong>iolR</strong> 175 Rev</td>
<td>5’-CATACCGGTATTCCCCCTACACA-3’</td>
</tr>
<tr>
<td>P <strong>pat</strong> 45 FAM</td>
<td>5’-CAACCAGTTAGTGGAACATTTTCACCCCGTTCCAAAACGCCTATCT-3’</td>
</tr>
<tr>
<td>P <strong>pat</strong> 45</td>
<td>5’-CAACCAGTTAGTGGAACATTTTCACCCCGTTCCAAAACGCCTATCT-3’</td>
</tr>
<tr>
<td>P <strong>pat</strong> 45 Rev</td>
<td>5’-AGATAGCGCGTTTTGAAACGGGTGAAATGTTCACCTAAGTGGTTG-3’</td>
</tr>
<tr>
<td>P <strong>pat</strong> 382 FAM</td>
<td>5’-CCCAGGACGCATGAAAATGTT-3’</td>
</tr>
<tr>
<td>P <strong>pat</strong> 382 Rev</td>
<td>5’-TTCATTTACTGCGCCACTAC-3’</td>
</tr>
<tr>
<td><strong>qRT-PCR</strong></td>
<td></td>
</tr>
<tr>
<td><strong>acs</strong> qRT-PCR F</td>
<td>ACCGACTCAGCTGGCAAAAAC</td>
</tr>
<tr>
<td><strong>acs</strong> qRT-PCR R</td>
<td>CTTGACGGCTTTCTCCACC</td>
</tr>
<tr>
<td><strong>cobB</strong> qRT-PCR F</td>
<td>CTGTCGATGGCGGATATTG</td>
</tr>
<tr>
<td><strong>cobB</strong> qRT-PCR R</td>
<td>GGCGTATGTCTTCTCTC</td>
</tr>
<tr>
<td><strong>pat</strong> qRT-PCR F</td>
<td>GATTGGAGCCGCCTTATG</td>
</tr>
<tr>
<td><strong>pat</strong> qRT-PCR R</td>
<td>CTGAAAATCGCTTTCTGCC</td>
</tr>
<tr>
<td><strong>gyrA</strong> qRT-PCR F</td>
<td>AACACCCATGACACCATC</td>
</tr>
<tr>
<td><strong>gyrA</strong> qRT-PCR R</td>
<td>GGCTGATACGCTTTGTC</td>
</tr>
</tbody>
</table>

*fluorophore 6-carboxyfluorescein (6FAM)*
overexpression of *iolR* (pIOLR3), the gene was inserted into plasmid pTEV6 (28) using *KpnI* and *HindIII* restriction sites (pTEV6 *iolR* 5') and (pTEV6 *iolR* 3'). The resulting overexpression plasmid, pIOLR3, directed the synthesis of IolR protein with an *N*-terminal hexahistidinemaltose-binding protein (His$_6$-MBP) tag cleavable by tobacco etch virus (TEV) protease (47). The polymerase incomplete primer extension (PIPE) cloning method (48) was used to construct the arabinose-inducible plasmid for the expression of *crp* (pCRP2). The 633-bp *crp* gene from *S. enterica* and its ribosome-binding site (643-bp total) were amplified using primers (pBAD30 *crp* 5') and (pBAD30 *crp* 3'). Plasmid pTEV5 (28), which directs the synthesis of the protein with a cleavable *N*-terminal hexahistidine tag, was engineered with BspQI restriction sites (pTEV16, C. M. VanDrisse & J. C. Escalante-Semerena, unpublished data, (49)), and used for overexpression of CRP, using primers (pTEV16 *crp* 5') and (pTEV16 *crp* 3') (pCRP3). DNA sequencing was used to verify the cloned sequences in all plasmids.

**Isolation of Tn10d(tet$^+$) insertion in iolR.** A pool of ~100,000 *S. enterica* strains each assumed to contain one Tn10d(tet$^+$) element randomly inserted in the chromosome was generated as described (50). A P22 lysate grown on this pool of strains was used to transduce recipient strain JE7449 (*metE ara pat::MudJ*) to tetracycline resistance (*Tc$^R$*) on NB agar plates containing X-gal and EGTA. Colonies displaying altered coloration were freed of phage and P22 phage lysates were generated to use as donors in crosses with the parental JE7449 strain. The location of the insertion was determined in the reconstructed strains by sequencing the DNA flanking the Tn10d(tet$^+$) element using a PCR-based protocol with degenerate primers (51). DNA sequencing was performed using BigDye® Terminator v3.1 protocols (Applied Biosystems).
**Growth studies.** Cultures were grown overnight at 37°C in NB and used to inoculate medium with (5% v/v) in a volume of 200 µl per well of a 96-well plate. NCE minimal medium containing MgSO₄ (1 mM), Wolfe’s trace minerals (1x), and carbon source [acetate (10 mM, 50 mM), glycerol (22 mM), or *myo-*inositol (50 mM)] was used. Plasmids were induced with L-(+)-arabinose, as described. Plates were incubated at 37°C in a Powerwave Microplate Reader (Bio-Tek Instruments). Data were analyzed using Prism v6 (GraphPad) software.

**β-Galactosidase assays.** β-Galactosidase activities were determined as described (52). Three independent overnight cultures were grown in NB + ampicillin, sub-cultured (1:100, v/v) into 200 ml of medium + ampicillin and induced with 100 µM L-(+)-arabinose. Cultures were grown at 37°C in NB, NB + ribose (10 mM), NB + glucose (10 mM), acetate (10 mM), or *myo-*inositol (55 mM) medium. Acetate cultures were inoculated with 2.5% (v/v) of overnight culture.

**qRT-PCR.** Cultures of strains JE10713 (*iolR::cat*) and JE6583 (*iolR*) were grown in NB to an OD₆₀₀ of 0.6. RNA extraction was performed as described (53). cDNA synthesis was performed using iScript cDNA synthesis kit (Bio-Rad). qRT-PCR reactions were performed using Fast SYBR® Green Master Mix (Thermo Fischer) and a 7500 Fast Real-Time PCR System (Applied Biosystems).

**Overproduction and purification of proteins.** Plasmid pIOLR3 and pCRP3 were transformed into *E. coli* C41(λDE3), and overnight cultures sub-cultured 1:100 (v/v, inoculum/medium) into 4 L of LB and ampicillin (150 µg ml⁻¹). Cultures were grown at 37°C with shaking to an OD₆₀₀ of 0.6, induced with isopropyl-β-D-thiogalactopyranoside (IPTG, 1 mM), and shaken overnight at 28°C. Cells were harvested by centrifugation at 8,394 × g for 15 min at 4°C. Cell lysate was re-suspended in bind buffer A [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (50 mM, pH 7.5), 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 2 min (2 s, 50% duty, setting 4) in 30 s intervals on ice using a 550 Sonic Dismembrator (Fisher Scientific). Clarified cell lysates were obtained after centrifugation for 45 min at 4°C at 43,667 × g.

**IolR Purification:** Protein was resolved using ÄKTA FPLC system (GE Healthcare) equipped with a HisTrap 10 mL column. After loading the protein, the column was washed with 7 column volumes (CV) of bind buffer, followed by 10 CV of wash buffer [HEPES buffer (50 mM, pH 7.5), NaCl (500 mM), and imidazole (40 mM)]. rTEV-cleavable His₆-MBP tagged IolR eluted off the column using a 30 min linear gradient of elution buffer [HEPES buffer (50 mM, pH 7.5), NaCl (500 mM), and imidazole (500 mM)]. To cleave the His₆-MBP tag from IolR, recombinant His₆-TEV protease (rTEV) was purified as described (47), and added to the His₆-MBP-IolR solution containing dithiothreitol (DTT, 1 mM) at a 1:100 [rTEV:tagged IolR] ratio for 3 h at RT. Cleaved IolR protein was dialyzed against dialysis buffer A [HEPES (50 mM, pH 7.5), NaCl (500 mM), and tris(2-carboxyethyl)phosphine hydrochloride (TCEP, 0.5 mM)] and EDTA (0.5 mM). The second purification was carried out using the purification described above, and un-tagged IolR eluted off the column during the wash step. IolR was dialyzed in dialysis buffer B [HEPES (50 mM, pH 7.5), NaCl (350 mM), and 20% (v/v) glycerol], followed by dialysis buffer C [HEPES (50 mM, pH 7.5), NaCl (250 mM) and 20% (v/v) glycerol], and lastly into storage buffer A [HEPES (50 mM, pH 7.5), NaCl (150 mM) and 20% (v/v) glycerol]. IolR was drop-frozen in liquid nitrogen and stored at -80°C. The IolR protein was purified to 96% homogeneity as determined using Total Lab v2005 software.

**Crp Purification:** Protein was 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
bind buffer. Followed by that His₆-tagged Crp eluted in the same buffer system that contained 500 mM imidazole. Protein cleaved with rTEV (1:50 mg:mg ratio) for 3 hr at 25°C, then dialyzed at 4°C against dialysis buffer A, dialysis buffer B, and dialysis buffer C, as described above. Cleaved protein ran over equilibrated column (Crp eluted in flow-through), followed by 10 CV bind buffer, 5 CV of wash buffer, and 5 CV of elution buffer. Protein was dialyzed at 4°C against dialysis buffer E [50 mM HEPES, 400 NaCl (pH 7.0)], dialysis buffer F [50 mM HEPES, and 250 NaCl (pH 7.0)], and storage buffer B [50 mM HEPES, 150 NaCl, and 20% glycerol (pH 7.0)]. Crp was drop-frozen in liquid nitrogen and stored at -80°C.

**Analytical gel filtration.** Experiments were performed at 4°C with a 0.5 ml/min flow rate. Per run, a sample volume of 100 µl containing 150 µg of IolR was injected onto a Superdex 200 HR 10/30 gel filtration column (GE Healthcare) attached to an ÄKTA FPLC system that was equilibrated with HEPES (50 mM, pH 7.5) containing NaCl (500 mM). A calibration standard containing a mixture of molecular masses (Bio-Rad Laboratories; vitamin B₁₂ (1.35 kDa), equine myoglobin (17 kDa), chicken ovalbumin (44 kDa), bovine γ-globulin (158 kDa), and thyroglobulin (670 kDa)) was used to generate a standard curve. Analysis was performed using UNICORN v4.11 software. Data were graphed and analyzed using Prism v6 (GraphPad) software. Typical linear regression analyses of the standard curves yielded $r^2$ values of 0.98.

**DNA-binding assays.** Electrophoretic mobility shift DNA-binding assays were performed using probes with a fluorophore 6-carboxyfluorescein (6FAM) attached at the 5’ prime end. Probes were generated by PCR amplification from the S. enterica chromosome. The probe containing the putative promoter region of pat ($P_{pat}$, 150 nt) was generated using a 6FAM-5’-labeled primer $P_{pat}$ 150 FAM and 3’ primer $P_{pat}$ 150 Rev. The positive control probe $P_{iolR}$ (175 nt) was generated using primers 6FAM-5’-labeled ($P_{iolR}$ 175 FAM) and 3’ primer ($P_{iolR}$ 175 Rev).
The negative control probe $P_{argS}$ (196 nt) was generated using the primers described elsewhere (35). Probes were purified with the Wizard® SV Gel and PCR Clean-Up System (Promega). Reactions (10 µl) contained 6FAM-5’-labeled probe (50 ng), Tris-HCl buffer (50 mM, pH 7.5), KCl (50 mM), MgCl$_2$ (10 mM), EDTA (0.5 mM), glycerol (10%, v/v), and IolR protein [shown in molar excess to probe (2.5-10 pmol)]. Reactions were incubated at 22°C for 45 min and resolved on a Criterion 10% native polyacrylamide gel (BioRad) in 0.5X Tris-borate-EDTA buffer (TBE, pH 8.3; Tris-HCl (45 mM), boric acid (45 mM), EDTA (1 mM)). The signal was detected using a Typhoon Trio+ Variable Mode Imager (GE Healthcare) with ImageQuant v5.2 software.

**DNA-footprinting analysis.** The $pat$ promoter ($P_{pat}$) was amplified from *S. enterica* genomic DNA using 6-FAM-5’-labeled primer ($P_{pat}$ 382 FAM) and 3’ primer ($P_{pat}$ 382 Rev). Appropriate digestion levels were tested using varying concentrations of both fluorescently labeled probe and DNase. Reactions contained varying amounts of IolR or bovine serum albumin (BSA, negative control), Tris-HCl (50 mM, pH 7.5), KCl (50 mM), MgCl$_2$ (10 mM), EDTA (0.5 mM), glycerol (10%, v/v), and incubated for 10 min at 25°C. The labeled DNA probe (120 ng) was added to the reaction and incubated for 20 min at 25°C, followed by addition of DNase for 5 min at 25°C and heat-inactivated for 10 min at 78°C. Reaction mixtures were purified using the MinElute PCR Purification Kit (Qiagen). Samples were analyzed with an Applied Biosystems 3730 DNA Analyzer (Plant-Microbe Genomics Facility, Ohio State University) set to default run module for LIZ600 dye, with 0.1 µl of size standard (LIZ600), 0.5-1.0 µl of sample and 9 µl of HiDi per well. Electropherogram overlays comparing IolR and BSA as a negative control per condition were generated using GeneMapper v4.0 software and putative binding sites were analyzed by aligning the sequenced probe data (electropherogram data not shown; available upon request).
**In vitro acetylation assay.** Protein acetylation assays were performed as described, using radiolabeled [1-\(^{14}\)C] Ac-CoA (2, 7, 14). Briefly, reactions contained SelolR or SeCrp (5 \(\mu\)M) with or without SePat (3 \(\mu\)M). SeAcs was used as a positive control. Laboratory stocks of purified SePat and SeAcs proteins were used in these studies. Reactions were resolved and visualized by SDS-PAGE. Radiolabeled proteins were visualized using a Typhoon Trio+ Variable Mode Imager (GE Healthcare) with ImageQuant v5.2 software.

