2-AMINOACRYLATE STRESS: A TOOL FOR EXPLORING METABOLIC NETWORK STRUCTURE

by

ANDREW JOHN BORCHERT

(Under the Direction of Diana Downs)

ABSTRACT

Reactive metabolites are often generated as deliberate or unintended products from a number of reactions within the cell. While the reactivity of these metabolites can play a role in improving the metabolic robustness of an organism, their unbridled accumulation an perturb the metabolic network and hinder cell fitness, necessitating the presence of cellular mechanisms controlling their abundance. The well-conserved Rid superfamily of enzymes are responsible for the hydrolysis of enamine/imine species generated as reactive intermediates from a number of amino acid metabolic pathways. 2-aminoacrylate, a particularly reactive enamine, is generated by the pyridoxal 5'-phosphate-dependent biosynthetic serine/threonine α,β-eliminase. Loss of RidA in *Salmonella enterica* allows for the accumulation of 2AA, which can damage a number of other pyridoxal 5'-phosphate-dependent enzymes and perturb cell growth. The work that follows describes an examination into the metabolic network factors that control 2AA generation by the biosynthetic serine/threonine α,β-eliminase. The data comparing 2AA stress between *Escherichia coli* and *S. enterica* highlight how conserved components can be resuctured into distinct network architectures to influence the physiology of an organism. The second part of this work focuses on using transcriptomics and metabolomics methods as a high-throughput means to effectively

describe the rewiring of metabolic network structure during 2AA stress. Transcriptomic and metabolomic readouts are increasingly being used as a rapid means to inform the physiological state of an orgaism/cell and offer a cost effective and accelerated means for physiological characterization. Importantly, generation of a model describing regulatory and metabolic consequences of a given stress using these tools is not confined to the observation of a growth phenotype or rigorous mechanistic/kinetic characterization, as has been the case during genetic and biochemical characterization of 2AA stress. Nonetheless, implementing classical biochemical genetic approaches to 'omics' workflows can prove valuable as a means to connect metabolic outputs to physiological phenomena and establish causal relationships. Generally, the work herein advocates for the continued use of careful genetic and environmental manipulations within these experiments as a means to separate physiologically relevant effects of a particular disease state or environmental stress from the total of all effects observed.

INDEX WORDS: RidA, Enamine, 2-aminoacrylate stress, Metabolic network structure, Metabolic stress, Pyridoxal 5'-phosphate

2-AMINOACRYLATE STRESS: A TOOL FOR EXPLORING METABOLIC NETWORK STRUCTURE

by

ANDREW JOHN BORCHERT

B.S., The University of Minnesota, 2013

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

2019

2-AMINOACRYLATE STRESS: A TOOL FOR EXPLORING METABOLIC NETWORK STRUCTURE

by

ANDREW JOHN BORCHERT

Major Professor: Diana M. Downs

Committee: Jorge C. Escalante-Semerena
William B Whitman
Mary A Moran

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
December 2019

DEDICATION

To my family and friends, who have shown me compassion and support; to my teachers and academic advisors, who have encouraged my pursuit of rigourous science; and to my partner, Flavia, who has provided both.

ACKNOWLEDGEMENTS

"If I have seen further than others, it is by standing upon the shoulders of giants" – Sir Isaac Newton.

During a PhD, we all ascribe to see just a bit further than others in our small niche of the academic field. I would not have been successful in achieving this without the support from all the giants in my life.

Special thanks to my advisor, Dr. Diana Downs who has enthusiastically supported and provided critical feedback for my project. Many shy away from perceived conflict; however, healthy challenge brings about the best in science. I am thankful to Diana for providing this challenge to my work and for being receptive to the challenges I have brought forward. I also valued the insight and feedback offered by my committee members, Dr. Jorge Escalante-Semerena, Dr. Mary Ann Moran, and Dr. William (Barny) Whitman. The Downs lab members, past and present, have offered day-to-day support, both scientifically and personally, and I am grateful for their camaraderie. I am particularly grateful to Dr. Dustin Ernst, Dr. Jannell Bazurto, and Mike Paxhia, as I have leaned on their insight most while formulating scientific models and while deciding what I want for life following graduate school.

I am thankful to Joe Helm for sparking my interest in biology, Andrea Abeln for introducing me to the beautiful world of microbiology (and fermentation!), and Mara Corey for broadening my perspective and sharpening my communication skills along the way. My first foray into research was with Dr. Bonnie Westra in the School of Nursing at the University of Minnesota while my interest into microbial research was sparked under the direction of Dr. Jessica Eichmiller

and Dr. Chanlan Chun in Dr. Mike Sadowsky's group, and I am greatful to all of them. My brother, Patrick Borchert, suggested I apply to graduate school and consider Diana's lab, so thank you for being right (at least this time)!

Finally, the support of my family and friends has been instrumental to shaping me into who I am today. My parents have always encouraged my curiosity and provided me the opportunity to pursue my dreams. I am grateful to my grandparents for teaching me kindness, a sense of community/culture, and introducing me to the joy of creating something beautiful. In particular, I am thankful for the continuing support of my Grandma Borchert. You are my greatest hero and mentor and the source of my success (unless it as a cribbage partner). You sometimes mention that us grandkids must get our 'smarts' somewhere else, but it was your love and support that enabled us to achieve academically. This PhD is as much yours as it is mine.

To my partner, Flavia—You embody all of the traits that I have listed above, and you act as the bridge between my academic and personal life. Thank you for taking time to edit my work and challenge my scientific models as well as for jumping into the world of homebrewing with me. I am grateful for your intellect, humor, kindness, compassion, and patience.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

Metabolic networks are comprised of biochemical reactions organized to maximize growth while minimizing the detrimental impact of reactive metabolites on cell fitness. 2-aminoacyrlate (2AA) is an internally generated reactive metabolite that inactivates pyridoxal 5'-phosphate(PLP)-dependent enzymes involved in amino acid metabolism. Members of the RidA/YER057c/UK114 superfamily of proteins play an important role in the degradation of reactive enamine/imine species, like 2AA, thereby preempting damage to PLP-dependent enzymes. Work in *Salmonella enterica* has identified a number of 2AA targets that are damaged when RidA is eliminated. However, the global consequence of these specific perturbations and the ensuing metabolic response of the organism remains relatively unexplored. This introduction provides historical context for the metabolism of reactive species and their influence on metabolic network structure which leads into a brief overview of the RidA paradigm of 2AA stress. The content that follows examines metabolic network structure and how reorganization of this network can influence the occurrence of 2AA stress in different organisms and/or environments. Then, *S. enterica*'s global transcriptomic and metabolomic response to 2AA stress is utilized as a high-throughput means to explore the integrated nature of amino acid metabolism.

1.1 REACTIVE METABOLITES AND METABOLIC NETWORK STRUCTURE

The occurrence of increasingly complex metabolic pathways has been critical as organisms have evolved from simple prebiotic forms into more complex, self-replicating structures able to sense and respond dynamically to their environmental surroundings [1-3]. During this metabolic

expansion, the inherent reactivity of many compounds help facilitate the assembly of increasing complex macromolecules [4, 5]. However, while the inherent reactivity of chemicals helps drive the expansion of metabolism required for an organism to compete and survive within a given environmental niche, unbridled chemical interaction can hinder cell fitness and must be carefully monitored [6-9]. To this end, the evolution of enzyme-mediated metabolic pathways not only helped control against such erroneous product generation [10, 11], but also allowed early organisms the ability to generate new methods of energy/carbon utilization [12, 13] and escape their geochemical constraints [14, 15]. Nonetheless, despite the numerous strategies organisms have developed to prevent, repair, and circumvent metabolic stress inflicted by reactive metabolites, these labile compounds continue to play an important role within metabolic networks.

For instance, since genetic variability of a population controls the capacity of an organism's ability to evolve, a chemical stress that increases the mutation rate of an organism can be beneficial for inducing rapid evolution toward inhabiting a particular environmental niche [7, 16, 17]. Indeed, general stress responses such as the SOS response and expression of RpoS induce a state of hypermutation [18, 19], and this stress-induced hypermutation can allow for rapid evolution overcoming the particular stress, such as amino acid starvation [20, 21]. Additionally, chemical reactivity can be harnessed to allow for alternative means of essential metabolite generation. Examples of reactive metabolites contributing to the metabolic plasticity of an organism include the use of 3-hydroxypyruvate in an alternative biosynthesis pathway for pyridoxal 5'-phosphate [22] or the repurposing of a reactive enamine generated as an intermediate in isoleucine biosynthesis, 2-aminocroatonate (2AC), in synthesis of the phosphoribosyl amine (PRA) moiety of thiamine [23-26]. Altogether, the utility of reactive species in acting as an import evolutionary

cues as well as their involvement in conferring metabolic plasticity and allowing novel chemistries results in the continued selection of these reactive metabolites in metabolic networks [16, 27].

Despite their advantages, reactive species often perturb the metabolic network and disrupt cell fitness, if not carefully controlled properly [9]. Examples of harmful metabolites that the cell must devise strategies to control lest they accumulate and damage the metabolic network include reactive oxygen species (ROS) [28, 29], reactive nitrogen species (RNS) [30], and aldehyde intermediates [31, 32]. Recently, another type of metabolic stress that has come to the forefront involves the occurrence of reactive enamine species [9]. Work in the Downs lab has focused on characterizing stress associated with the enamine product, 2-aminoacrylate (2AA), in the context of the enzyme dedicated to preempting its accumulation *in vivo*, RidA.

1.2 THE RIDA PARADIGM AND REACTIVE ENAMINES

Reactive enamine species can be generated by PLP-dependent enzymes and are critical pathway intermediates during the metabolism of many amino acids [24, 33-37]. Since enamines/imines can be hydrolyzed to their keto-acid products non-enzymatically *in vitro*, the physiological relevance of amino acid-derived reactive enamine species *in vivo* went largely overlooked until the identification of RidA as an enamine/imine deaminase[34, 38]. Many RidA homologs are encoded within operons dedicated to amino acid catabolism and these RidA enzymes may play an important role in facilitating the efficient flux through enamine/imine intermediates that are generated as intermediates within these pathways [34, 39, 40]. In the work preceding characterization of RidA as an enamine/imine deaminase, the elimination of *ridA* identified a *purF*-independent mechanism for PRA formation [26, 41]. Further characterization demonstrated that 2AC, an intermediate in isoleucine biosynthesis produced by biosynthetic threonine/serine dehydratase (IlvA), accumulated in a *ridA* mutant and could be repurposed in a anthranilate

phosphoribosyl transferase (TrpD)-dependent manner for PRA synthesis [25, 42]. Interestingly, *E. coli* seems to utilize this *purF*-independent mechanism to generate up to 50% of its PRA, despite the presence on intact RidA [23]. Importantly, IlvA is subject to allosteric inhibition by L-isoleucine, so the presence of isoleucine in the growth medium inhibits this mode of PRA formation. This finding suggests that reactive enamines can play a critical metabolic function, even in the presence of dedicated enamine/imine deaminases.

IlvA is also able to catalyze the formation of the enamine 2AA, using L-serine as a substrate [38]. In the absence of RidA, 2AA accumulates and is able to damage a number of PLP-dependent enzymes, including alanine racemase (Alr/DadX; EC:5.1.1.1) [43], aspartate aminotransferase (AspC; EC: 2.6.1.83) [44], isoleucine aminotransferase (IlvE; EC: 2.6.1.42) [38], and serine hydroxymethyltransferase (GlyA; EC: 2.1.2.1) [45]. This damage perturbs the metabolic processes orchestrated by these enzymes, causing a metabolic stress, and under certain conditions, leads to observable growth phenotypes. A summary of the RidA paradigm from *S. enterica* is provided in Figure 1.1. The RidA paradigm of enamine metabolism presents an interesting dichotomy, where reactive enamines/imines can play an important role relating to metabolic plasticity, but their unhindered accumulation can also pose a serious threat to cell fitness. This may help explain why some PLP-dependent enzymes, like IlvA generate and release reactive 2AA, despite the existence of enzymes which precede through 2AA intermediates but do not release it from the active site [36, 46, 47]. Nonetheless, 2AA accumulation and stress following RidA elimination is a far-reaching phenomenon, with examples in *Escherichia coli* [44], *Pseudomonas aeruginosa* [39], *Campylobacter jejuni* [39], *Arabidopsis* [48], maize [48], tomato [49], and *Saccharomyces cerevisiae* [50].

1.3 DISSERTATION OUTLINE

In the following dissertation, I describe my efforts to utilize the RidA paradigm of 2AA stress as a tool to expand our understanding of metabolic network structure in bacteria. In this regard, 2AA stress provides a unique perspective from which to characterize metabolic network integration, since the primary targets for 2AA damage have been well-characterized biochemically and genetically. The underlying assumption is that the regulatory and/or metabolic response or a *ridA* mutant under 2AA stress can be traced back to damage of a particular PLP-dependent enzyme. In so doing, integrated nodes made up from metabolic pathways that are both directly and indirectly affected by 2AA stress can be identified and used to expand our knowledge of metabolic network integration.

Chapter 2 begins with a literature review providing a comprehensive summary into the reactive nature of enamine/imine species, with a particular focus on the reactive enamine, 2AA, and how it fits into the RidA paradigm of metabolic stress. To date, the only enzymatic targets for 2AA damage described are PLP-dependent enzymes, and a brief discussion of potential structural factors that influence 2AA damage susceptibility is provided. Enamines are well-known for their reactivity with molecules containing nucleophilic side chains. This reactivity allows for RidA-independent mechanisms of 2AA quenching and may also allow for, as yet undiscovered, physiologically relevant perturbation of metabolic networks via direct interaction of 2AA with free metabolites.

Chapter 3 describes the surprising finding that 2AA stress outcomes differed significantly between *E. coli* and *S. enterica ridA* mutant strains. This observation prompted the discussion into metabolic structure differences between these two organisms and provides insight into a number of factors that influence the RidA paradigm.

In Chapter 4, kinetic characterization of variant forms of the 2AA generator, IlvA, provided further insight into 2AA metabolism in *S. enterica*. We demonstrate that these two IlvA variants suppress 2AA stress in an *S. enterica ridA* mutant via independent mechanisms. We also use a kinetic model to successfully predict that threonine acts as an indirect modulator of 2AA generation, through its ability to influence isoleucine generation, which is itself an allosteric inhibitor of IlvA.

In Chapter 5, attention is turned to exploring the transcriptional response of an *S. enterica ridA* mutant undergoing 2AA stress. Genes encoding many flagellar assembly components were down-regulated in a *ridA* mutant and this finding helped identify a previously unknown 2AA-dependent motility defect.

Chapter 6 details the first ^{1}H NMR metabolomics study looking into the global metabolic shifts that occur for a *ridA* mutant grown in a 2AA-generating condition (minimal glucose medium) and considers these findings alongside the transcriptomics data from Chapter 5 to construct a model describing the effect of 2AA stress in *S. enterica*. By this model, damage of serine hydroxymethyltransferase (GlyA) by 2AA causes a limitation to 5,10-methylenetetrahydrofolate (5,10-mTHF) and other folate metabolites. The indirect consequence of this damage is disruption of methionine, purine, thiamine, histidine, biotin, and mixed-acid metabolism, which the cell attempts to overcome via induction of genes in these pathways. Amino acid metabolism is also perturbed in a *ridA* mutant, a likely consequence from 2AA-dependent damage of other PLP-dependent enzyme targets.