**Acetyl-CoA synthetase assay.** Acs activity was measured using an NADH-consuming assay[53]. Cultures of the *iolR\(^{+}\) (JE6583) and *iolR::cat\(^{+}\) (JE10713) strains were grown in NCE minimal medium with acetate (10 mM) at 37°C with shaking to an OD\(_{600}\) of 0.2. Cell cultures were harvested by centrifugation at 6,000 x g for 30 min and cell pellets were stored at -80°C until used. Cell paste was re-suspended in HEPES buffer (pH 7.5), plus lysozyme (1 mg ml\(^{-1}\)), DNase I (25 \(\mu\)g ml\(^{-1}\)), phenylmethanesulfonyl fluoride (PMSF, 0.5 mM), and SIGMAFAST protease inhibitor cocktail tables (Sigma)]. Cells were lysed by sonication for 1 min (2 s, 50% duty, setting 80) on ice using a QSonica Q55 cell disruptor (QSonica LLC). Lysed cells were centrifuged at 16,000 x g at 4°C for 20 min. The lysed cell extract was concentrated and dialyzed simultaneously using Amicon \(^{\circledR}\) Ultra concentrators (Millipore) with a 30 kD MWCO (Acs is \(\sim\)72 kD), per the manufacturer’s instructions. Samples were washed twice with 5 ml of HEPES buffer (pH 7.5), and total protein concentration was measured using a NanoDrop device (Thermo Scientific). Reactions 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 (PEP, 3 mM), NADH (0.1 mM), pyruvate kinase (1 U), myokinase (5 U), lactate dehydrogenase (1.5 U) and acetate (0.2 mM). Reactions were started by addition of the treated cell extract (50 \(\mu\)g). Absorbance at 340 nm was monitored in a 96-well plate using the
Spectramax Plus UV-visible spectrophotometer (Molecular Devices). Enzyme activity (of total protein concentration) was calculated as described [53]. Samples were analyzed in triplicate. Error bars represent standard deviation. Statistics were analyzed using unpaired t test (GraphPad Prism v6 software).

ACKNOWLEDGEMENTS

This work was supported by PHS grant R01-GM62203 to J.C.E.-S. S.T. was supported by the PHS Molecular Biosciences Training Grant T32-GM07215 and NRSA Predoctoral Fellowship F31-GM083668. We thank Chelsey M. VanDrisse, Flavia G. Costa, and Michael Ullmer for technical assistance and Michael Zianni (Plant-Microbe Genomic Facility, Ohio State University) for DNA footprinting analysis.

REFERENCES


18. **Yamamoto H, Serizawa M, Thompson J, Sekiguchi J.** 2001. Regulation of the *glv* operon in *Bacillus subtilis*: YfiA (GlvR) is a positive regulator of the operon that is repressed through CcpA and Cre. *J. Bacteriol*. **183:**5110-5121.


CHAPTER 4

IN SALMONELLA ENTERICA, THE GCN5-RELATED ACETYLTRANSFERASE MDDA (FORMERLY YNCA) ACETYLATES METHIONINE SULFONE AND METHIONINE SULFOXIMINE, BLOCKING THEIR TOXIC EFFECTS\(^3\)

ABSTRACT

Protein and small molecule acylation are widespread in nature. Many of the enzymes catalyzing acylation reactions belong to the Gcn5-related N-acetyltransferase family (GNAT, PF00583), named after the yeast Gcn5 protein. The genome of Salmonella enterica serovar Typhimurium LT2 encodes 26 GNATs, 11 of which have no known physiological role. Here we provide in vivo and in vitro evidence for the role of the MddA (Methionine derivative detoxifier A; formerly YncA) GNAT in the detoxification of oxidized forms of methionine, including methionine sulfoximine (MSX) and methionine sulfone (MSO). MSX and MSO inhibited growth of an S. enterica ΔmddA strain unless glutamine or methionine was present in the medium. We used an in vitro spectrophotometric assay and mass spectrometry to show that MddA acetylated MSX and MSO. An mddA+ strain displayed biphasic growth kinetics in the presence of MSX and glutamine. Deletion of two amino acid transporters (GlnHPQ and MetNIQ) in a ΔmddA strain restored growth in the presence of MSX. Notably, MSO was transported by GlnHPQ, but not by MetNIQ. In summary, MddA is the mechanism used by S. enterica, to respond to oxidized forms of methionine, which MddA detoxifies by acetyl-CoA-dependent acetylation.

INTRODUCTION

The Gcn5 related N-acetyltransferase (GNAT, PF00583) superfamily of proteins (>10,000 members) is present in all domains of life. GNATs transfer the acetyl group from acetyl-CoA to proteins or small molecules (for reviews see (1, 2). Acetylation targets of GNATs include the N-termini of proteins (3, 4), aminoglycoside antibiotics (5), glutamate (6), spermidine (7), aminoalkylphosphonic acid (8), dTDP-fucosamine (9), and transfer RNAs (10). Some of the first bacterial GNATs characterized were the aminoglycoside N-acetyltransferases from Enterococcus
faecium (5) and Serratia marcescens (11), demonstrating GNAT-dependent acetylation and inactivation of antibiotics.

GNATs provide protection against a myriad of cellular stressors, and the number of stressors controlled by GNATs appears to correlate with the environment encountered by the cell. Therefore, the relevance of GNAT function to cell physiology varies amongst organisms. For example, S. enterica and E. coli each contain ~26 GNATs, yet actinomycetes such as Streptomyces lividans encode up to ~70 putative GNATs, suggesting that S. lividans occupies a more challenging habitat.

At present, there is limited to no information available on the cellular processes several putative S. enterica GNATs may affect. Not surprisingly, the signals that trigger the synthesis of GNATs, the transcription factors involved in sensing such signals, and the determinants of GNAT substrate specificity remain unknown.

In S. enterica, MddA (formerly YncA, STM1590) is a putative GNAT with no characterized function. Homology searches reveal the presence of MddA-like proteins in Pseudomonas aeruginosa (63% identity) and Acinetobacter baylyi (36% identity), and suggest a role for SeMddA in controlling the toxic effects of methionine sulfoximine (MSX) and methionine sulfone (MSO) (Fig. 4.1) (12, 13). Protein structures of SeMddA homologues have been solved in P. aeruginosa (PDB 2J8R) and A. baylyi (PDB 2JLM), showing these enzymes contain the structural core representative of members of the GNAT family (12-14). The P. aeruginosa MddA homologue was solved in complex with MSX and showed a conformational change in the active site upon binding to MSX (13).

MSX is similar in structure to phosphinothricin (PHO) (Fig. 4.1), and at least some MddA homologues have been incorrectly annotated as PHO acetyltransferases (12, 13), an activity
performed by the Bar acetyltransferase (15). The Bar protein is a GNAT of *Streptomyces* spp. involved in protection against a self-produced toxin Bialaphos, a natural herbicide consisting of the tripeptide PHO-Ala-Ala (16). The toxin is activated when PHO, a glutamate analogue, is cleaved from the peptide. Bialaphos is a potent herbicide and plants have been genetically engineered to be resistant by encoding the *bar* gene (15, 17). PHO and MSX both inhibit glutamine synthetase (GlnA), which converts glutamate to glutamine, and plays an important role in the regulation of nitrogen metabolism (18, 19).

MSX is found in the roots and seeds of members of the Connaraceae plant species, and has been identified to be toxic component of these plants (20). While this is the only example of natural occurring MSX, in the late 1940’s MSX was identified as a toxic by-product in bleached flour (21-23). It was produced by addition of nitrogen trichloride, which reacted with wheat proteins in the flour. This had severe effects on individuals that consumed flour treated this way, and around 1950 this method of flour bleaching was discontinued. Researchers have hypothesized that long-time exposure to MSX in processed foods may account for an increase in neurodegenerative disorders in humans such as Alzheimer's disease, Parkinson’s disease, and amyotrophic lateral sclerosis (ALS) (24, 25).

Here we present *in vivo* and *in vitro* evidence that *SeMddA* (methionine derivative detoxifier A) is necessary for survival of *S. enterica* in the presence of MSX and MSO, addition of glutamine or methionine prevents this toxicity, and that the *SeMddA* protein acetylates the amino moiety of MSX. We report that deletion of two amino acid transporters, MetNIQ and GlnHPQ, fully restores growth of a ∆*mddA* strain exposed to MSX. Our data also indicate transport of MSO differs from MSX, as MetNIQ cannot transport MSO. Taken together, these data demonstrate that *SeMddA* acetylates MSX and MSO, thereby blocking their toxic effects.
Figure 4.1. Chemical structures of methionine analogs. (A) methionine (Met); (B) methionine sulfoxide (MS); (C) methionine sulfone (MSO); (D) methionine sulfoximine (MSX); and (E) phosphinothricin (PHO).
MATERIALS AND METHODS

Culture media and chemicals. Nutrient broth (NB, Difco) containing NaCl (85 mM) was used as rich medium. The minimal medium used was no-carbon essential (NCE) minimal medium (26) containing MgSO₄ (1 mM), Wolfe’s trace minerals (1x) (27), and glycerol (22 mM) as the sole source of carbon and energy. When used, antibiotics were added at the following concentrations: tetracycline, 20 µg ml⁻¹; kanamycin, 50 µg ml⁻¹; chloramphenicol, 20 µg ml⁻¹; and ampicillin, 100 µg ml⁻¹. When added to the medium, the calcium chelator ethyleneglycol tetraacetic acid (EGTA) was present at 10 mM, and X-gal was added to a final concentration of 40 µg ml⁻¹. All chemicals were purchased from Sigma-Aldrich unless noted otherwise; Kanamycin, ampicillin, NaCl, and 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (HEPES, Fischer Scientific); tris(2-carboxyethyl)phosphine hydrochloride (TCEP, Soltec Ventures); isopropyl β-D-1-thiogalactopyranoside (IPTG, IBI Scientific); dithiothreitol (DTT, Gold BioTechnology); and [₁⁴C-1]-Acetyl-CoA (Moravek).

Bacterial strains and primers. All strains are derivatives of S. enterica serovar Typhimurium strain LT2 (unless specified), and are listed in Table 4.1. Tn₁₀d(tet⁺) refers to the transposase-defective mini-Tn₁₀ (Tn₁₀Δ₁₆Δ₁₇) (28). All primers used in this study were synthesized by IDT (Coralville, Iowa) and are listed in Table 4.2.

Phage transductions. P22 phage-mediated transduction crosses were performed as described previously (29), using the high-frequency general transducing mutant of bacteriophage P22 HT105/1 int-210 (30, 31). Phage-free, phage-sensitive transductants were isolated on non-selective green indicator plates as described previously (32).
Table 4.1. Strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Reference/source*</th>
</tr>
</thead>
<tbody>
<tr>
<td>JE10079</td>
<td>ara-9 mddA</td>
<td>Laboratory strain</td>
</tr>
<tr>
<td><strong>Derivatives of JE10079</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JE18333</td>
<td>mddA1::cat</td>
<td></td>
</tr>
<tr>
<td>JE18543</td>
<td>mddA1::cat / pNK972</td>
<td></td>
</tr>
<tr>
<td>JE18622</td>
<td>ΔmddA2</td>
<td></td>
</tr>
<tr>
<td>JE18955</td>
<td>pMDD8</td>
<td></td>
</tr>
<tr>
<td>JE18961</td>
<td>mddA1::cat / pMDD8</td>
<td></td>
</tr>
<tr>
<td>JE19029</td>
<td>mddA1::cat / pMDD11</td>
<td></td>
</tr>
<tr>
<td>JE19583</td>
<td>metNI2703::kan</td>
<td></td>
</tr>
<tr>
<td>JE19730</td>
<td>ΔmddA2 metNI2703::kan</td>
<td></td>
</tr>
<tr>
<td>JE20027</td>
<td>ΔmddA2 glnP1561::Tn10Δ16Δ17 (ter')</td>
<td></td>
</tr>
<tr>
<td>JE20064</td>
<td>glnPQ1562::cat</td>
<td></td>
</tr>
<tr>
<td>JE20065</td>
<td>ΔmddA2 glnPQ1562::cat</td>
<td></td>
</tr>
<tr>
<td>JE20067</td>
<td>ΔmddA2 metNI2703::kan&quot; glnPQ1562::cat'</td>
<td></td>
</tr>
<tr>
<td>JE20073</td>
<td>ΔmddA2 glnPQ1562::cat / pGLN2</td>
<td></td>
</tr>
<tr>
<td>JE20329</td>
<td>ΔmddA2 metNI2703::kan&quot; / pMETN1</td>
<td></td>
</tr>
<tr>
<td>JE6583</td>
<td>metE205 ara-9</td>
<td>K. Sanderson via J. Roth</td>
</tr>
<tr>
<td><strong>E. coli strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em> C41(DE3)</td>
<td>ompT hsdS (rB mB) gal λ (DE3) including at least one non-characterized mutation</td>
<td>(1, 2)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMDD7</td>
<td>mddA+ cloned into pKLD66</td>
<td></td>
</tr>
<tr>
<td>pMDD8</td>
<td>mddA+ cloned into pBAD24</td>
<td></td>
</tr>
<tr>
<td>pMDD10</td>
<td>mddA3 cloned into pKLD66 (encodes MddA\textsuperscript{EI82Q})</td>
<td></td>
</tr>
<tr>
<td>pMDD11</td>
<td>mddA3 cloned into pBAD24, (encodes MddA\textsuperscript{EI82Q})</td>
<td></td>
</tr>
<tr>
<td>pGLN2</td>
<td>glnPQ+ cloned into pBAD24</td>
<td></td>
</tr>
<tr>
<td>pMETN1</td>
<td>metNI+ cloned into pBAD24</td>
<td></td>
</tr>
<tr>
<td>pNK972</td>
<td>tpn&quot; bla&quot;</td>
<td>(3)</td>
</tr>
</tbody>
</table>