Chapter 7 dives further into the 1H NMR metabolomic characterization of a *ridA* mutant of *S. enterica*. Supplementation of the growth medium with L-isoleucine, which allosterically inhibits the generator for 2AA, IlvA, confirmed that the bulk of metabolic shifts observed for a

ridA mutant were the consequence from 2AA stress. Glycine supplementation, which overcomes the glycine/5,10-mTHF limitation in a *ridA* mutant supported the conclusion that the perturbations to mixed acid fermentation, branched-chain amino acid metabolism, and many other metabolites were downstream consequences from GlyA damage by 2AA.

Chapter 8 comprises a summary the main conclusions from this work and outlines potential future directions for the continued study of metabolic network structure using 2AA stress outcomes.

1.4 REFERENCES

1. Ralser, M. (2014) The RNA world and the origin of metabolic enzymes. Biochem. Soc. Trans. 42 (4), 985-988.

2. Orgel, L.E. (1998) The origin of life—a review of facts and speculations. Trends Biochem. Sci. 23 (12), 491-495.

3. Woese, C.R. (1987) Bacterial evolution. Microbiol. Rev. 51 (2), 221-71.

4. Lavado, N. *et al.* (2019) From prebiotic chemistry to supramolecular oligomers: urea-glyoxal reactions. Org. Biomol. Chem. 17 (23), 5826-5838.

5. Liu, Z. *et al.* (2019) How prebiotic chemistry and early life chose phosphate. Life (Basel) 9 (1).

6. Lilja, E.E. and Johnson, D.R. (2017) Metabolite toxicity determines the pace of molecular evolution within microbial populations. BMC Evol. Biol. 17 (1).

7. Pérez-Pantoja, D. *et al.* (2013) Endogenous stress caused by faulty oxidation reactions fosters evolution of 2,4-dinitrotoluene-degrading bacteria. PLoS Genet. 9 (8), e1003764.

8. Attia, S.M. (2010) Deleterious effects of reactive metabolites. Oxid. Med. Cell. Longev. 3 (4), 238-253.

9. Linster, C.L. *et al.* (2013) Metabolite damage and its repair or pre-emption. Nat. Chem. Biol. 9 (2), 72-80.

10. Eliot, A.C. and Kirsch, J.F. (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. Annu. Rev. Biochem. 73 (1), 383-415.

11. Jensen, R.A. (1976) Enzyme recruitment in evolution of new function. Annu. Rev. Microbiol. 30 (1), 409-425.

12. Braakman, R. and Smith, E. (2012) The emergence and early evolution of biological carbon-fixation. PLoS Comp. Biol. 8 (4), e1002455.

13. Mulkidjanian, A.Y. *et al.* (2006) The cyanobacterial genome core and the origin of photosynthesis. Proc. Natl. Acad. Sci. USA 103 (35), 13126-13131.

14. Weiss, M.C. *et al.* (2016) The physiology and habitat of the last universal common ancestor. Nat. Microbiol. 1 (9).

15. Fuchs, G. (2011) Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? Annu. Rev. Microbiol. 65 (1), 631-658.

16. Galhardo, R.S. *et al.* (2007) Mutation as a stress response and the regulation of evolvability. Crit. Rev. Biochem. Mol. Biol. 42 (5), 399-435.

17. Bjedov, I. *et al.* (2003) Stress-induced mutagenesis in bacteria. Science 300 (5624), 1404-9.

18. Lombardo, M.-J. *et al.* (2004) General stress response regulator RpoS in adaptive mutation and amplification in *Escherichia coli*. Genetics 166 (2), 669-680.

19. Echols, H. (1981) SOS functions, cancer and inducible evolution. Cell 25 (1), 1-2.

20. Sung, H.M. and Yasbin, R.E. (2002) Adaptive, or Stationary-Phase, Mutagenesis, a Component of Bacterial Differentiation in *Bacillus subtilis*. J. Bacteriol. 184 (20), 5641-5653.

21. Wright, B.E. *et al.* (1999) Hypermutation in derepressed operons of *Escherichia coli* K12. Proc. Natl. Acad. Sci. USA 96 (9), 5089-5094.

22. Kim, J. and Copley, S.D. (2012) Inhibitory cross-talk upon introduction of a new metabolic pathway into an existing metabolic network. Proc. Natl. Acad. Sci. USA 109 (42), E2856-E2864.

23. Bazurto, J.V. *et al.* (2016) An unexpected route to an essential cofactor: *Escherichia coli* relies on threonine for thiamine biosynthesis. mBio 7 (1), e01840-15.

24. Lambrecht, J.A. *et al.* (2012) Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. J. Biol. Chem. 287 (5), 3454-3461.

25. Lambrecht, J.A. *et al.* (2010) Members of the YjgF/YER057c/UK114 family of proteins inhibit phosphoribosylamine synthesis *in vitro*. J. Biol. Chem. 285 (45), 34401-34407.

26. Browne, B.A. *et al.* (2006) PurF-independent phosphoribosyl amine formation in *yjgF* mutants of *Salmonella enterica* utilizes the tryptophan biosynthetic enzyme complex anthranilate synthase-phosphoribosyltransferase. J. Bacteriol. 188 (19), 6786-6792.

27. Mittler, R. *et al.* (2011) ROS signaling: the new wave? Trends Plant Sci. 16 (6), 300-309.

28. Voth, W. and Jakob, U. (2017) Stress-activated chaperones: A first line of defense. Trends Biochem. Sci. 42 (11), 899-913.

29. Imlay, J.A. (2008) Cellular defenses against superoxide and hydrogen peroxide. Annu. Rev. Biochem. 77 (1), 755-776.

30. Patra, S.K. *et al.* (2019) Reactive nitrogen species induced catalases promote a novel nitrosative stress tolerance mechanism in *Vibrio cholerae*. Nitric Oxide 88, 35-44.

31. Kato, N. *et al.* (1986) Formaldehyde dismutase, a novel NAD-binding oxidoreductase from *Pseudomonas putida* F61. Eur. J. Biochem. 156 (1), 59-64.

32. Cheng, S. *et al.* (2008) Bacterial microcompartments: their properties and paradoxes. Bioessays 30 (11-12), 1084-1095.

33. Borchert, A.J. *et al.* (2019) Reactive enamines and imines *in vivo*: Lessons from the RidA paradigm. Trends Biochem. Sci. 44 (10), 849-860.

34. Hodge-Hanson, K.M. and Downs, D.M. (2017) Members of the Rid protein family have broad imine deaminase activity and can accelerate the *Pseudomonas aeruginosa* D-arginine dehydrogenase (DauA) reaction *in vitro*. PloS one 12 (9), e0185544.

35. Ernst, D.C. *et al.* (2016) L-2,3-diaminopropionate generates diverse metabolic stresses in *Salmonella enterica*: targets of diaminopropionate in *Salmonella*. Mol. Microbiol. 101, 210-223.

36. Ernst, D.C. *et al.* (2014) Endogenous synthesis of 2-aminoacrylate contributes to cysteine sensitivity in *Salmonella enterica*. J. Bacteriol. 196 (18), 3335-3342.

37. Walsh, C. (1982) Suicide substrates: mechanism-based enzyme inactivators. Tetrahedron 38 (7), 871-909.

38. Lambrecht, J.A. *et al.* (2013) RidA proteins prevent metabolic damage inflicted by PLP-dependent dehydratases in all domains of life. mBio 4 (1), e00033-13-e00033-13.