* All strains and plasmids were constructed during the course of this work, unless otherwise stated.
* pNK972 is a pBR332 derivative carrying the IS10 transposase gene described in (3).
* Tn10Δ16Δ17 (ter') is an abbreviation of Tn10Δ16Δ17 (4).
* pKLD66 is an overexpression vector described in (2).
* pBAD24 is a cloning vector described in (5).
Table 4.2. Primers used in this study.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cloning</strong></td>
<td></td>
</tr>
<tr>
<td>5’ glnPQ pBAD24</td>
<td>NNGCTCTTCTTTATCATGCAGTTTGACTGGAGCGCCATCTCT</td>
</tr>
<tr>
<td>3’ glnPQ pBAD24</td>
<td>NNGCTCTTCTTTATCATGCAGTTTGACTGGAGCGCCATCTCT</td>
</tr>
<tr>
<td>5’ metNI pBAD24</td>
<td>NNGCTCTTCTTTATCATGCAGTTTGACTGGAGCGCCATCTCT</td>
</tr>
<tr>
<td>3’ metNI pBAD24</td>
<td>NNGCTCTTCTTTATCATGCAGTTTGACTGGAGCGCCATCTCT</td>
</tr>
<tr>
<td>5’ mddA pBAD24</td>
<td>NNGCTCTTCTTTATCATGCAGTTTGACTGGAGCGCCATCTCT</td>
</tr>
<tr>
<td>3’ mddA pBAD24</td>
<td>NNGCTCTTCTTTATCATGCAGTTTGACTGGAGCGCCATCTCT</td>
</tr>
<tr>
<td>5’ mddA pTEV16</td>
<td>NNGCTCTTCTTTATCATGCAGTTTGACTGGAGCGCCATCTCT</td>
</tr>
<tr>
<td>3’ mddA pTEV16</td>
<td>NNGCTCTTCTTTATCATGCAGTTTGACTGGAGCGCCATCTCT</td>
</tr>
<tr>
<td><strong>Mutagenesis</strong></td>
<td></td>
</tr>
<tr>
<td>5’ mddA G244C</td>
<td>GTTTTCGCTATACCGTCCAGCAGTTTGATGTCGTTTTC</td>
</tr>
<tr>
<td>3’ mddA G244C</td>
<td>GAACATAAACCGAGTTGACTTTGACGAAAGAAACAG</td>
</tr>
<tr>
<td><strong>Strain deletions</strong></td>
<td></td>
</tr>
<tr>
<td>5’ glnPQ DEL</td>
<td>TATCCCTTATCACCCACCACGCTACATGGCTAGGCTGCTGAGCTGCTCT</td>
</tr>
<tr>
<td>3’ glnPQ DEL</td>
<td>CGTCCCTGCGCGCTGGAGGGCGATATCTCCACATGATATAGTTAGCTGCTCT</td>
</tr>
<tr>
<td>5’ metNI DEL</td>
<td>CAGTCGCTCTCATATTCTACTACGATAATAAAAGATATGCTGCTAGGCGCTGCTGCTG</td>
</tr>
<tr>
<td>3’ metNI DEL</td>
<td>CGTCCCTGCGCGCTGGAGGGCGATATCTCCACATGATATAGTTAGCTGCTGCTG</td>
</tr>
<tr>
<td>5’ mddA DEL</td>
<td>TATTGTAACACATTCCCTGGGTTTCTTATGGGTAGGCTGCTGCTG</td>
</tr>
<tr>
<td>3’ mddA DEL</td>
<td>AAAAGATGAGCGTCGAGCTGCTG</td>
</tr>
</tbody>
</table>

Construction of gene deletions. Strains carrying in-frame deletions of each gene of interest were constructed following the protocol established for *E. coli* (33). Briefly, PAGE-purified primers with 36 bp of homology with the 5' and 3' ends of the gene of interest (IDT, Coralville, Iowa) were used to amplify the *cat* or *kan* cassette from template plasmid pKD3 or pKD4, respectively. PCR products were analyzed on a 1% agarose gel, purified using the Wizard® SV Gel and PCR Clean-Up System (Promega), per the manufacturer’s protocol, followed by subsequent construction of the deletion of the gene as described in the original protocol (33). PCR was used to confirm the insertion of the *cat* or *kan* amplicons compared to the wild type allele size. All mutations were re-constructed in the appropriate strain background using P22 phage-mediated transduction crosses. Briefly, P22 phage was propagated on a strain containing an allele of interest. P22 lysates were generated from these strains to transduce the original recipient strain to the appropriate antibiotic resistance, yielding a re-constructed strain with the allele of interest (Table 4.1).

Plasmid construction. All plasmids used in this work are listed in Table 4.1. The cloning method using unique BspQ1 restriction sites as published previously (34) was used to construct all plasmids in this study unless otherwise stated. DNA sequencing (Georgia Genomics Facility, UGA) was used to verify all plasmids constructed in this study. Each gene of interest was amplified from *S. enterica* sv. Typhimurium LT2 genomic DNA.

Genes of interest (*mddA*, *glnPQ*, *metNI*) were cloned into the L-(+)-arabinose inducible vector pBAD24 (35) engineered to contain BspQ1 sites (C. M. VanDrisse & J. C. Escalante-Semerena, unpublished data). The resulting plasmids (pMDD8, pGLN2, and pMETN1) were used in complementation studies.
Plasmid pTEV5 (36), which directs the synthesis of the protein with a cleavable N-terminal hexahistidine tag, was modified to contain BspQ1 restriction sites. The resulting plasmid was named pTEV16 (C. M. VanDrisse & J. C. Escalante-Semerena, unpublished data), and used for overexpression of SeMddA\textsuperscript{WT} (pMDD7).

Site-directed mutagenesis was performed using primers designed from PrimerX (available at http://www.bioinformatics.org/primerx/) to mutate the predicted catalytic residue of SeMddA\textsuperscript{WT}, glutamate (E82), to a non-catalytic glutamine residue (Q82), to construct a catalytic variant (SeMddA\textsuperscript{E82Q}) in both the complementation (pMDD11) and overexpression vectors (pMDD10).

**Growth behavior analyses.** Starter cultures were grown overnight at 37°C with shaking in nutrient broth containing the appropriate drug marker and used to inoculate fresh medium (1\% v/v) in a volume of 200 µl per well of a 96-well plate with appropriate antibiotics. Strains containing plasmids were induced with varying concentrations of L-(+)-arabinose, as described in figure legends. Additional chemicals such as glutamine, methionine, MSX, MSO, or PHO were added at concentrations indicated in figures and figure legends. Plates were incubated at 37°C with shaking for 20-48 h in a microplate reader (Bio-Tek Instruments). Growth curves 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 the standard deviation.

**Overproduction and purification of the SeMddA\textsuperscript{WT} and SeMddA\textsuperscript{E82Q} proteins.** Vectors encoding SeMddA\textsuperscript{WT} (pMDD7) and SeMddA\textsuperscript{E82Q} (pMDD10) 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 LB containing ampicillin (100 µg ml\textsuperscript{-1}). Cultures were grown at 37°C with shaking to an OD\textsubscript{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 for 15 min at 4°C. The collected cell paste was re-suspended in binding buffer A [HEPES buffer (50 mM, pH 7.5) 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 sec, 50% duty) for 2 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 followed by filtration of the supernatant through a 0.45 µm filter (Millipore). Samples were loaded onto a 2 ml HisPur™ Ni-NTA resin column (Thermo Scientific) at 4°C, pre-equilibrated with binding buffer. His₆-tagged SeMddA proteins eluted with buffer B (HEPES buffer (50 mM, pH 7.5) containing NaCl (500 mM) and imidazole (500 mM)] following a wash step in the same buffer system with a lower amount of imidazole (40 mM).

To cleave the tag, His₆-TEV protease (rTEV) was purified as described (37), and cleavage of tagged SeMddA proteins was performed as follows: rTEV was added to the eluted protein supplemented with DTT (1 mM) at a 1:100 protease:tagged mg protein ratio, and the mixture was incubated at room temperature for 3 h. Proteins were dialyzed at 4°C in buffer C [HEPES buffer (50 mM, pH 7.5) containing NaCl (500 mM), tris(2-carboxyethyl)phosphine hydrochloride (TCEP, 0.5 mM], and ethylenediaminetetraacetic acid (EDTA, 0.5 mM). Dialyzed, cleaved protein was reloaded onto the column and were eluted using a 40 mM imidazole wash step, followed by an imidazole wash (500 mM, which allowed separation of the untagged (SeMddA<sup>WT</sup> or SeMddA<sup>E82Q</sup>) from tagged (His₆-TEV protease) protein. SeMddA<sup>WT</sup> and SeMddA<sup>E82Q</sup> eluted from the column during the wash step. The proteins were 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. Both proteins were purified to 99% homogeneity as determined using Total Lab v2005 software.

**Analytical gel filtration.** Experiments were performed at 4°C. Per run, a sample volume of 500 µl of 100 µg of SeMddA<sup>WT</sup> or SeMddA<sup>E82Q</sup> protein was injected onto a Superdex 200 HR 10/30 gel filtration column (GE Healthcare) attached to an ÄKTA purifier FPLC system that was equilibrated with buffer D [HEPES (50 mM, pH 7.4) and NaCl (100 mM)]. A calibration standard containing a mixture of molecular masses ranging from 1.35 to 670 kDa (Bio-Rad) was used to generate a standard curve to determine the molecular mass. The standards mixtures contained vitamin B<sub>12</sub> (1.35 kDa), equine myoglobin (17 kDa), chicken ovalbumin (44 kDa), bovine γ-globulin (158 kDa), and thyroglobulin (670 kDa). A flow rate of 0.5 ml min<sup>-1</sup> was used to develop the column and elution peak analysis was performed using UNICORN v4.11 software. Data were graphed and analyzed using Prism v6 (GraphPad) analytical software. Linear regression analyses of the standard curves yielded <i>r</i><sup>2</sup> values of 0.98.

**Thin layer chromatography.** Reaction mixtures included HEPES buffer (50 mM, pH 7.0), containing TCEP (1 mM), [1-<sup>14</sup>C]-Ac-CoA or [1-<sup>14</sup>C]-Pro-CoA (20 µM), substrate (0.5 mM), and SeMddA<sup>WT</sup> or SeMddA<sup>E82Q</sup> (1 µg). Reactions were incubated at 37°C for 30 minutes, spotted onto a polyester backed silica gel plate (Whatman Ltd), and developed in a chamber pre-equilibrated with a mobile phase of <i>n</i>-butanol, acetic acid, and dH<sub>2</sub>O (3:1:1). TLC plates were incubated for 3-4 h, dried, and developed with a phosphor screen overnight. The resulting phosphor-image was detected using a Typhoon Trio+ Variable Mode Imager (GE Healthcare) with ImageQuant v5.2 software.
Spectrophotometric enzyme activity assay. Specific activities of SeMddA<sup>WT</sup> and SeMddA<sup>E82Q</sup> were measured using a continuous spectrophotometric assay that employed 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent) as a reporter of free sulfhydryl groups (38, 39). Reaction mixtures contained HEPES buffer (50 mM, pH 7.2) containing Ac-CoA or Pro-CoA (100 µM), amine substrate (150 µM), DTNB (0.3 mM) and SeMddA<sup>WT</sup> or SeMddA<sup>E82Q</sup> protein (500 ng). Reactions were incubated for 5 min at 30°C in a Spectramax 384 Plus (Molecular Devices). Absorbance was measured continuously at 412 nm using SoftMax Pro v6.2 software. Data graphing and analysis was performed using Prism v6 software. Activity assays were performed in technical duplicate in three independent experiments.

Synthesis and identification of acetyl-methionine sulfoximine and acetyl-methionine sulfone. To confirm the location of acetylation of MSX and MSO, acetyl-MSX and acetyl-MSO were generated enzymatically. Reaction components included acetyl-CoA (1 mM), MSX or MSO (0.5 mM), SeMddA<sup>WT</sup> (20 µg), and ammonium bicarbonate (20 mM) in a reaction volume of 500 µl. A no-enzyme control reaction was also performed. Reactions were incubated at 37°C for 2 h. SeMddA<sup>WT</sup> was removed from the reaction mixture by filtration using Amicon® Ultra centrifugal filters (Millipore) with a 3K molecular weight cut-off, according to manufacturer’s protocol. Samples were concentrated in a Vacufuge plus speed vacuum (Eppendorf) at 30°C, and resuspended in 50% acetonitrile, 50% dH<sub>2</sub>O with 1% formic acid. The identity of acetyl-MSX and acetyl-MSO was confirmed by mass spectrometry (Protein and Mass Spectrometry Facility, UGA). ESI-MS was performed after sample dilution in acetonitrile and run on an Esquire 3000 Plus (Bruker) Ion Trap Mass Spectrometer at 0.3 ml/h.

Isolation of a Tn10<sup>d(tet<sup>+</sup>)</sup> insertion in an mddA::cat<sup>+</sup> strain by transposon mutagenesis. To identify mutations that allow for growth of an mddA::cat<sup>+</sup> strain in the presence of MSX, a
mutagenesis screen using Tn10d(tet⁺) transposons was utilized. To obtain a phage pool lacking the mddA gene, a transposition experiment was performed by transducing a P22 phage stock carrying a pool of Tn10d(tet⁺) transposons inserted throughout the S. enterica genome into an mddA::cat⁺ strain containing a plasmid carrying a transposase (JE18543). This was done to prevent recovery of growth due to repair of the mddA::cat⁺ deletion. Transduction reactions were plated on NB plates containing tetracycline (20 µg ml⁻¹) to select for transposon insertion. An estimated total of 61,308 colonies were pooled, resulting in ~13.5x coverage of the S. enterica genome.

A P22 lysate grown on this pool of strains was used to transduce strain mddA::cat⁺ (JE18333) to tetracycline resistance (20 µg ml⁻¹) on NB plates, followed by replica printing onto NCE minimal medium plates containing glycerol (22 mM) and MSX (10 µM). Colonies arising on the selection plates were freed of phage, patched onto NB plates, incubated for 4-6 h at 37°C, and replica printed to several selection plates including NB with tetracycline (20 µg ml⁻¹), NB with chloramphenicol (20 µg ml⁻¹), and NCE minimal medium with MSX (10 µM). Mutants that grew in all of these conditions were freed of phage and P22 lysates were generated. The resulting phage lysate was used as a donor in crosses with parental mddA::cat⁺ strain (JE18333). Transductions were plated on NB with tetracycline, selecting for the transposon, and after 24 h of growth at 37°C replica printed to plates containing MSX (10 µM) to confirm the phenotype.

The location of the insertion on the chromosome was determined in the re-constructed strain by sequencing the DNA flanking the Tn10d(tet⁺) element using a PCR-based protocol. A DNA product was amplified with degenerate primers and primers derived from the Tn10d(tet⁺) insertion sequences as reported previously (40) and used as template for sequencing reactions.
DNA sequencing was performed using BigDye® Terminator v3.1 protocols (Applied Biosystems) and the reactions were analyzed at the University of Georgia Genomics Facility.