39. Irons, J. *et al.* (2018) PA5339, a RidA homolog, is required for full growth in *Pseudomonas aeruginosa*. J. Bacteriol. 200 (22).

40. Niehaus, T.D. *et al.* (2015) Genomic and experimental evidence for multiple metabolic functions in the RidA/YjgF/YER057c/UK114 (Rid) protein family. BMC Genomics 16 (1).

41. Enos-Berlage, J.L. *et al.* (1998) Complex metabolic phenotypes caused by a mutation in *yjgF*, encoding a member of the highly conserved YER057c/YjgF family of proteins. J. Bacteriol. 180 (24), 6519-6528.

42. Lambrecht, J.A. and Downs, D.M. (2013) Anthranilate phosphoribosyl transferase (TrpD) generates phosphoribosylamine for thiamine synthesis from enamines and phosphoribosyl pyrophosphate. ACS Chemical Biology 8 (1), 242-248.

43. Flynn, J.M. and Downs, D.M. (2013) In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5'-phosphate cofactor. J. Bacteriol. 195 (16), 3603-3609.

44. Borchert, A.J. and Downs, D.M. (2017) The response to 2-aminoacrylate differs in *Escherichia coli* and *Salmonella enterica*, despite shared metabolic components. J. Bacteriol. 199 (14), e00140-17.

45. Flynn, J.M. *et al.* (2013) Decreased coenzyme A levels in *ridA* mutant strains of *Salmonella enterica* result from inactivated serine hydroxymethyltransferase: *ridA* mutants are deficient in one carbon metabolism. Mol. Microbiol. 89 (4), 751-759.

46. Xu, X.L. and Grant, G.A. (2016) Mutagenic and chemical analyses provide new insight into enzyme activation and mechanism of the type 2 iron-sulfur l-serine dehydratase from *Legionella pneumophila*. Arch. Biochem. Biophys. 596, 108-117.

47. Vederas, J.C. *et al.* (1978) Stereochemistry and mechanism of reactions catalyzed by tryptophanase *Escherichia coli*. J. Biol. Chem. 253 (15), 5350-5354.

48. Niehaus, T.D. *et al.* (2014) *Arabidopsis* and maize RidA proteins preempt reactive enamine/imine damage to branched-chain amino acid biosynthesis in plastids. Plant Cell 26 (7), 3010-3022.

49. Leitner-Dagan, Y. *et al.* (2006) CHRD, a plant member of the evolutionarily conserved YjgF family, influences photosynthesis and chromoplastogenesis. Planta 225 (1), 89-102.

50. Ernst, D.C. and Downs, D.M. (2018) Mmf1p couples amino acid metabolism to mitochondrial DNA maintenance in *Saccharomyces cerevisiae*. mBio 9 (1).

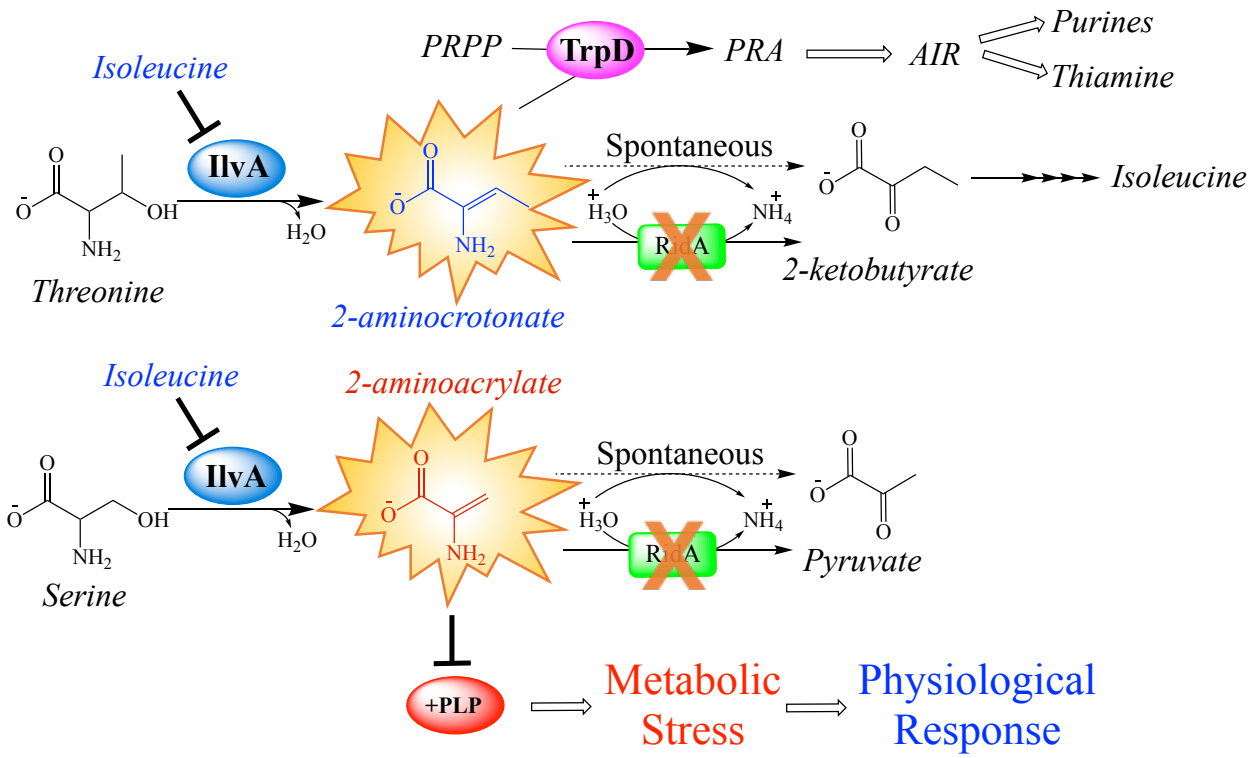


Fig. 1.1. The RidA Paradigm in *S. enterica*. Biosynthetic serine/threonine dehydratase (IlvA), which is allostericly inhibited by L-isoleucine, catalyzes the α,β-elimination of L-threonine to the reactive enamine 2-aminocrotonate (2AC), as the first comital step toward isoleucine biosynthesis. RidA facilitates the hydrolysis of 2AC to its keto acid form, 2-ketobutyrate. In a *ridA* mutant, 2AC accumulates and can be utilized in a TrpD-dependent manner for the formation of PRA, a precursor in purine and thiamine biosynthesis. IlvA also generates 2-aminoacrylate (2AA) from L-serine substrate. In RidA's absence, 2AA accumulates and damages a number of PLP-dependent enzymes, eliciting a metabolic stress.

CHAPTER 2

LITERATURE REVIEW: REACTIVE ENAMINES AND IMINES IN VIVO;

LESSONS FROM THE RIDA PARADIGM[1]

[1]Borchert AJ, Ernst DC, Downs DM. 2019. *Trends in Biochemical Sciences* 44(10): 849-860 doi: 10.1016/j.tibs.2019.04.011

Reprinted here with permission of the publisher.

2.1 ABSTRACT

Metabolic networks are webs of integrated reactions organized to maximize growth and replication while minimizing the detrimental impact that reactive metabolites can have on fitness. Enamines and imines, such as 2-aminoacrylate (2AA), are reactive metabolites produced as short-lived intermediates in a number of enzymatic processes. Left unchecked, the inherent reactivity of enamines and imines may perturb the metabolic network. Genetic and biochemical studies have outlined a role for the broadly conserved reactive intermediate deaminase (Rid) (YjgF/YER057c/UK114) protein family, in particular RidA, in catalyzing the hydrolysis of enamines and imines to their ketone product. Herein, we discuss new findings regarding the biological significance of enamine and imine production and outline the importance of RidA in controlling the accumulation of reactive metabolites.

2.2 METABOLIC NETWORKS MUST ACCOMMODATE PRODUCTION OF REACTIVE METABOLITES

Metabolism comprises a network of biochemical reactions organized according to the chemical constraints of the cell and responsive to ever-changing environmental stimuli [1–3]. These networks are arranged to maximize the output of chemicals necessary for cell growth and survival, while minimizing the detrimental impact that reactive metabolite accumulation can have on cell fitness [4–6]. The biochemical pathways in central and secondary metabolism have been elucidated over the years by combinations of *in vitro* biochemistry and *in vivo* genetics. Importantly, a dichotomy exists where reactive metabolites act as obligatory catalytic intermediates, facilitating the chemistry for the synthesis of essential compounds. However, if left unchecked their inherent reactivity can lead to detrimental disruption of the metabolic network. Examples of obligatory reactive intermediates include carbon monoxide in the Wood–Ljungdahl

pathway of acetyl-CoA synthesis [7], acetaldehyde produced during ethanolamine catabolism [8], and nitric oxide generated by aerobic ammonia oxidation [9], among many others. Reactive metabolites can also contribute to metabolic robustness, leading to the emergence of nonnative metabolic pathways. That is, the reactivity of certain metabolic intermediates or labile side-products, often acted on by nonspecific or promiscuous enzymes, can enable alternative methods of essential metabolite production [10,11].

Enamines and imines are short-lived reactive metabolites produced as reaction intermediates by multiple enzymes central to amino acid metabolism [12–16]. Despite the prevalence of these reactive molecules in ubiquitous biochemical pathways, the potential for free enamine and imine species to persist *in vivo* was largely ignored prior to the characterization of the RidA proteins as enamine deaminases [17–20]. This review summarizes the biochemical mechanisms for enamine and imine production and the *in vivo* consequences of enamine (specifically 2AA) accumulation and outlines the role of RidA in controlling enamine accumulation. We conclude with thoughts prompted by the investigation of the RidA paradigm and propose challenges and opportunities that lie ahead in efforts to understand how reactive enamines and imines fit into emerging models of the cellular metabolic network.

2.3 ENAMINES AND IMINES ARE REACTIVE METABOLITES PREVALENT THROUGHOUT METABOLISM

Enamines and imines are nitrogen-containing compounds often derived from the condensation of an aldehyde or ketone with a secondary or primary amine, respectively [21]. Peptide-bound enamines are produced biochemically via β-elimination of polypeptide residues, forming reactive peptide species that play important roles in pharmacology, enzyme engineering, and food science by facilitating formation of complex molecules [22–26]. Despite the practical use

of peptide-bound enamines, the occurrence of free enamine (and imine) species *in vivo* and any physiological implications for the chemical make-up of the cell are poorly understood.

Biochemical studies show that many pyridoxal 5′-phosphate (PLP)-dependent enzymes generate enamine and/or imine intermediates from free α-amino acid substrates (Figure 1) [27]. These α-amino acid substrates contain electronegative side chains (e.g., L-serine, L-cysteine, 3-chloro-L-alanine), which are cleaved by dedicated (or promiscuous) α,β-eliminases to generate an enamine product. Some α,β-eliminases release an enamine directly into the reaction milieu (Figure 1A) [28,29]. This behavior is observed for a number of fold-type II PLP-dependent α,β-eliminases (see below for a description of PLP enzyme fold-type), such as biosynthetic serine/threonine dehydratase (IlvA; EC 4.3.1.19), cysteine desulfhydrase (EC 2.5.1.47), and diaminopropionate ammonia-lyase (EC 4.3.1.15) [15–17,29,30]. In some cases, the cofactor-bound enamine is protonated to an iminium ion prior to release from the enzyme (Figure 1B) [31,32]. One example includes tryptophanase (TnaA; EC 4.1.99.1), a versatile fold-type I PLP-dependent lyase responsible for tryptophan catabolism (via α,β-elimination) [33], cysteine detoxification (α,β-elimination) [34], and tryptophan synthesis (β-substitution) [35]. Iminium ions are also produced as catalytic intermediates by FAD-dependent dehydrogenases

(Figure 1C). In most cases, the iminium ion intermediates are enzymatically hydrolyzed in the active site, but a number of FAD-dependent dehydrogenases release the iminium ion into the cytoplasm [19,36]. The biological relevance of unbound enamine and imine species has generally been overlooked due to their short half-lives determined *in vitro* (less than 3 min) [31,37]. However, recent studies have shown that the accumulation of enamines, specifically 2AA, can alter the physiological state of an organism, most notably through covalent damage of PLP-

dependent enzymes (Figure 1E) [18,38–41]. To that end, many organisms encode RidA, which facilitates the catalysis of enamines and imines *in vivo*, as described below (Figure 1D).

PLP-dependent enzymes are categorized according to fold-type. PLP-dependent enzymes are grouped according to structural similarities, which are generally indicative of catalytic mechanism and evolutionary relatedness [27]. Seven different fold-types of PLP-dependent enzymes have been identified (See reviews by Schneider et al., and Percudani and Peracchi for a thorough coverage of PLP-dependent enzyme fold-type) [42, 43]. Briefly, fold-type I (aspartate aminotransferase family) is the most diverse and contains the largest number of aminotransferases, as well as a number of decarboxylases, lyases, and other enzymes. Fold-type II (tryptophan synthase β family) contains alkyltransferases, ammonia lyases (α,β-eliminases), and a number of racemases, while fold-type III (alanine racemase family) consists of all other PLP-dependent racemases and decarboxylases. The other class containing aminotransferases is fold-type IV (D-amino acid aminotransferase family), which includes D-amino acid aminotransferases, L-branched chain aminotransferases, and 4-amino-4-deoxychorismate lyases. The three remaining fold-types (V, glycogen phosphorylases; VI, D-lysine-5,6-aminomutases; and VII, L-lysine-2,3-aminomutases) each contain members belonging to one reaction type. The chemistry behind the breadth of reactions catalyzed by PLP-dependent enzymes, structural differences between fold-types, and their biological and industrial significance have been outlined in a number of reviews [27, 44-47]. Reaction specificity between PLP-dependent enzyme fold types is mediated through stereoelectronic control (cleaved bond must be aligned perpendicular to the PLP plane), electrophilic strength of the Schiff base, and catalytic side chain placement [47]. To the second point, the residues coordinating the pyridine nitrogen and the O3' phenol oxygen from PLP appear to be largely responsible for controlling the electrophilic strength differences of the Schiff base

between fold-type I-IV enzymes, and their identities separate well with fold type and the respective reactions these enzymes catalyze (Figure 2) [48, 49].

2.4 CHARACTERIZATION OF RIDA AS A NEW ENAMINE AND IMINE DEAMINASE

The Rid Superfamily and Ability of RidA to Hydrolyze Enamine and/or Imine Species. The Rid family of proteins (formerly YjgF/YER057c/UK114) is a class of proteins identified by intersubunit clefts made from seven well-conserved residues and found in all domains of life (Figure 3) [50,51]. Recent phylogenetic analysis used synteny to split the Rid superfamily members into eight subfamilies: an archetypal RidA subfamily and seven other subfamilies (Rid1–Rid7) [36]. Rid1–Rid7 subfamily members are largely confined to bacteria and are often encoded in genomes also encoding RidA members [36]. Using the residue numbers from the *Salmonella enterica ridA* sequence, the most highly conserved residue across the Rid superfamily is glutamate 120 (E120). S30 and G31 are also common across all subfamilies and thus assumed to be significant in the function of Rid enzymes [36].