RESULTS

Methionine sulfoximine (MSX) and methionine sulfone (MSO) inhibit growth of an 
mddA::cat$^+$ strain. We examined the ability of an S. enterica mddA::cat$^+$ strain (JE18333) to grow in the presence of MSX or MSO (Fig. 4.2). In the absence of MSX or MSO, no growth differences were observed between the mddA$^+$ and mddA::cat$^+$ strains in either rich or minimal media (data not shown). Addition of MSX (10 µM) caused complete growth inhibition of the mddA::cat$^+$ strain in minimal medium (Fig. 4.2A, solid diamonds). The same concentration of MSX did not have any effect on the growth of the mddA$^+$ strain (Fig. 4.2A, inverted solid triangles). The observed effects were different when strains were grown in rich medium. Under these conditions MSX partially inhibited growth of the mddA::cat$^+$ strain, and only when the MSX concentration was at least 50 µM (Fig. 4.2B, solid triangles).

Addition of methionine sulfone (MSO, 50 µM) also negatively affected growth of the mddA::cat$^+$ strain relative to that of the mddA$^+$ in minimal medium, but was less severe than the effect caused by MSX (Fig. 4.2C, solid triangles). A short delay in the onset of exponential growth was observed for the mddA::cat$^+$ strain in rich medium containing 200 µM MSO (Fig. 4.2D, solid triangles).

SeMddA activity blocks the negative effects of MSX and MSO. We performed in vivo experiments to determine whether or not SeMddA played a role in circumventing the toxic effects of MSX and MSO. For this purpose, plasmid pMDD8 (mddA$^+$) was introduced into the mddA::cat$^+$ strain (JE18333), yielding strain JE18961. As a control, an inactive variant of
Figure 4.2. MSX and MSO inhibit growth of an mddA::cat\(^{+}\) strain. Growth of S. enterica mddA\(^{+}\) and mddA::cat\(^{+}\) strains was examined in the presence of MSX (10 \(\mu\)M or 50 \(\mu\)M) or MSO (50 \(\mu\)M or 100 \(\mu\)M) as indicated. Non-essential E (NCE) minimal medium with glycerol (22 mM) or nutrient broth were used in these experiments. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. The following strains were analyzed: mddA\(^{+}\) (JE10079), mddA::cat\(^{+}\) (JE18333), mddA::cat\(^{+}\)/pMDD8 mddA\(^{+}\) (JE18961), and mddA::cat\(^{+}\)/pMDD11 mddA\(^{+}\) (encoding SeMddA\(^{E82Q}\)) (JE19029). Error bars represent standard deviation.
SeMddA was constructed by site-directed mutagenesis, in which the predicted catalytic residue E82 was changed to Q82 (SeMddA^{E82Q}). The presence of plasmid pMDD11 encoding the SeMddA^{E82Q} variant did not restore growth of the mddA::cat\textsuperscript{+} strain in the presence of MSX (Fig. 4.2A, solid circles). However, the wild-type mddA\textsuperscript{+} allele \textit{in trans} (pMDD8) supported growth of the mddA::cat\textsuperscript{+} strain in the presence MSX or MSO (Fig. 4.2A, solid squares; Fig. 4.2D, inverted solid triangles). It should be noted that at higher levels of inducer, \textit{i.e.} ≥250 µM (L-(+)-arabinose), an mddA::cat\textsuperscript{+} strain synthesizing the SeMddA^{E82Q} variant grew in medium containing MSX (10 µM), albeit with an extended lag phase, indicating that SeMddA^{E82Q} retained some catalytic activity (Fig. 4.3).

\textbf{High levels of MSX are inhibitory to wild-type S. enterica.} An increase in the lag phase of a culture of the mddA\textsuperscript{+} strain (JE10079) was seen at higher levels of MSX (20-50 µM). However, the strain grew at the same rate observed in medium devoid of MSX, and reached the same optical density of a culture unexposed to MSX (Fig. 4.4A). The effect of MSX in rich medium was notably different as the mddA\textsuperscript{+} strain showed no inhibition when the MSX concentration was at 500 µM (data not shown), and only partial inhibition of growth occurred at 1 mM (Fig. 4.4B, solid triangles). In sharp contrast, as stated previously, no growth of the mddA::cat\textsuperscript{+} strain was observed in the presence of as low as 10 µM MSX in minimal medium (Fig 4.4C), however, in rich medium, growth of the mddA::cat\textsuperscript{+} strain was only abolished at a concentration of 1 mM MSX (Fig 4.4D, inverted solid triangles). High concentrations (\textit{e.g.} 200 µM) of other methionine derivatives such as methionine sulfoxide (MS) and buthionine sulfoximine (BSX) did not affect the growth rate of either the mddA\textsuperscript{+} or mddA::cat\textsuperscript{+} strains in minimal medium (doubling time: + MS 1.4, 1.5, respectively; + BSX 1.6, 1.7, respectively).
Figure 4.3. High levels of MddA$^{E82Q}$ variant allow an mddA1::cat$^{+}$ strain to grow in the presence of MSX. Growth of the S. enterica strain mddA1::cat$^{+}$ / pMDD11 (encodes MddA$^{E82Q}$) in NCE minimal medium (glycerol, 22 mM) was examined in the presence of MSX (10 µM) and increasing concentrations of inducer (L-(+)-arabinose; 250 µM, 500 µM, and 1000 µM), as indicated. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. The strain analyzed was JE19029 (mddA1::cat$^{+}$ / pMDD11 mddA3$^{+}$, which encodes MddA$^{E82Q}$). Error bars represent standard deviation.
Figure 4.4. The effect of increasing MSX concentration on *mddA*<sup>+</sup> and *mddA::cat*<sup>+</sup> strains is detrimental to growth. Growth of the *S. enterica* strains *mddA*<sup>+</sup> and *mddA1::cat*<sup>+</sup> was examined with increasing concentrations of MSX, as indicated. The minimal medium used was NCE minimal medium with glycerol (22 mM) and the rich medium used was nutrient broth. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. The following strains were analyzed: *mddA*<sup>+</sup> (JE10079) and *mddA1::cat*<sup>+</sup> (JE18333). Error bars represent standard deviation.
SeMddA does not block the inhibitory effects of phosphinothricin (PHO). Addition of PHO (100 µM) was inhibitory to S. enterica growth, however, both the mddA and mddA::cat\(^+\) strains were equally affected, indicating that the deleterious effect of PHO could not be blocked by SeMddA (Fig. 4.5). At present, it is unclear whether PHO is acetylated in S. enterica, and if so, which acetyltransferase catalyzes the reaction.

The inhibition of a S. enterica mddA::cat\(^+\) strain by MSX or MSO is alleviated by the addition of glutamine or methionine. Glutamate analogues, such as MSX and MSO, inhibit glutamine synthetase (GlnA), which catalyzes the ATP-dependent condensation of glutamate with ammonia to produce glutamine. GlnA function is essential in nitrogen metabolism and MSX has been shown to bind tightly to the enzyme, causing irreversible inhibition of activity (41, 42). This inhibition is partially resolved by the addition of the product, glutamine, bypassing the requirement for GlnA (43). When glutamine (200 µM) was present in the culture medium, an mddA::cat\(^+\) strain exposed to MSX (10 µM) resulted in a modest increase in growth yield, plateauing at OD\(_{630}\) ~0.2 (Fig 4.6A). The addition of higher concentrations of glutamine (i.e. 500 and 1000 µM) correlated with higher growth yields (Fig. 4.6A), suggesting that saturation of the GlnA active site with glutamine outcompeted binding of the inhibitor.

Growth of an mddA::cat\(^+\) strain exposed to MSO (50 µM) was restored to wild type levels with higher levels of glutamine (i.e. 500 and 1000 µM) in the medium (Fig. 4.6C). Addition of glutamine to a mddA::cat\(^+\) strain exposed to either MSX or MSO in rich medium allowed the strain to grow at a rate similar to that of the mddA\(^+\) strain, and reached a similar growth yield compared to wild type levels (Fig. 4.7). It is clear that MSX and MSO inhibit GlnA activity (41, 44-48), however, it is less clear whether these compounds are inhibiting other cellular processes that could account for the observed growth phenotypes.
Figure 4.5. SeMddA does not prevent growth inhibition by PHO. Growth of the *S. enterica* strains *mddA*+ and *mddA::cat*+ was examined with or without PHO (100 µM). NCE minimal medium supplemented with glycerol (22 mM) was used in these experiments. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. The following strains were analyzed: *mddA*+ (JE10079) and *mddA::cat*+ (JE18333). Error bars represent standard deviation.
Figure 4.6. Glutamine and methionine counteract the deleterious effects of MSX and MSO on growth in the absence of MddA. Growth of *S. enterica* mddA<sup>+</sup> and mddA::cat<sup>+</sup> strains was examined in the presence of MSX (10 µM) and MSO (50 µM), with the addition of glutamine (200 µM, black triangles and squares; 500 µM, gray triangles; 1000 µM, white triangles) or methionine (500 µM, white squares and triangles) to the medium. NCE minimal medium supplemented with glycerol (22 mM) was used in these experiments. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under *Materials and Methods*. The following strains were analyzed: mddA<sup>+</sup> (JE10079) and mddA::cat<sup>+</sup> (JE18333). Error bars represent standard deviation.
Figure 4.7. Addition of glutamine fully restores growth of an mddA strain exposed to MSX and MSO in rich medium. Growth of the S. enterica strains mddA+ and mddA1::cat+ was examined with (A) MSX (50 µM) and (B) MSO (100 µM), with the addition of glutamine (200 µM). The medium used was nutrient broth. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. The following strains were analyzed: mddA+ (JE10079) and mddA1::cat+ (JE18333). Error bars represent standard deviation.
Previous work showed that the addition of methionine to wild type *S. enterica* exposed to MSX restored growth (for wild type cultures within 8 h), possibly alleviating toxicity by competing with the uptake of MSX (43). The growth behavior of *mddA*<sup>+</sup> and *mddA::cat*<sup>+</sup> strains in medium containing methionine (500 µM) and MSX (10 µM), or MSO (50 µM), was investigated (Fig. 4.6B, 4.6D). Addition of methionine improved the growth of a *mddA::cat*<sup>+</sup> strain exposed to MSX, but not to wild type levels. Methionine also decreased MSX toxicity in the *mddA*<sup>+</sup> strain, resulting in cultures reaching stationary phase ~10 h sooner than cultures growing in the presence of MSX and the absence of methionine (Fig. 4.6B). When MSO was used in lieu of MSX, growth of *mddA::cat*<sup>+</sup> was restored to wild type levels when methionine was added to the culture medium (Fig. 4.6D).

**SeMddA is a dimer in solution.** *SeMddA<sub>WT</sub>* is a 516-residue, 19.2-kDa protein. *SeMddA<sub>WT</sub>* and the *SeMddA<sub>E82Q</sub>* catalytic variant were isolated to 99% homogeneity using Ni-affinity chromatography (see Materials and Methods) (Fig. 4.8A). To determine the oligomeric state of the proteins in solution, FPLC gel filtration analysis was performed using commercially available molecular mass standards. Under the conditions tested, *SeMddA<sub>WT</sub>* and *SeMddA<sub>E82Q</sub>* eluted ~30 min after injection. The retention time was consistent for a protein whose mass was approximately 40-kDa, when compared to the elution times of molecular mass standards. Since the calculated molecular mass *SeMddA<sub>WT</sub>* was approximately 19-kDa, we inferred that the *SeMddA<sub>WT</sub>* was a dimer in solution (Fig. 4.8B). The oligomeric state of *SeMddA* was consistent with MddA homologues from *P. aeruginosa* (PA4866) and *A. baylyi* (ACIAD1637) (12, 13).

**SeMddA<sub>WT</sub> acetylates toxic methionine derivatives.** Thin layer chromatography (TLC) was used to identify the substrate(s) of *SeMddA<sub>WT</sub>*. An *in vitro* activity assay was used to monitor the *SeMddA*-dependent transfer the <sup>14</sup>C-labeled acetyl moiety from [<sup>14</sup>C-2]-acetyl-CoA to putative
Figure 4.8. SeMddA is a dimer in solution. (A) SeMddA WT and the catalytic variant SeMddA E82Q (~19-kDa) were purified using a two-step nickel affinity purification. An SDS-PAGE gel shows the molecular mass standards (kDa) (lane 1), purified SeMddA WT protein (lane 2), and purified SeMddA E82Q (lane 3). Both proteins purified to >99% homogeneity. (B) The molecular mass of SeMddA WT (black circle) and SeMddA E82Q (gray circle) in solution was estimated by gel filtration chromatography as described under Materials and Methods. Molecular mass standards (open circles) are thyroglobulin (bovine; 670-kDa), γ-globulin (bovine; 158-kDa), ovalbumin (chicken; 44-kDa), myoglobin (horse; 17-kDa) and vitamin B₃₂ (1.35-kDa).
substrates that were analogous in structure to MSX (see Materials and Methods). Phosphor imaging showed SeMddA<sup>WT</sup>-dependent acetylation of MSX, MSO, methionine sulfoxide (MS), buthionine sulfoximine (BSX), methionine, and glutamine (Fig. 4.9). Transfer of the radiolabel was not seen when PHO, glutamate, or arginine were used as substrates (Fig. 4.9). To examine the specificity of the SeMddA<sup>WT</sup> protein to acetylate the substrates identified by TLC, a spectrophotometric assay using 5,5'-dithiobis-2-nitrobenzoic acid (DTNB, Ellman’s reagent), was utilized to measure the specific activity of SeMddA<sup>WT</sup> for each substrate (see Materials and methods) (Table 4.3). SeMddA<sup>WT</sup> had increased activity for MSX (24 ± 0.5) and MSO (23 ± 0.1) µmol CoA min<sup>-1</sup> mg<sup>-1</sup>. The other compounds tested (MS, BSX, methionine, and glutamine) were poor substrates for the enzyme compared to MSX, under the conditions tested, with specific activities of ~0.5-2 µmol CoA min<sup>-1</sup> mg<sup>-1</sup> (data not shown).

Activity of the SeMddA<sup>E82Q</sup> catalytic variant (negative control) was reduced by >20-fold (1 ± 0.5) compared to SeMddA<sup>WT</sup> (Table 4.3). Although the activity of SeMddA<sup>E82Q</sup> was substantially decreased in vitro, the residual activity observed may account for the growth of an mddA<sup>cat</sup>+ strain containing a plasmid encoding SeMddA<sup>E82Q</sup> in medium with MSX (10 µM) at high levels of induction (250-1000 µM L- (+)-arabinose) (Fig. 4.3).

GNATs can transfer not only acetyl groups but also longer acyl groups (e.g. propionyl, succinyl) to their substrates, and several reports demonstrate the physiological relevance of this modification (49, 50). The ability of SeMddA<sup>WT</sup> to propionylate its substrates was also examined. SeMddA<sup>WT</sup> propionylated MSX (14 ± 2.3) and MSO (15 ± 0.1) µmol min<sup>-1</sup> mg<sup>-1</sup>, but the specific activities were lower the one observed when acetyl-CoA was used as the acyl donor (Table 4.3). These data indicate that, although SeMddA<sup>WT</sup> has propionylation activity, acetyl-CoA is the preferred co-substrate under the conditions tested.
Figure 4.9. *SeMddA*<sup>WT</sup> acetylates methionine derivatives. The substrate specificity of *SeMddA*<sup>WT</sup> was examined using thin layer chromatography. Reactions included 1 µg of *SeMddA*<sup>WT</sup> or *SeMddA*<sup>E82Q</sup> (neg. ctrl), [1<sup>14</sup>C]-acetyl-CoA, and 0.5 mM of substrate. After exposure of the TLC plate to a phosphor screen the resulting image was detected using a Typhoon Trio+ Variable Mode Imager (GE Healthcare).
Table 4.3. Kinetic Parameters\(^a\) for *S. enterica* MddA\(^{WT}\).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>(K_{M(app)}) ((\mu)M)</th>
<th>(k_{cat(app)}) (s(^{-1}))</th>
<th>(k_{cat}/K_M) (M(^{-1}) s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSX</td>
<td>576.3 ± 89</td>
<td>34.5 ± 2.0</td>
<td>6.0 (\times) 10(^4)</td>
</tr>
<tr>
<td>MSO</td>
<td>229.5 ± 56</td>
<td>30.6 ± 1.7</td>
<td>1.3 (\times) 10(^5)</td>
</tr>
<tr>
<td>Acetyl-CoA</td>
<td>155.7 ± 41</td>
<td>23.4 ± 1.4</td>
<td>1.5 (\times) 10(^5)</td>
</tr>
</tbody>
</table>

\(^a\) Values represent average mean ± standard deviation

Table 4.3. *SeMddA\(^{WT}\)* acetylates MSX and MSO. Specific activity of the *SeMddA\(^{WT}\)* protein for MSX and MSO was measured using a spectrophotometric assay described under *Materials and Methods*. Both acetyl-CoA and propionyl-CoA were tested as acyl donors. The catalytic variant *SeMddA\(^{E82Q}\)* was used as negative control. Experiments were performed in technical duplicates in three independent experiments. Values represent specific activity (\(\mu\)mol of CoA produced per min per mg of protein); ± standard deviation for each substrate.
Location of acetylation of MSX and MSO. To confirm the location of acetylation of MSX and MSO, acetyl-MSX and acetyl-MSO were generated enzymatically (see Materials and Methods), and the structures resolved by mass spectrometry (Protein and Mass Spectrometry Facility, UGA). Signals for the predicted masses of acetylated MSX (221 m/z) and acetylated MSO (222 m/z) were observed using ESI MS (data not shown). LC/MS/MS of the acetyl-MSX signal resulted in a strong peak at 142 m/z (data not shown), indicative of the acetyl group being located at the carboxy end of the molecule.

S. enterica cannot use MSX or acetyl-MSX as a source of methionine. To try to understand the fate of Ac-MSX, the ability of this compound to be utilized by S. enterica as a methionine source was tested. A methionine auxotroph (JE6583, metE mddA\(^+\)) was grown in minimal medium that lacked methionine (negative control), contained methionine (positive control; 100 µM), MSX (100 µM), or methionine and MSX (100 µM each). Growth was only observed when methionine was present (Fig. 4.10). These results indicated that neither MSX nor its acetylated form, as the strain was mddA\(^+\), could be utilized by S. enterica to generate methionine, or at least not to levels high enough to restore growth of the methionine auxotroph.

An mddA\(^+\) strain displays biphasic growth at higher MSX concentrations with the addition of glutamine. Growth of the mddA\(^+\) strain (JE10079) exposed to MSX with the addition of glutamine (200 µM) was examined at varying concentrations of MSX (50-200 µM). Under these conditions, cultures of the mddA\(^+\) strain exhibited biphasic growth, with the initial growth onset occurring at the same time, followed by a plateau with an increasing lag phase. The observed lag in growth correlated with higher concentrations of MSX when greater than 5 µM MSX (Fig. 4.11A). Interestingly, after the onset of growth the growth rate for all conditions was similar [doubling time: 1.4 (5 µM), 2.1 (50 µM), 2.4 (100 µM), 3.2 (200 µM)].
Figure 4.10. Neither MSX nor acetyl-MSX permits growth of a methionine auxotroph in S. enterica. Growth of an S. enterica strain auxotrophic for methionine (metE mddA\textsuperscript{+}) was tested in NCE minimal medium (glycerol, 22 mM) in the absence of methionine, methionine (100 µM), MSX (100 µM), or methionine and MSX (100 µM each). Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. The strain analyzed was metE205 mddA\textsuperscript{+} (JE6583). Error bars represent standard deviation.
Figure 4.11. An mddA\(^+\) strain exhibits biphasic growth in minimal medium containing glutamine and MSX. (A) Growth of the *S. enterica* mddA\(^+\) strain in NCE minimal medium supplemented with glycerol (22 mM) was examined in the presence of glutamine (200 µM) and increasing concentrations of MSX (5-200 µM), as indicated. The strain analyzed was mddA\(^+\) (JE10079). (B) Growth of the *S. enterica* mddA\(^+\)/pMDD8 mddA\(^+\) strain was examined in the same medium in the presence of MSX (50 µM) and increasing concentrations of inducer (L-(+)-arabinose, 50-500 µM), as indicated. The strain analyzed was mddA\(^+\)/pMDD8 mddA\(^+\) (JE18955). Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under *Materials and Methods*. Error bars represent standard deviation.
**Ectopic overexpression of mddA** provides resistance to higher concentrations of MSX. A mddA strain carrying mddA on an plasmid under the control of an inducible promoter (JE18961) was grown in minimal medium with MSX (50 µM) and varying levels of inducer (50-500 µM L-(+)-arabinose). At higher concentrations of inducer, cultures of strain JE18961 grew with shorter lag times when compared to the non-induced control (Fig. 4.11B). These data were consistent with the idea that increased SeMddA protein levels provided greater protection against the toxic effects of MSX.

**Identification of genetic loci whose functions are required for MSX toxicity.** We took a genetic approach to find loss-of-function derivatives of a mddA::cat (JE18333) strain that grew in the presence of MSX (10 µM). For this purpose, a P22 phage lysate grown on a pool of S. enterica mddA::cat strains carrying Tn10d(tet) elements randomly inserted in the chromosome was used as donor to transduce the strain mddA::cat to tetracycline resistance (for details see Materials and Methods). Tetracycline-resistant (TetR) colonies were replica printed onto minimal medium supplemented with glycerol (22 mM), tetracycline (20 µg ml⁻¹), and MSX (10 µM). Strains that grew this condition were analyzed further. To confirm that growth of the mddA::cat strain was due to the inheritance of a Tn10d(tet) element, P22 phage was grown on these strains to generate a phage lysate that was used as donor to transduce the original mddA::cat recipient strain to TetR. After freeing the reconstructed strain from phage, the location of the transposon insertion was identified by arbitrary PCR and subsequent DNA sequencing (40).

One of ~50,000 TetR colonies grew on medium containing MSX after re-construction. The transposon insertion in this strain was located within the glnP gene (insertion at bp 384 out of 660), which encodes the membrane component subunit of the glutamine ABC transporter GlnHPQ (Fig. 4.12A). The re-constructed mddA::cat glnP::Tn10d(tet) strain (JE20027) grew
Figure 4.12. Growth of the mddA::cat+ glnP::Tn10d strain in the presence of MSX. (A) Transposon mutagenesis (Tn10d(tet+)) of an mddA::cat+ strain identified that inactivation of the glnHPQ (glutamine permease) restored growth in the presence of MSX (10 µM). (B) Growth of the reconstructed S. enterica mddA1::cat+ glnP1561::Tn10d strain identified in the mutagenesis screen was examined in NCE minimal medium supplemented with glycerol (22 mM) and MSX (10 µM). Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. The following strains were analyzed: mddA+ glnP+ (JE10079), ∆mddA2 (JE18622), and mddA1::cat+ glnP1561::Tn10d (JE20027). Error bars represent standard deviation.
on minimal medium containing MSX (10 µM) (Fig. 4.12B, solid triangles). Identification of this mutant was not surprising, as the amino acid transporters MetNIQ (methionine permease) and GlnHPQ (glutamine permease) are responsible for transporting MSX in *S. enterica*. To our knowledge, transport of MSO has not been investigated (51).

**Deletion of two amino acid transport systems relieves MSX toxicity.** A deletion of *glnP* in combination with *glnQ* (ATP-binding subunit, strain JE20064) was constructed to ensure that the cell was devoid of transporter. Growth of the *mddA*<sup>+</sup> and Δ*mddA* strains in the absence of GlnPQ was examined in medium containing MSX (20 µM) (Fig. 4.13A). A *glnPQ*::*cat*<sup>+</sup> strain (JE20064) exposed to MSX reached stationary phase faster than the wild type strain (~10 h) (Fig. 4.13A, inverted solid triangles). A Δ*mddA* *glnPQ*::*cat*<sup>+</sup> strain (JE20065) displayed an increased lag phase compared to the wild-type strain, but once the culture started growing it did so at a rate similar to that of the *glnPQ mddA*<sup>+</sup> strain (Fig. 4.13A, open diamonds). Ectopic synthesis of *glnPQ* in the Δ*mddA* *glnPQ*::*cat*<sup>+</sup> strain (JE20073) restored the transport of MSX and abolished growth in medium containing MSX (Fig. 4.13A, open circles).

Transport of MSX has been examined in *S. enterica* in relation to the transport of amino acids (Met) (52, 53). These studies demonstrated that the uptake of MSX and methionine sulfoxide was inhibited when both glutamine permease (GlnHPQ) and methionine permease (MetNIQ) were blocked. As seen in figure 3.13B, growth of a Δ*mddA* *glnPQ*::*cat*<sup>+</sup> *metNI*::*kan*<sup>+</sup> strain (JE20067) lacking both transporters was examined in the presence MSX. Strain JE20067 grew better than the wild-type strain, with no observed lag when MSX was added to the medium (Fig. 4.13B, open squares). The Δ*mddA* *metNI*::*kan*<sup>+</sup> strain behaved similarly to Δ*mddA* *glnPQ*::*cat*<sup>+</sup> strain (Fig. 4.13B, open diamonds).
Figure 4.13. Deletion of two amino acid transporters (GlnHPQ, MetNIQ) restores growth of a ΔmddA strain exposed to MSX and MSO. (A, B, C, D) Growth of *S. enterica* mdd*+* and ΔmddA strains with or without the glutamine permease (*glnPQ*) or methionine permease (*metNI*) was examined in the presence of MSX (20 µM) or MSO (50 µM) in NCE minimal medium with glycerol (22 mM). Plasmids were induced with 10 µM L- (+)-arabinose. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. The following strains were analyzed: mdd*+* glnPQ*+* metNI*+* (JE10079), ΔmddA (JE18622), glnPQ::cat*+* (JE20064), ΔmddA glnPQ::cat*+* (JE20065), metNI::kan*+* (JE19583), ΔmddA metNI::kan*+* (JE19730), ΔmddA metNI::kan* glnPQ::cat*+* (JE20067), ΔmddA glnPQ::cat* / pGLN2 glnPQ*+* (JE20073), and ΔmddA metNI::kan* / pMETN1 metNI*+* (JE20329). Error bars represent standard deviation.
Ectopic synthesis of MetNI in the $\Delta mddA\ metNI::kan^+$ strain background also failed to support growth in the presence of MSX, confirming that GlnHPQ and MetNIQ transport MSX into the cell, and that deletion of these two systems supports growth of an $mddA$ strain on medium containing MSX (Fig. 4.13B, open circles). It is interesting to note that in either the $\Delta mddA\ glnPQ::cat^+$ or the $\Delta mddA\ metNI::cat^+$, growth occurred at a rate similar to the $mddA^+$ strain with only a longer lag phase.

**MSO is transported by GlnHPQ but not by MetNIQ.** Similar growth experiments using the above-mentioned strains were carried out with the addition of MSO (50 µM) to investigate the transport of MSO (Fig. 4.13C, 4.13D). A deletion of $glnPQ$ in a $\Delta mddA$ strain background (JE20065) restored growth to wild-type levels (Fig. 4.13C, open diamonds), indicating a decrease in the amount of MSO being transported into the cell. This effect was reversed when $glnPQ^+$ were expressed in trans (Fig. 4.13C, solid circles).

Surprisingly, the kinetics of growth of the $\Delta mddA^+\ metNI::kan^+$ strain (JE19730; Fig. 4.13D, solid diamonds) and $\Delta mddA$ strain (JE18622; Fig. 4.13D, open triangles) in the presence of MSO were very similar, demonstrating that deletion of the methionine transporter did not prevent sufficiently inhibitory levels of MSO from entering the cell. Expression of $metNI^+$ in trans in the $\Delta mddA^+\ metNI::kan^+$ strain (JE20329; Fig. 4.13D open circles) resulted in growth that was similar to that of the $\Delta mddA$ strain (Fig. 4.13D, open triangles). Taken together, these results imply that unlike MSX and methionine sulfoxide, MSO is not transported through the methionine permease.
DISCUSSION

**SeMddA is necessary for cell survival in the presence of MSX in minimal but not rich medium conditions.** On minimal medium, *S. enterica* strains lacking MddA cannot grow in the presence of MSX; however, in rich medium a higher concentration (5x) of MSX is needed to observe even a delay in growth (Fig. 4.2). This raises the question as to what enzymes or metabolites are present under rich medium conditions that help prevent MSX toxicity in the absence of MddA. One explanation could be differences in the availability of glutamine or methionine in rich versus minimal medium, since the addition of these compounds restore growth of a ΔmddA strain in the presence of MSX (Fig. 4.6). Other possibilities include increased expression of *glnA* (the gene encoding the enzyme known to be affected by MSX) or the presence of antioxidant compounds, in rich medium conditions.