While multiple crystal structures of RidA homologs proved useful in identifying the putative active site of this class of proteins [52,53], a number of genetic observations ultimately led to the identification of the biochemical function and physiological role of RidA [54–57]. As mentioned above, *S. enterica* IlvA is an α,β-eliminase with activity on serine (K_M = 43 mM) or threonine (K_M = 6 mM) as substrates [58]. *In vitro* IlvA alone produces α-ketobutyrate from threonine or pyruvate from serine, showing that solvent water is sufficient for hydrolysis of the enamine to the corresponding keto acid [17]. The addition of purified RidA to the IlvA reactions increased the rate of ketoacid formation, showing that RidA catalyzes the hydrolysis of enamines derived from threonine [2-aminocrotonate (2AC)] and serine (2AA) [17]. These and other data culminated in the designation of RidA as an enamine/imine deaminase [55,56]. Purified RidA

orthologs from organisms across the tree of life displayed conserved enamine hydrolysis-enhancing activity *in vitro*, confirming the designation of this protein subfamily as enamine deaminases [17].

RidA has a broad substrate range, acting not only on 2AA and 2AC but also on the imines of glutamine, histidine, leucine, methionine, and phenylalanine generated by FAD-dependent oxidases [19,36]. Currently, a physiological role for RidA in the deamination of 2AC and 2AA has been identified in *S. enterica* and other organisms [18,38,59,60], but the significance of its activity on other substrates remains to be determined.

Proposed Mechanism of RidA-Dependent Enamine and Imine Catalysis. Efforts to define the catalytic residues involved in the deaminase activity of RidA by site-directed mutagenesis identified a conserved arginine (R105) as the sole residue essential for function *in vivo*; variants altered at this site had greatly decreased activity *in vitro* [17]. RidA variants E120A and Y17F also had decreased deaminase activity *in vitro*, but retained the ability to complement *ridA* mutant phenotypes [17]. These data, the conserved residue observations above, and the crystal structure of the RidA homolog TdcF [52] suggested a mechanism for how RidA catalyzes the hydrolysis of iminium ion substrates. A representation of this mechanism using 2-iminiopropionate (2IP) as an example substrate is shown in Figure 3. In the mechanistic model, R105 forms a bidentate salt bridge with the carboxylic acid of the substrate. Substitution of this residue is expected to disrupt the critical interaction with the substrate carboxylic acid, consistent with the loss of activity of the RidA$_{R105A}$ variant [17]. R105 is absolutely conserved in RidA and Rid1–Rid3 subfamily members, predicting that members of these protein subfamilies act on amino acid-derived substrates [36]. Furthermore, the Rid4–Rid7 subfamilies lack R105 and no amino acid-derived enamine/imine deaminase activity has been detected for these subfamilies [19]. Based

on the crystal structure of TdcF, the backbone carbonyl groups of R105 and G31 appear to stabilize the iminium ion (e.g., 2IP) formed from the substrate, while the backbone of C107 and the side chain of E120 coordinate the water involved in hydrolysis of 2IP [52]. Although the TdcF crystal structure suggests that the hydroxyl group of Y17 is too far away to hydrogen bond with the substrate, this residue may directly interact with the substrate/product or play an important role in the proper folding of the active site. The TdcF crystal structure also shows a number of solvent water molecules localized near Y17, making it reasonable that the Tyrosine hydroxyl group coordinates a water molecule near the imine nitrogen, where it acts as a proton donor for the formation of ammonia [52]. Importantly, this model proposes a mechanism for iminium ion hydrolysis. Consistent with RidA enzymes acting on iminium ion substrates, when iminium ions are generated directly via FAD-dependent oxidases (Figure 1C), inclusion of RidA increases the rate of hydrolysis *in vitro* [19,20]. However, it remains unclear whether RidA decreases enamine accumulation by binding the enamine and facilitating both iminium ion formation and subsequent hydrolysis or whether RidA decreases enamine levels simply through increased consumption of its iminium ion tautomer [17].

2.5 CHEMISTRY OF FREE ENAMINE SPECIES: FROM *IN VITRO* DISCOVERY TO BIOLOGICAL RELEVANCE

Inactivation of PLP-Dependent Enzymes by 2AA. Many non-eliminase PLP-dependent enzymes are promiscuous and capable of catalyzing adventitious α,β-elimination reactions using amino acid substrates with strong electronegative leaving groups (e.g., -Cl, -SO_3-) bound to their β-carbon [27]. *In vitro* studies showed that the 2AA generated during these promiscuous elimination reactions can immediately and irreversibly damage the source PLP-dependent enzymes before 2AA escapes the active site [61–63]. The inactivation could be caused by: (i)

release of 2AA from PLP and subsequent nucleophilic attack of the electrophilic enzyme-bound PLP Schiff base by the β-carbon of 2AA [62] (Figure 4A); or (ii) attack of the 2AA/PLP adduct by active site nucleophilic residues [63] (Figure 4B). Both scenarios produce an irreversibly inactivated enzyme [64]. Critically, these *in vitro* systems contained a single enzyme that both generated and was inactivated by 2AA. The precedent of 2AA's inactivation of PLP-dependent enzymes, coupled with biochemical characterization of RidA and genetic observations in *ridA* mutant backgrounds, led to a model of the biologically relevant function of RidA [17,38]. An extrapolation of the *in vitro* results suggested that, *in vivo*, PLP-dependent α,β-eliminases generate 2AA that could diffuse from the generator enzyme(s) and covalently modify susceptible PLP-dependent target enzymes [17]. The critical assumption was that, in the absence of RidA, 2AA would persist long enough in the cellular milieu to interact with and inactivate target enzymes. Given this assumption, several phenotypes of *ridA* mutant strains were predicted to be the consequence of 2AA-dependent inactivation of PLP enzymes [38,40,41].

IlvA is the Primary Source of 2AA Stress in *S. enterica*. The *ridA* mutant phenotypes were considered first with the purpose of identifying the proposed enzyme generator(s) of 2AA. *S. enterica ridA* mutants display a number of phenotypes, as shown in Table 1. It was particularly telling that the strain was sensitive to exogenous serine,

and this sensitivity was reversed by the addition of isoleucine [54]. Isoleucine controls the activity of the serine/threonine dehydratase IlvA by feedback inhibition [65]. When the wild-type *ilvA* allele was replaced with an allele (*ilvA219*) encoding the feedback-insensitive IlvAL447F variant, isoleucine failed to reverse the serine sensitivity, indicating that the effect of isoleucine was through allosteric regulation of IlvA [55,56]. These and other data led to the key conclusion that multiple phenotypes of a *ridA* mutant in *S. enterica* were due to the IlvA-dependent generation of

2AA from serine [17,38]. Although IlvA-dependent dehydration of L-serine is the predominant source of 2AA in *S. enterica* under standard growth conditions, other PLP-dependent α,β-eliminases are capable of generating substantial levels of free 2AA when their substrates are supplied exogenously [15,16] (Table 1).

Free 2AA Inactivates Multiple Cellular PLP-Dependent Enzymes. While a serine/threonine dehydratase is the dominant 2AA generator in *S. enterica* and in most other organisms tested thus far [18,60,66], multiple enzymes are potentially damaged by 2AA. For instance, branched-chain amino acid aminotransferase (IlvE) activity is decreased by 30–50% in *ridA* mutants; this decreased activity was confirmed to be due to covalent modification of the IlvE active site by 2AA [38]. Importantly, the generation of 2AA (by IlvA), inactivation of IlvE, and preventive quenching of 2AA by RidA were all reconstituted *in vitro* [38]. Growth phenotypes showed that multiple metabolic pathways were compromised in a *ridA* strain experiencing 2AA stress (Table 1). Thus far, assays have shown that serine hydroxymethyltransferase (GlyA; EC 2.1.2.1), aspartate aminotransferase (AspC; EC 2.6.1.1), and alanine racemases (Alr/DadX; EC 5.1.1.1), in addition to branched-chain amino acid aminotransferase (IlvE; EC 2.6.1.42), have lower specific activity in *ridA* mutant strains than in wild-type [38,40,41,59]. The mechanism of inactivation has not been rigorously defined for target PLP enzymes other than IlvE. Significantly, in a *ridA* mutant the activity of multiple target proteins is compromised but not eliminated, resulting in perturbation of the metabolic network that can result in multiple pleiotropic phenotypes. For some of these phenotypes (i.e., motility defect), the presumed target enzyme has not been identified [59,66].

The work summarized above established the RidA paradigm in *S. enterica*, which in its simplest form states that 2AA produced by PLP-dependent α,β-eliminases can be deaminated by

RidA. In the absence of RidA, the 2AA can: (i) persist *in vivo*; and (ii) diffuse into and inactivate distinct target PLP-dependent enzymes. This definition provided a starting point to explore the causes and implications of 2AA stress in a variety of organisms. Beyond *S. enterica*, enamine-related stress extends into other species, and RidA plays an important part in controlling enamine stress in these organisms [18,19,36,60,66,67].

2.6 THE CONUNDRUM OF PLP ENZYMES BEING BOTH GENERATORS AND TARGETS OF 2AA

Enzymes Differ in their Susceptibility to 2AA. A dichotomy exists between PLP-dependent enzymes that generate 2AA and those damaged by free 2AA. Presumably, 2AA generators are largely recalcitrant to 2AA-dependent damage, although these enzymes do not appear to be entirely immune to mechanism-based inactivation by 2AA [13,68]. This form of inactivation involves the enzyme-catalyzed generation of 2AA from a 2AA precursor (β-substituted alanine species); 2AA is then displaced and nucleophilically attacks the enzyme used to generate it, rendering it permanently inactive. Some PLP-dependent enzymes can become sensitized to mechanism-based 2AA inactivation through site-directed mutagenesis [69]. This is exemplified by tryptophan synthase (EC 4.2.1.20), which catalyzes the last two reactions in the biosynthesis of tryptophan [70]. The PLP-dependent β-subunit of tryptophan synthase catalyzes both β-elimination and β-substitution reactions, but variant β-subunits bias reactions toward β-elimination relative to wild-type tryptophan synthase [71]. These same variants are susceptible to inactivation by 2AA produced from L-serine or 3-chloro-L-alanine, whereas the wild-type enzyme was undamaged by serine and less damaged than the variant enzyme by 3-chloro-L-alanine [72]. These data showed that changing a single active-site residue profoundly influenced susceptibility to 2AA inactivation, suggesting that a fine line exists between substrate-based inactivation and

successful elimination reactions. This work provides insights about how enzymes routinely exposed to 2AA could be protected from damage. The fact that enzymes proficient at α,β-elimination and 2AA generation are still susceptible to mechanism-based inactivation [13,68], albeit less so than promiscuous eliminases [68], highlights the reactivity and potential toxicity of 2AA.

PLP Enzyme Acitive-Site Residues May Determine Susceptibility to 2AA Damage. PLP-dependent enzymes are generally grouped into one of seven fold-types according to structural similarities [73,74]. Pairing 2AA interaction profiles with the relevant PLP-dependent enzymes revealed a correlation between fold-type and 2AA production/inactivation; these data included four enzymes inactivated by 2AA and four that generated 2AA (Table 1). Generators, or enzymes most immune to 2AA damage, belonged to fold-type II [15–17]. Conversely, enzymes damaged via the covalent addition of 2AA/PLP mechanism (Figure 4B) were fold-type I and IV enzymes [38,41] and those damaged via the PLP/pyruvate adduct mechanism (Figure 4A) were fold-type III racemases [38,40,41]. Reactions catalyzed by fold-type I and IV PLP-dependent enzymes involve a quinonoid intermediate [48]. Enzymes belonging to these fold-types satisfy this requirement through the placement of a negatively charged amino acid adjacent to the pyridine N from PLP, which promotes the nitrogen's protonation and allows subsequent quinonoid formation [48]. Alternatively, L-alanine racemization (fold-type III) and α,β-elimination (fold-type II) do not require the formation of a quinonoid intermediate [48]. Formation of a quinonoid intermediate is an essential step in the mechanism of damage via 2AA/PLP (Figure 4B) and may explain why the enzymes believed to be damaged in this way belong to fold-type I and IV. By contrast, the PLP/pyruvate adduct damage mechanism (Figure 4A) requires the 4′C of PLP to act as an electrophile for successful attack by the 2AA-derived carbanion [64]. Therefore, the finding that

fold-type III racemases resist the formation of a quinonoid intermediate may help to explain why these enzymes are damaged via the PLP/pyruvate adduct mechanism and not the 2AA/PLP mechanism [42]. These data support a model where fold-type II enzymes are more recalcitrant to 2AA damage via the 2AA/PLP mechanism due to their decreased propensity for forming a quinonoid species. By contrast, differences in the relative electrophilic strength of the C4′ may account for fold-type II enzymes being more immune than fold-type III to the PLP/pyruvate adduct mechanism of 2AA-dependent damage. Further work using variant enzymes will be valuable in better defining features of the active site that determine the susceptibility of an enzyme to 2AA damage.

2.7 UNDERSTANDING THE GLOBAL SCOPE OF ENAMINE/IMINE REACTIVITY

The Global Metabolic Effects of 2AA are Largely Unexplored. A number of PLP-enzymes targeted by 2AA have been identified, primarily by work in *S. enterica* [38,40,41,67]. However, in the context of metabolic network configurations, the large number of potential targets (>40 PLP-dependent enzymes are encoded in *S. enterica* [73,75]) complicates efforts to define the global physiological consequences of 2AA stress. Moreover, the detectable physiological impact (e.g., phenotype) of 2AA stress varies by organism, reflecting how organism-specific metabolic network architecture dictates 2AA damage outcome [60,66,67]. As a first step toward understanding the global impact of 2AA stress in *S. enterica*, one study used differential gene expression in a *ridA* mutant as a downstream indicator of the metabolic perturbation caused by 2AA stress [59]. The resulting data identified a previously unknown 2AA-dependent motility defect but failed to reflect dramatic changes in the expected metabolic pathways linked to 2AA production or damage. These initial global efforts have emphasized the need for innovative

technical approaches to better capture the complex metabolic environment in a living cell and identify the potentially subtle consequences of 2AA stress.

RidA-Independent Mechanisms for Quenching 2AA. Studies of RidA highlight the importance of reactive metabolite control mechanisms *in vivo* despite the spontaneous lability of some reactive metabolites *in vitro*. Multicopy suppressor analyses revealed alternative, RidA-independent mechanisms for reactive enamine control. For example, when overexpressed, cystathionine β-lyase (MetC; EC 4.4.1.8) and aspartate/glutamate racemase (YgeA, MMP0739; EC 5.1.1.13) quench 2AA in *S. enterica* mutants lacking RidA [30,76]. This work suggests that these enzymes do not catalyze enamine hydrolysis but instead produce another reactive metabolite that reacts with and quenches 2AA. In the case of MetC, the data support a model where MetC acts on a non-cystathionine sulfur-containing substrate to generate a free thiol group. The free thiol is then thought to generate a stable adduct with 2AA [30]. Similar thiol-mediated reactions with enzyme-bound enamines, such as during lanthionine formation, are well established [24,83,84]. Alternatively, YgeA and MMP0739 contain a threonine in place of an active-site cysteine residue found in well-characterized PLP-independent racemases [76]. The absence of this cysteine residue is proposed to allow the formation of a carbanion intermediate that participates in a nucleophilic attack on 2AA, quenching it through derivatization [76]. This model is reminiscent of alkyl radical quenching of enzyme-bound dehydroalanine used for post-translational enzyme engineering [22,23]. The discovery of these two alternative systems for quenching 2AA underscores a number of points. First, RidA may represent one of several strategies that have evolved to limit the accumulation of reactive enamines/ imines *in vivo*. A handful of non-RidA enzymes have been reported to quench imines, including aldehyde oxidase (AOX; EC 1.2.3.1) [79], pyrroline 5-carboxylate reductase (ProC; EC 1.5.1.2) [80], and glutamate

dehydrogenase (GdhA; EC 1.4.1.4) [81]. Second, while 2AA has a demonstrated role in damaging PLP-dependent enzymes, this molecule may also perturb the metabolic network through direct interaction with other free reactive metabolites.