**SeMddA does not acetylate phosphinothricin (PHO).** SeMddA and its homologues have been annotated as PHO acetyltransferases (like the Bar protein of *Streptomyces* spp.), an activity that MddA homologues in several organisms do not possess (12, 13). Although PHO is inhibitory to *S. enterica* when present at 100 µM, it is clear that the deleterious effect of PHO is affected by the absence or presence of MddA (Fig. 4.5). In contrast, our data support the conclusion that a physiological role of MddA in *S. enterica* is to block the harmful effects of oxidized methionine derivatives such as methionine sulfoximine (MSX) and methionine sulfone (MSO) (Fig. 4.2, 4.11, Table 4.3).

The fact that SeMddA can acetylate MSX and MSO, but not PHO, suggests a relative high degree of specificity of the enzyme for structural analogues to its *bona fide* substrates (Fig. 4.2, 4.5). There are, however, examples of homologues of the *Streptomyces* Bar protein that efficiently acetylate PHO and MSX. Such homologues are found in *Streptomyces hygroscopicus*...
and *Rhodococcus* spp. (19, 54). Whether PHO is acetylated by another *S. enterica* GNAT remains an open question.

*Expression of mddA is induced in response to MSX.* In the presence of glutamine with increasing concentrations of MSX, an *mddA* strain grows biphasically. Clearly, the addition of glutamine allows for the initial onset of growth, but at a certain point the cell either runs out of available glutamine, or GlnA is inhibited by MSX, leading to growth arrest (Fig. 4.11). Not surprisingly, the higher the concentration of MSX, the longer growth arrest lasts. We hypothesize that during growth arrest there may be an increase in the expression of *mddA* in response to the level of MSX. Once the latter is acetylated and rendered innocuous, the cell can resume growth, a conclusion that is supported by the fact that growth rates remain similar as soon as growth is restored.

*How does a strain devoid of MddA and either transporter grow in the presence of MSX?* On the basis of growth data presented (Fig. 4.12, 4.13), we conclude that a deletion of only one of the MSX transport systems still allows inhibitory concentrations of MSX into the cell, which raises the question of how the *mddA glnPQ* and *mddA metNI* strains can survive the toxic effects of MSX. A plausible explanation is the existence of an enzyme with *Se*MddA-like activity whose *k*\(_{\text{cat}}\) for MSX is slower than that of *Se*MddA, but sufficient to support growth under such conditions. Examples of redundant functions in cells are not rare [*e.g.* MetE, MetH (methionine synthases); PurN, PurT (glycinamide ribonucleotide transformylase), etc]). It should also be noted that mutations in GlnA have been characterized which prevent MSX binding and inhibition while retaining activity (55, 56). It is also interesting to note that the while both MSX and methionine sulfoxide are transported through the methionine permease, MetNIQ, our data
indicate that the structural analogue MSO is not (Fig. 4.13), shedding light on the specificity of the MetNIQ transporter.

**Why does SeMddA acetylate oxidized methionine derivatives?** There is precedent for the role of GNATs in detoxifying toxic compounds (5, 11), by using acetylation as a means to inactivate antibiotics. While our data clearly demonstrate that MddA of *S. enterica* can prevent the deleterious effects of MSX and MSO (Fig. 4.2, 4.11), it is less clear what environmental conditions expose *S. enterica* to MSX and MSO.

MSX has been identified in processed foods (1940’s) and naturally occurs in the roots and stems of some plant species (20). It is also possible that these compounds are produced as a consequence of the host response and reactive nitrogen or oxygen species (57), or synthesized endogenously as a by-product of a normal metabolic process (58), which would not be unprecedented for *S. enterica*. Oxidation of methionine residues of proteins can generate methionine sulfoxide, but no studies looking at protein oxidation have detected the production of MSO (59-61). Although the source of MSX and MSO are currently unknown, MddA function is necessary for *S. enterica* growth when these compounds are present. Future studies aim to understand the role of acetylation with respect to environmental stressors.

**ACKNOWLEDGEMENTS**

The authors have declared that no competing interests exist. This work was supported by USPHS grant R01 GM062203 to J.C.E.-S. We thank James N. Workman and Rachel Burckhardt for technical assistance.
REFERENCES


42. **Liaw SH, Eisenberg D.** 1994. Structural model for the reaction mechanism of glutamine synthetase, based on five crystal structures of enzyme-substrate complexes. *Biochemistry.* **33:**675-681.


CHAPTER 5

PROBING SUBSTRATE SPECIFICITY OF PHOSPHINOTHRICIN ACETYLTTRANSFERASE HOMOLOGUES IN *DEINOCOCCUS RADIODURANS* AND *GEOBACILLUS KAUSTOPHILUS*\(^4\)

---

ABSTRACT

Acetylation of small molecules is widespread and used as a mechanism to detoxify harmful chemicals. Many of the enzymes catalyzing acetylation reactions belong to the Gcn5 N-acetyltransferase (GNAT) family. *Streptomyces* species utilize a GNAT, named Bar, to acetylate and detoxify a self-produced toxin, phosphinothricin (PPT). Many Bar homologues, such as MddA from *Salmonella enterica*, were previously annotated incorrectly as PPT acetyltransferases, but instead have been shown acetylate the toxic structural homologues methionine sulfoximine (MSX) and methionine sulfone (MSO). These findings raise questions as to the substrate selectivity and function of putatively annotated PPT acetyltransferases. Here we provide evidence for the specificity of Bar homologues from *Deinococcus radiodurans* and *Geobacillus kaustophilus*, each of which contain two genes annotated to encode PPT acetyltransferases (*Dr*1057, *Dr*1182; and *Gk*0593, *Gk*2920). Previous work with *S. enterica* demonstrated that MddA was required for growth in medium containing MSX or MSO. Growth studies using an *S. enterica mddA1::cat* strain as a heterologous host revealed that ectopic expression of *D. radiodurans* and *G. kaustophilus* Bar homologues showed specificity in their ability to restore growth of an *mddA* strain in the presence of MSX and MSO. In wild-type *S. enterica*, synthesis of *Dr*1182, *Gk*0593, and *Gk*2920 blocked the inhibitory effect of PPT. Results of *in vitro* activity assays confirmed *in vivo* results, demonstrating selectivity among the enzymes annotated as PPT acetyltransferases. Here, we describe a genetic method to test substrate specificity of annotated PPT acetyltransferases.
INTRODUCTION

The Gcn5 N-acetyltransferase (GNAT, PF00583) superfamily of proteins is present in all domains of life and catalyze the transfer of the acetyl group of acetyl-CoA to proteins or small molecules (for review see (1)). These enzymes were first discovered to acetylate and inactivate aminoglycoside antibiotics (2-6) and protect against several cellular stressors (7-10), providing a precedent for the role of GNAT-mediated detoxification via acetylation. Previous work has shown a subset of GNAT enzymes annotated as phosphinothricin (PPT) acetyltransferases protect the cell against toxic amino acid derivatives, including PPT, methionine sulfoximine (MSX), and methionine sulfone (MSO) (7, 8, 11-14) (Fig. 5.1).

PPT is a component of a potent toxin known as Bialaphos, a tripeptide (phosphinothricyl-alanyl-alanine) produced by Streptomyces species. (15). The toxic effect occurs when PPT, a glutamate analogue, is cleaved from the tripeptide. PPT causes growth inhibition by irreversibly binding to the glutamine synthetase (GlnA) enzyme, responsible for the conversion of glutamate to glutamine using ATP and ammonium (16, 17). A GNAT of Streptomyces species, known as Bar (also Pat for phosphinothricin acetyltransferase), is responsible for the acetylation and detoxification of the self-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 have been classified as PPT acetyltransferases based on sequence similarity at the protein level to Streptomyces Bar. Some Bar homologues have been incorrectly annotated as PPT acetyltransferases, including Pseudomonas aeurginosa PITA (PA4866), Acinetobacter baylyi (AC1AD1637) and Salmonella enterica MddA (STM1590) (8, 12). Instead, these enzymes acetylate the toxic analogues MSX and MSO, which also inhibit GlnA function.
Figure 5.1. **Chemical structure of PPT analogues.** (A) Phosphinothricin (PPT); (B) Methionine sulfone (MSO); and (C) Methionine sulfoximine (MSX).
Here we present *in vivo* and *in vitro* data characterizing putative PPT acetyltransferases from *Geobacillus kaustophilus* and *Deinococcus radiodurans*. Using *S. enterica* as a heterologous host we assayed for function of the PPT acetyltransferases in an *S. enterica mddA* mutant or wild-type strain in conditions containing MSX, MSO, or PPT. Significantly, enzyme activity *in vitro* correlated with the observed growth phenotypes. Taken together, these data demonstrate varying specificity of these enzymes for MSX, MSO, and PPT. These data provide a potential model to assay for PPT acetyltransferase specificity.

**MATERIALS AND METHODS**

*Culture media and chemicals.* The minimal medium used in this study was no-carbon essential (NCE) minimal medium (19) containing MgSO$_4$ (1 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).

*Bacterial strains.* All strains studied were derivatives of *S. enterica* serovar Typhimurium strain LT2 (unless specified), and are listed in Table 5.1. All primers used in this study were synthesized by IDT (Coralville, Iowa) and are listed in Table 5.2.

*Plasmid construction for complementation and overexpression.* All plasmids used in this work are listed in Table 5.1. The cloning method using unique BspQI restriction sites as
Table 5.1. Strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Reference/source</th>
</tr>
</thead>
<tbody>
<tr>
<td>JE10079</td>
<td>ara-9 mddA⁺</td>
<td>Laboratory strain</td>
</tr>
<tr>
<td><strong>Derivatives of JE10079</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JE18333</td>
<td>mddA::cat⁺</td>
<td>(7)</td>
</tr>
<tr>
<td>JE18961</td>
<td>mddA::cat⁺ / pMDD8</td>
<td>(7)</td>
</tr>
<tr>
<td>JE20780</td>
<td>pDR1182-2</td>
<td></td>
</tr>
<tr>
<td>JE20781</td>
<td>mddA::cat⁺ / pDR1182-2</td>
<td></td>
</tr>
<tr>
<td>JE20782</td>
<td>pGK0593-2</td>
<td></td>
</tr>
<tr>
<td>JE20783</td>
<td>mddA::cat⁺ / pGK0593-2</td>
<td></td>
</tr>
<tr>
<td>JE20857</td>
<td>pDR1057-2</td>
<td></td>
</tr>
<tr>
<td>JE20858</td>
<td>mddA::cat⁺ / pDR1057-2</td>
<td></td>
</tr>
<tr>
<td>JE20864</td>
<td>mddA::cat⁺ / pVOC</td>
<td></td>
</tr>
<tr>
<td>JE20865</td>
<td>mddA::cat⁺ / pGK2920-2</td>
<td></td>
</tr>
<tr>
<td>JE20866</td>
<td>pGK2920-2</td>
<td></td>
</tr>
<tr>
<td>JE20973</td>
<td>pVOC</td>
<td></td>
</tr>
<tr>
<td><strong>E. coli strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli C41(lDE3)</td>
<td>ompT hsdS (rB mB) gal λ (DE3) including at least one non-characterized mutation</td>
<td>(22, 23)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMDD7</td>
<td>mddA⁺ cloned into pKLD66</td>
<td>(7)</td>
</tr>
<tr>
<td>pMDD8</td>
<td>mddA⁺ cloned into pBAD24</td>
<td>(7)</td>
</tr>
<tr>
<td>pDR1057-1</td>
<td>Dr1057⁺ cloned into pKLD66</td>
<td>(7)</td>
</tr>
<tr>
<td>pDR1057-2</td>
<td>Dr1057⁺ cloned into pBAD24</td>
<td>(7)</td>
</tr>
<tr>
<td>pDR1057-3</td>
<td>Dr1057⁺ cloned into pBAD24</td>
<td>(7)</td>
</tr>
<tr>
<td>pDR1057-4</td>
<td>Dr1057⁺ cloned into pKLD66</td>
<td>(7)</td>
</tr>
<tr>
<td>pDR1182-1</td>
<td>Dr1182⁺ cloned into pKLD66</td>
<td>(7)</td>
</tr>
<tr>
<td>pDR1182-2</td>
<td>Dr1182⁺ cloned into pBAD24</td>
<td>(7)</td>
</tr>
<tr>
<td>pGK0593-1</td>
<td>Gk0593⁺ cloned into pKLD66</td>
<td>(7)</td>
</tr>
<tr>
<td>pGK0593-2</td>
<td>Gk0593⁺ cloned into pBAD24</td>
<td>(7)</td>
</tr>
<tr>
<td>pGK2920-1</td>
<td>Gk2920⁺ cloned into pKLD66</td>
<td>(7)</td>
</tr>
<tr>
<td>pGK2920-2</td>
<td>Gk2920⁺ cloned into pBAD24</td>
<td>(7)</td>
</tr>
</tbody>
</table>

* All strains and plasmids were constructed during the course of this work, unless otherwise stated.
* pKLD66 is an overexpression vector described in (22) engineered with BspQ1 sites (C. M. VanDrisse and J. C. Escalante-Semerena, unpublished).
* pBAD24 is a cloning vector described in (23) engineered with BspQ1 sites (C. M. VanDrisse and J. C. Escalante-Semerena, unpublished).
<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cloning</strong></td>
<td></td>
</tr>
<tr>
<td>5' Dr1057 pBAD24</td>
<td>NNGCTCTTCNTTATGCCGAGGAGCAGCGGCAGCCCAT</td>
</tr>
<tr>
<td>5' Dr1057 pTEV16</td>
<td>NNGCTCTTCNAGCATGCCGAGAGCGGCAACCGGCCCAT</td>
</tr>
<tr>
<td>3' Dr1057 pVector</td>
<td>NNGCTCTTCNTTATCGTCCAGTACGCAGCGACC</td>
</tr>
<tr>
<td>5' Dr1182 pBAD24</td>
<td>NNGCTCTTCNNTTATGCCGAGGAGCAGCGGCAGCCCAT</td>
</tr>
<tr>
<td>5' Dr1182 pTEV16</td>
<td>NNGCTCTTCNAGCATGCCGAGAGCGGCAACCGGCCCAT</td>
</tr>
<tr>
<td>3' Dr1182 pVector</td>
<td>NNGCTCTTCNNTTATCGTCCAGTACGCAGCGACC</td>
</tr>
<tr>
<td>5' Gk0593 pBAD24</td>
<td>NNGCTCTTCNTTATGCCGAGGAGCAGCGGCAGCCCAT</td>
</tr>
<tr>
<td>5' Gk0593 pTEV16</td>
<td>NNGCTCTTCNNTTATGCCGAGGAGCAGCGGCAGCCCAT</td>
</tr>
<tr>
<td>3' Gk0593 pVector</td>
<td>NNGCTCTTCNNTTATCGTCCAGTACGCAGCGACC</td>
</tr>
<tr>
<td>3' Gk2920 pBAD24</td>
<td>NNGCTCTTCNNTTATGCCGAGGAGCAGCGGCAGCCCAT</td>
</tr>
<tr>
<td>5' Gk2920 pTEV17</td>
<td>NNGCTCTTCNNTTATGCCGAGGAGCAGCGGCAGCCCAT</td>
</tr>
<tr>
<td>3' Gk2920 pVector</td>
<td>NNGCTCTTCNNTTATGCCGAGGAGCAGCGGCAGCCCAT</td>
</tr>
<tr>
<td><strong>Mutagenesis</strong></td>
<td></td>
</tr>
<tr>
<td>5' Dr1057 N114E</td>
<td>CCCGACCGCTACGAACTCAGGTCAGGTCAG</td>
</tr>
<tr>
<td>3' Dr1057 N114E</td>
<td>GTGACCGTGACCTCGTAGGCGGTCGAGGTCAGG</td>
</tr>
</tbody>
</table>

Table 5.2. Primers used in this study.
published previously (21) was used to construct all plasmids in this study and DNA sequencing (Georgia Genomics Facility, UGA) was used to verify all plasmids. Genes of interest were amplified from *D. radiodurans* R1, or *G. kaustophilus* HTA426, genomic DNA. *G. kaustophilus* HTA426 genomic DNA was from laboratory collection strains and *D. radiodurans* R1 genomic DNA was gift from John Batista (Louisiana State University).

Plasmid pKLD66 (22), which directs the synthesis of the protein with a cleavable N-terminal hexahistidine tag, was engineered with BspQI restriction sites (pTEV16, C. M. VanDrisse & Escalante-Semerena, unpublished data) and used for overexpression, resulting in plasmids pDR1057-2, pDR1182-1, pGK0593-1, and pGK2920-1.

Each gene of interest was cloned into the *L*-(+)-arabinose inducible vector pBAD24 (23) engineered with BspQI sites (C. M. VanDrisse & J. C. Escalante-Semerena, unpublished data) for complementation studies resulting in plasmids pDR1057-2, pDR1182-2, pGK0593-2, and pGK2920-2.

Site-directed mutagenesis was performed using primers designed from PrimerX (available at http://www.bioinformatics.org/primerx/) to mutate the asparagine (N) 114 of *Dr*1057<sup>WT</sup> to a glutamate residue (Q114), to construct a catalytically active variant (*Dr*1057<sup>N114E</sup>) in the pBAD24 overexpression vector (pDR1057-3).

**Growth behavior analyses.** Starter cultures were grown overnight at 37°C with shaking in nutrient broth containing the appropriate antibiotic and used to inoculate fresh medium (1% v/v, 200 µl per well) of a 96-well plate with appropriate antibiotics. Strains containing plasmids were induced with varying concentrations of *L*-(+)-arabinose, as described in figures and figure legends. Additional chemicals such as phosphinothricin (PPT), methionine sulfoximine (MSX), and methionine sulfone (MSO) were added at concentrations indicated in figures and figure
legends. 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 the standard deviation.

**Protein overproduction and purification.** Plasmids encoding each gene of interest (Dr1182, Gk0593, or Gk2920) 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 LB containing ampicillin (100 µg ml⁻¹). Cultures were grown at 37°C with shaking to an OD₆₀₀ of 0.6, induced with IPTG (1 mM), and shaken overnight at ~28°C. Cells were harvested by centrifugation at 6,000 x g 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 sec, 50% duty) for 2 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 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 40 mM imidazole to remove nonspecifically bound proteins. Followed by that His₆-tagged proteins eluted in the same buffer system that contained 500 mM imidazole. 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.
**Mass spectrometry analysis.** Protein bands of interest resolved by SDS-PAGE were excised and submitted to the Protein and Mass Spectrometry (PAMS) facility at the University of Georgia. After a trypsin digest the proteins were analyzed by MALDI mass finger printing an LC-MS/MS to determine protein identity and sequence.

**Thin layer chromatography.** Reactions were performed as described (7). Reaction mixtures included HEPES buffer (50 mM, pH 7.0), containing TCEP (1 mM), [1-14C]-acetyl-CoA (20 µM), substrate (0.5 mM), and enzyme (Dr1057, Dr1182, Gk0593, or Gk2920; 1 µg). Reactions were incubated at 37°C for 30 min, spotted onto a polyester backed silica gel plate (Whatman Ltd), and developed in a chamber pre-equilibrated with a mobile phase of n-butanol, acetic acid, and dH2O (3:1:1). After incubation TLC plates were developed with a phosphor screen overnight, and the resulting phosphor-image was detected using a Typhoon Trio+ Variable Mode Imager (GE Healthcare) with ImageQuant v5.2 software.

**Activity assays.** Reactions were performed as described (7), using a SpectraMax Plus 384 microplate spectrophotometer (Molecular Devices) equipped with SoftMax Pro v4 software was used for data acquisition. Assays were performed at 30°C in 100-µl volumes in 96-well microplates using a continuous spectrophotometric assay that employed 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent) as a reporter of free sulfhydryl groups at 412 nm (24, 25). Reaction mixtures contained HEPES buffer (50 mM, pH 7.2), DTNB (0.3 mM), acetyl-CoA, protein of interest (Dr1057, Dr1182, Gk0593, or Gk2920), and substrate of interest (PPT, MSO, or MSX). Reactions were initiated by the addition of enzyme. A control containing enzyme but no acetyl-CoA was used to correct for the background. Data were acquired every 10 s over a 5 min time period. Data were collected using an average of a technical triplicate, in experimental triplicate. Reactions included HEPES (pH 7.5, 50 mM), protein of interest (0.5 µg),
acetyl-CoA (100 µM), and PPT, MSX, or MSO (150 µM), and DTNB (0.3 mM). Data was graphed using Prism v6 (GraphPad) analytical software. The molar extinction coefficient used for the concentration of the TNB$^{2-}$ anion was 14,150 M$^{-1}$ cm$^{-1}$.

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

**RESULTS**

*S. enterica* can be used as a heterologous host to screen activity of putative phosphinothricin (PPT) acetyltransferases in vivo. Previous work from our lab characterized the annotated PPT acetyltransferase, MddA (formerly YncA, STM1590) from *S. enterica* (7). We demonstrated that this enzyme acetylated the methionine derivatives methionine sulfoximine (MSX) and methionine sulfone (MSO), but not PPT. Growth of an *S. enterica mddA1::cat$^+$* strain is abolished when MSX is present (10 µM) and inhibited in the presence of MSO (50 µM), compared to the *mddA$^+$* strain (Fig. 5.2) (7). The *mddA1::cat$^+$* strain was used as a genetic tool to examine the ability of the *D. radiodurans* and *G. kaustophilus* putative PPT acetyltransferases to detoxify MSX or MSO in vivo when expressed in trans, by restoring growth to wild-type levels (Fig. 5.3). Control growth studies are shown in Figure 5.2. Expression of *Dr*1182 or *Gk*2920 in trans was able to restore growth of the *S. enterica mddA1::cat$^+$* strain to wild-type levels at low levels of induction (10 µM *L-(+)* arabinose) in conditions with either MSX or MSO. Induction of *Gk*0593 was able to restore the growth defects only when expressed at high levels [MSX (1 mM), MSO (200 µM) *L-(+)* arabinose], indicating these may be poor substrates for the enzyme.
Figure 5.2. 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 (10 µM), MSO (50 µM), or PPT (20 or 100 µM). Vectors were induced with 10 µM of L-(+)-arabinose. Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under Materials and Methods. Strains analyzed: _ara-9_ (JE10079), _ara-9 mddA1::cat*+ (JE18333), _ara-9 mddA1::cat*+ / pVOC (JE20864), _ara-9 mddA1::cat*+ / pMDD8 mddA*+ (JE18961). Error bars represent standard deviation. pVOC, vector only control. Symbols in D apply to all panels.
Figure 5.3. Complementation of annotated PPT acetyltransferases from *D. radiodurans* and *G. kaustophilus* in a heterologous host. Growth of an *S. enterica* mddA1::cat<sup>+</sup> strain carrying an L- (+)-arabinose inducible plasmid encoding Dr1057, Dr1182, Gk0593, or Gk2920 was examined in the presence of MSX (10 µM) or MSO (50 µM) in NCE minimal medium with glycerol (22 mM). Plasmids were induced with varying concentrations of L- (+)-arabinose (10-1000 µM), as indicated. Strains analyzed: ara-9 (JE10079), ara-9 mddA1::cat<sup>+</sup> (JE18333), ara-9 mddA1::cat<sup>-</sup> / pVOC (JE20864), ara-9 mddA1::cat<sup>+</sup> / pDr1057-2 (JE20858), ara-9 mddA1::cat<sup>+</sup> / pDr1182-2 (JE20781), ara-9 mddA1::cat<sup>-</sup> / pGk0593-2 (JE20783), ara-9 mddA1::cat<sup>-</sup> / pGk2920-2 (JE20865). Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under *Materials and Methods*. Error bars represent standard deviation. pVOC, vector only control. Symbols in B apply to panels A and B; symbols in D apply to panels C and D.
Dr1057 was not able to complement the phenotype of the mddA1::cat\textsuperscript{+} strain in either condition even at high levels of induction (1 mM $L$-$(\pm)$ arabinose).

These data provide insights into the substrate specificity of the putative PPT acetyltransferases from *D. radiodurans* and *G. kaustophilus* when using *S. enterica* as a heterologous host to monitor growth in the presence of toxic compounds.

**Increased expression provides protection against higher levels of MSO, MSO, and PPT.**

Growth of wild-type *S. enterica* is inhibited by MSX (20 µM), MSO (500 µM), and PPT (100 µM) when present at high concentrations (7). An *S. enterica* mddA\textsuperscript{+} strain carrying pDr1057, pDr1182, pGk0593, or pGk2920 *in trans* under the control of an inducible promoter was grown in minimal medium with MSX (20 µM), MSO (500 µM), or PPT (100 µM), and varying levels of inducer (10-1000 µM $L$-$(\pm)$-arabinose). When induced, cultures expressing Dr1182 or Gk2920 grew with shorter lag times when compared to the wild-type *S. enterica* strain (Fig. 5.4). These data were consistent with the idea that increased protein levels provided greater protection against the toxic effects of MSX, MSO, and PPT. Gk0593 was able to completely reduce the lag phase of wild type *S. enterica* in conditions containing PPT at low induction (10 µM $L$-$(\pm)$ arabinose) (Fig. 5.4E), but was not as efficient in providing protection against conditions including MSX or MSO, as a higher level of induction was needed to revert the mddA1::cat\textsuperscript{+} growth defect (Fig. 5.4D, 5.4F). Dr1057 was unable to improve growth of the wild type *S. enterica* strain under any conditions even at high induction (1 mM $L$-$(\pm)$ arabinose); indicating this protein does not acetylate these products under the conditions tested (Fig. 5.4A-C).

**Functional analysis of the putative PPT acetyltransferases.** Biochemical analysis of the putative PPT acetyltransferases was performed to better understand their substrate specificity (Fig. 5.5). Assays were carried out as described previously (7) using a SpectraMax Plus 384
Figure 5.4. Overexpression provides resistance to higher levels of PPT, MSX, and MSO in a heterologous host. Growth of wild-type *S. enterica* (JE10079) carrying an L-(+)-arabinose inducible plasmid encoding *Dr*1057, *Dr*1182, *Gk*0593, or *Gk*2920 was examined in presence of PPT (100 µM), MSX (20 µM), or MSO (500 µM), in NCE minimal medium with glycerol (22 mM) with varying concentrations of L-(+)-arabinose (10-1000 µM), as indicated. Strains analyzed: *ara*-9 (JE10079), *ara*-9 *mddA1::cat* (JE18333), *ara*-9 / pVOC (JE20973), *ara*-9 / pDr1057-2 (JE20857), *ara*-9 / pDr1182-2 (JE20780), *ara*-9 / pGk0593-2 (JE20782), and *ara*-9 / pGk2920-2 (JE20866). Growth curves were performed using a microplate reader (Bio-Tek Instruments) as described under *Materials and Methods*. Error bars represent standard deviation. pVOC, vector only control. Symbols in C apply to panels A-C; symbols in F apply to panels D-F.
microplate spectrophotometer (Molecular Devices). Assays were performed using a continuous spectrophotometric assay that employed 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent) as a reporter of free sulphydryl groups at 412 nm (24, 25). Reactions were initiated by the addition of enzyme (0.5 µg), and a control containing enzyme but no acetyl-CoA was used to correct for background. The *S. enterica* MddA acetyltransferase and a catalytically inactive variant, SeMddA<sup>E82Q</sup>, were used as positive and negative controls, respectively, for MSX and MSO (7). The ability of the putative PPT acetyltransferases, *Dr*1057, *Dr*1182, *Gk*0593, or *Gk*2920, to acetylate MSX, MSO, or PPT was examined (Fig. 5.5). The *in vitro* acetylation activity of *Dr*1182 correlated with the obtained *in vivo* data demonstrating the enzyme can acetylate each of the three substrates tested, with higher activity seen when MSX or MSO were used as substrates (Fig. 5.5). The activity of *Gk*0593 also correlated with the observed phenotypes, and showed the enzyme acetylated all three substrates equally well, but overall with lower activity than observed with *Dr*1182 (Fig. 5.5).

Because *Gk*2920 was recalcitrant to purification, treated whole cell lysates containing the overexpressed enzyme were used to test activity. As a negative control, a treated whole cell lysate carrying an empty vector was used. These data show *Gk*2920 acetylates MSX and MSO, but had no activity for PPT (Fig. 5.5). Based on the *in vivo* data in which expression of *Gk*2920 *in trans* improved growth in conditions with PPT, this was an unexpected result. However, mass spectrometry analysis revealed the first 55 amino acids of the N-terminus of *Gk*2920 were being cleaved during overexpression at high protein concentrations by some unknown mechanism (data not shown). This cleavage could account for the ability of the enzyme to still acetylate MSX and MSO, but not PPT.
Figure 5.5. Substrate specificity of annotated PPT acetyltransferases. Specific activity of the annotated PPT acetyltransferases of *D. radiodurans* (Dr1182) and *G. kaustophilus* (Gk0593, Gk2920) for MSO, MSX, or PPT was measured using a spectrophotometric assay (see *Materials and Methods*).
**Bioinformatic analyses of putative PPT acetyltransferases.** An alignment was generated for putative PPT acetyltransferases across a range of bacterial species (NCBI COBALT Multiple Alignment Tool, Fig. 5.6A). Looking at the alignment of the catalytic glutamate residue, we discovered that Dr1057 from *D. radiodurans*, which had no activity for MSX, MSO, or PPT under the conditions tested, lacked the catalytic glutamate, and instead encoded an asparagine. To determine if the loss of activity was due specifically to the lack of the glutamate residue, we engineered a variant of Dr1057 in which the asparagine residue was altered by site directed mutagenesis to a glutamate. Testing this variant *in trans* showed no restoration of an mddA mutant in conditions containing MSX or MSO and did not improve growth of the *S. enterica* wild type in conditions containing PPT (data not shown). Therefore we conclude that this enzyme is not lacking actively specifically due to the lack of the glutamate residue, and may have completely lost its ability to acetylate PPT and / or PPT analogues.

A phylogenetic tree was created comparing 19 PPT homologues (FigTree, Fig. 5.6B). We observed clustering of enzymes that acetylated MSX but not PPT, with the exception of Dr1182, which clustered more closely with the non-PPT utilizing enzymes. Organisms containing two annotated PPT acetyltransferases tend to be cluster separately in different nodes of the tree (*i.e.* *D. radiodurans, B. xenovorans,* and *G. kaustophilus*). A recent report examining the specificity of the two annotated PPT acetyltransferases from *P. putida*, which were also clustered in separate nodes of the tree, demonstrated that each enzyme was specific for either PPT or MSX (26). From *in vivo* data presented here for *G. kaustophilus* and *D. radiodurans*, we see that enzyme function overlaps and each of the two enzymes can acetylate both PPT and MSX, although to varying degrees.
Figure 5.6. Bioinformatic analyses of annotated PPT acetyltransferases. (A) Alignments were generated using NCBI COBALT Multiple Alignment Tool. (B) A phylogenetic tree of 19 annotated PPT acetyltransferases was generated using FigTree. Legend: Ab, *Acinetobacter baylyi*; Af, *Agrobacterium fabrum*; Bb, *Bifidobacterium bifidum*; Bs, *Bacillus subtilis*; Bx, *Burkholderia xenovorans*; Dr, *Deinococcus radiodurans*; Ec, *Escherichia coli*; Gk, *Geobacillus kaustophilus*; Lm *Listeria monocytogenes*; Pa, *Pseudomonas aeruginosa*; Pp, *Pseudomonas putida*; Sa, *Staphylococcus aureus*; Se, *Salmonella enterica*; Sc, *Streptomyces coelicolor*; and Sm, *Sinorhizobium melloti*. *For Dr1057 only the sequence of the GNAT domain that aligned to the other putative PPT acetyltransferases (a.a. 1-180) was used for the phylogenetic tree.*
DISCUSSION

Genes predicted to encode PPT acetyltransferases having varying specificities for PPT and the related analogues MSX and MSO. There is a pattern of enzymes annotated to be PPT acetyltransferases that do not retain this activity (7, 8, 12). The enzymes tested from *Geobacillus kaustophilus* and *Deinococcus radiodurans* exhibit varying degrees of specificity for the structurally related compounds MSX, MSO, and PPT (Fig. 5.3, 5.4, 5.5) (7, 8). More detailed mechanistic analyses are needed to verify the activity of this subclass of GNAT enzymes so that they are not misclassified based on sequence homology to Bar of *Streptomyces* species.

Why is there overlapping function? These studies raise question as to why an organism needs to encode multiple GNATs to combat the effects of PPT, MSX, and MSO. The only natural occurrence of MSX is from the members of the Connaraceae plant species, located in tropical regions around the world (27). MSX was also identified in the late 1940’s as a toxic by-product in flour bleached with nitrogen trichloride (28-30). PPT would be a more pervasive threat in any soil condition where *Streptomyces* species were present. On average bacteria encode ~20-25 GNATs and it is interesting that some organisms encode multiple that are specific to these related toxic amino acid derivatives.

ACKNOWLEDGEMENTS

The authors have declared that no competing interests exist. This work was supported by USPHS grant R01-GM62203 to J.C.E.-S.
REFERENCES


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


SUMMARY AND CONCLUSIONS

Overview. Acetylation of biomolecules (e.g. proteins and small molecules) is an important modification that can affect structure, DNA binding, and / or activity and is a conserved mechanism to rapidly modify cellular components in order to respond to environmental cues (1-5). The first portion of this work explored the transcriptional regulation of the reversible lysine acetylation system in *S. enterica*, providing the first example of a transcriptional regulator integrating the regulation of the RLA system with that of a target substrate, the acetyl-CoA synthetase (Hentchel *et al*, unpublished data). These studies demonstrated that slight perturbations in the ratios of the RLA components caused by mis-regulation drastically effected growth in conditions requiring a functional RLA system.

The second portion of this work focused on the characterization of GNAT enzymes annotated to encode phosphinothricin acetyltransferases. Phosphinothricin is a potent toxin produces by *Streptomyces* species and has important agricultural uses (6, 7). Characterization of the putative phosphinothricin acetyltransferase of *S. enterica*, MddA, revealed this enzyme did not have activity for phosphinothricin, but instead acetylated the closely related structural analogues methionine sulfoximine and methionine sulfone, which are toxic amino acid derivatives (8). The activities of other putative phosphinothricin acetyltransferases from *D. radiodurans*, *G. kaustophilus*, *B. xenovorans*, and *B. subtilis* were examined and revealed not only do certain
organisms encode multiple of these enzymes, but they have varying specificities for phosphinothricin, methionine sulfoximine, and methionine sulfone (Hentchel et al, unpublished data). These studies expand our understanding of the role of acetylation in cellular physiology.

**IolR regulates the reversible lysine acetylation (RLA) system.** Very little information is available on how the protein acetyltransferase (Pat) and protein deacetylase (CobB) of *S. enterica* are transcriptionally regulated (9). Work reported in Chapter 3 outlines the regulatory circuitry that integrates the expression of genes encoding the RLA system (*pat* and *cobB*), with that of their target substrate, acetyl-CoA synthetase (*acs*), and its impact on carbon metabolism (Hentchel et al, unpublished data). A mutagenesis screen identified IolR, a repressor of myo-inositol catabolism activated expression of *pat*. Genetic analyses revealed subtle effects on the RLA system that greatly impacted the growth behavior of the cell, due to the dysregulation of Acs by altering the levels of acetylated (inactive) vs. unacetylated (active) Acs. This work provides the first insights into the complexities of the system responsible for controlling RLA at the transcriptional level in *S. enterica* and how the cell regulates protein acetylation on various carbon sources.

**MddA detoxifies methionine sulfoximine and methionine sulfone, but not phosphinothricin.** GNATs have been shown to detoxify harmful compounds by acetylation (10-13). Work reported in Chapter 4 characterized a GNAT of *S. enterica*, MddA, responsible for the acetylation and detoxification of harmful amino acid derivatives (8). *In vivo* and *in vitro* analysis revealed MddA was necessary and sufficient for detoxification of methionine sulfoximine and methionine sulfone when present in the environment. Genetic and biochemical analyses identified ways in which MddA-deficient strains could prevent toxicity, including removal of the
methionine sulfoximine transport system, and the addition of exogenous amino acids. This work identified a mechanism of how *S. enterica* combats specific environment stresses.

*S. enterica* can be used as a heterologous host to assay for function of putative phosphinothricin acetyltransferases. Many putative phosphinothricin acetyltransferases are misannotated and instead, some of these enzymes have specificity for the phosphinothricin analogues methionine sulfoximine and methionine sulfone (8, 14, 15). In Chapter 5, the activities of putative phosphinothricin acetyltransferases from *D. radiodurans*, *G. kaustophilus*, *B. xenovorans*, and *B. subtilis* were determined using *S. enterica* as a heterologous host (Hentchel *et al*, unpublished data). This study provided the first example of using *S. enterica* to characterize the specificities of this subgroup of GNAT enzymes, and provided a model to predict functions of other putative phosphinothricin acetyltransferases.
FUTURE DIRECTIONS

*Does IolR directly regulate the CobB deacetylase the acetyl-CoA synthetase?* Our data show that IolR directly binds to and regulates the *pat* in *S. enterica* (Hentchel *et al*, unpublished data). While there is evidence that IolR also regulates expression of *acs* and *cobB*, the mechanism of this regulation is unclear (Hentchel *et al*, unpublished data). Gel shift assays need to be performed with IolR protein and the *acs* and *cobB* promoters to determine if IolR directly binds to and regulates these regions, or if there is another mechanism of indirect regulation. It has also been previously difficult to determine an IolR consensus site (16). If *acs* and *cobB* are directly regulated by IolR, DNA footprinting studies could be performed and these data added to previously known regulation targets could provide a framework for the development of an IolR consensus site.

*Why is IolR needed to integrate the expression of the RLA system with that of its substrate, Acs?* It is interesting that the repressor of myo-inositol catabolism (IolR) is responsible for activation of *pat*, *cobB*, and *acs* expression on various carbon sources. The data suggests that IolR is needed for full activation of the RLA system in growth containing low acetate, a condition in which activity of the RLA system is known to be important (Hentchel *et al*, unpublished data). Other data show that *pat* and *acs*, but not *cobB* expression are controlled in part by the catabolite repressor protein Crp. The link between carbon regulators and the RLA system needs further elucidation.

*Is RLA function required for optimum growth on myo-inositol?* Recently published acetylomes of *Bacillus subtilis* and *Erwinia amylovora* (17, 18) identified two enzymes involved in the degradation of myo-inositol, the malonate semialdehyde dehydrogenase (IolA) and carbohydrate kinase (IolC) amongst the acetylated proteins. It is not clear if the activity these
enzymes are under RLA control, but if so, it could provide a link between IolR regulation of RLA and RLA involvement in myo-inositol utilization.

**How does S. enterica sense and respond to methionine sulfoximine?** Preliminary work in *S. enterica* showed that the transcriptional regulators LRP and H-NS might play a role, either directly or indirectly, in the regulation of *mddA* (Hentchel *et al.*, unpublished data). It is not known at this time if and how these regulators affect expression of *mddA*. Studies examining the cascade of events from sensing methionine sulfoximine in the environment to up-regulation of *mddA* expression need to be performed before we fully understand how the cell responds to this environmental stress.

**What happens to acetylated MSX, MSO, and PPT?** After acetylation by MddA, acetyl-methionine sulfoximine, acetyl-methionine sulfone, and acetyl-phosphinothricin can no longer bind to and inhibit glutamine synthetase and are effectively rendered non-toxic. It would be interesting to know what happens to these compounds after they are detoxified. It is doubtful there is a deacetylase present in the cell that deacetylates them, as removal of the acetyl moiety would cause growth inhibition via inactivation of glutamine synthetase to resume. It is possible once acetylated the acetylated compounds are (i) exported into the supernatant, or (ii) reused in some manner by the cell to recycle carbon. If acetyl-methionine sulfoximine or acetyl-methionine sulfone were somehow reused by the cell, the data indicate that free methionine is not the outcome, as growth of a methionine auxotroph cannot be restored by either acetylated or non-acetylated methionine sulfoximine or methionine sulfone (8).

**Why do cells encode multiple phosphinothricin acetyltransferases?** Organisms on average encode ~25 GNATs. It is perplexing that 2 of these 25 would be dedicated to the acetylation and detoxification of two closely related compounds, methionine sulfoximine and phosphinothricin.
Many soil-dwelling organisms could be exposed to the presence of phosphinothricin routinely, as the compound is produced by *Streptomyces* spp., which are prolific in the environment (19). However, it is not known where bacteria would be exposed to methionine sulfoximine or methionine sulfone. In the 1900’s methionine sulfoximine was produced by the bleaching of flour by nitrogen trichloride, and is naturally produced by the *Connaraceae* plant species, only present in tropical locations (20-22) Oxidation of methionine residues typically results in production of methionine sulfoxide, not methionine sulfone (23-25). It is possible that a host response (nitric, oxidative stress responses) could generate these compounds *in vivo* (26), or that these compounds are present as the result of a normal metabolic pathway (27). However, none of these scenarios have been explored.

*What are the structural differences that determine specificity of annotated phosphinothricin acetyltransferases?* In order to fully understand the substrate specificity of the annotated phosphinothricin acetyltransferases, structural studies at the atomic level need to be performed. To date, there are crystal structures for this subclass of GNAT enzymes from *P. aeurginosa* and *A. baylyi* (14, 15). Comparisons of protein structure may allow us to determine how and why and particular GNAT annotated as a phosphinothricin acetyltransferase is able to acetylate specific compounds among closely related structural homologues. This in turn will allow us to properly predict and annotate the function of uncharacterized phosphinothricin acetyltransferases.

REFERENCES


237