2.8 CONCLUDING REMARKS

The study of RidA and its role in catalyzing the hydrolysis of enamines and imines has spurred a new interest in the prevalence and impact of these and other reactive species in the metabolic network. New questions have emerged concerning the strategies used by cells to control endogenous levels of enamines and imines as well as the metabolic fate and fitness impact if their accumulation is left unchecked. The prevalence and consequence of enamine and/or imine species in organisms remain largely unknown and are exciting areas for future studies. The focus of this review was limited to the deleterious consequences of uncontrolled enamine accumulation. However, less-explored physiological benefits of enamine accumulation have been described [17,82], suggesting that there may be additional instances where enamines positively influence metabolic robustness. Future work on probing the generation and consequences of enamines and imines in wild-type organisms, characterizing other Rid family proteins, and Rid-independent mechanisms of quenching enamines and imines will provide insights about the role of reactive metabolites and their control in the maintenance of a robust metabolic network.

2.9 REFERENCES

1. Keller, M.A. *et al.* (2015) The widespread role of non-enzymatic reactions in cellular metabolism. Curr. Opin. Biotechnol. 34, 153-161.

2. de Lorenzo, V. *et al.* (2015) Chemical reactivity drives spatiotemporal organisation of bacterial metabolism. FEMS Microbiol. Rev. 39 (1), 96-119.

3. Albert, R. *et al.* (2000) Error and attack tolerance of complex networks. Nature 406 (6794), 378-382.

4. de Crecy-Lagard, V. *et al.* (2018) Newly-discovered enzymes that function in metabolite damage-control. Curr. Opin. Chem. Biol. 47, 101-108.

5. Imlay, J.A. (2013) The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. Nat. Rev. Microbiol. 11 (7), 443-454.

6. Semchyshyn, H.M. (2014) Reactive carbonyl species in vivo: generation and dual biological effects. Sci. World J. 2014, 417842.

7. Menon, S. and Ragsdale, S.W. (1996) Evidence that carbon monoxide is an obligatory intermediate in anaerobic acetyl-CoA synthesis. Biochemistry 35 (37), 12119-12125.

8. Brinsmade, S.R. *et al.* (2005) Minimal functions and physiological conditions required for growth of *Salmonella enterica* on ethanolamine in the absence of the metabolosome. J. Bacteriol. 187 (23), 8039-8046.

9. Caranto, J.D. and Lancaster, K.M. (2017) Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase. Proc. Natl. Acad. Sci. USA. 114 (31), 8217-8222.

10. Kim, J. and Copley, S.D. (2012) Inhibitory cross-talk upon introduction of a new metabolic pathway into an existing metabolic network. Proc. Natl. Acad. Sci. USA. 109 (42), E2856-64.

11. Bazurto, J.V. *et al.* (2016) An unexpected route to an essential cofactor: *Escherichia coli* relies on threonine for thiamine biosynthesis. mBio 7 (1), e01840-15.

12. Chargaff, E. and Sprinson, D.B. (1943) Studies on the mechanism of deamination of serine and threonine in biological systems. J. Biol. Chem. 151, 273-280.

13. Phillips, A.T. and Wood, W.A. (1965) The mechanism of action of 5'-adenylic acid-activated threonine dehydratase. J. Biol. Chem. 240 (12), 4703-4709.

14. Clausen, T. *et al.* (1996) Crystal structure of the pyridoxal-5'-phosphate dependent cystathionine beta-lyase from *Escherichia coli* at 1.83 A. J. Mol. Biol. 262 (2), 202-224.

15. Ernst, D.C. *et al.* (2014) Endogenous synthesis of 2-aminoacrylate contributes to cysteine sensitivity in *Salmonella enterica*. J. Bacteriol. 196 (18), 3335-3342.

16. Ernst, D.C. *et al.* (2016) L-2,3-diaminopropionate generates diverse metabolic stresses in *Salmonella enterica*. Mol. Microbiol. 101 (2), 210-223.

17. Lambrecht, J.A. *et al.* (2012) Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. J. Biol. Chem. 287 (5), 3454-3461.

18. Niehaus, T.D. *et al.* (2014) *Arabidopsis* and maize RidA proteins preempt reactive enamine/imine damage to branched-chain amino acid biosynthesis in plastids. Plant Cell 26 (7), 3010-3022.

19. Hodge-Hanson, K.M. and Downs, D.M. (2017) Members of the Rid protein family have broad imine deaminase activity and can accelerate the *Pseudomonas aeruginosa* D-arginine dehydrogenase (DauA) reaction *in vitro*. PLoS One 12 (9), e0185544.

20. Degani, G. *et al.* (2018) Imine deaminase activity and conformational stability of UK114, the mammalian member of the Rid protein family active in amino acid metabolism. Int. J. Mol. Sci. 19 (4) 945.

21. Adams, J.P. (2000) Imines, enamines and oximes. J. Chem. Soc. Perkin Trans. 1 (2), 125-139.

22. Wright, T.H. *et al.* (2016) Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity. Science 354 (6312), aag1465.

23. Yang, A. *et al.* (2016) A chemical biology route to site-specific authentic protein modifications. Science 354 (6312), 623-626.

24. Field, D. *et al.* (2015) Bioengineering lantibiotics for therapeutic success. Front. Microbiol. 6, 1363.

25. Seebeck, F.P. and Szostak, J.W. (2006) Ribosomal synthesis of dehydroalanine-containing peptides. J. Am. Chem. Soc. 128 (22), 7150-7151.

26. Gorham, S.D. *et al.* (1992) Effect of chemical modifications on the susceptibility of collagen to proteolysis. II. Dehydrothermal crosslinking. Int. J. Biol. Macromol. 14 (3), 129-138.

27. Eliot, A.C. and Kirsch, J.F. (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. Annu. Rev. Biochem. 73, 383-415.

28. Flavin, M. and Slaughter, C. (1969) Enzymic reactions of enamines with N-ethylmaleimide. J. Biol. Chem. 244 (6), 1434-1444.

29. Datta, P. and Bhadra, R. (1978) Biodegradative threonine dehydratase. Reduction of ferricyanide by an intermediate of the enzyme-catalyzed reaction. Eur. J. Biochem. 91 (2), 527-532.

30. Ernst, D.C. *et al.* (2018) Increased activity of cystathionine beta-lyase suppresses 2-aminoacrylate stress in *Salmonella enterica*. J. Bacteriol. 200 (9), e00040-18.

31. Hillebrand, G.G. *et al.* (1979) Formation of an intermediate and its rate of conversion to pyruvate during the tryptophanase-catalyzed degradation of S-o-nitrophenyl-L-cysteine. Biochemistry 18 (9), 1751-1755.

32. Vederas, J.C. *et al.* (1978) Stereochemistry and mechanism of reactions catalyzed by tryptophanase from *Escherichia coli*. J. Biol. Chem. 253 (15), 5350-5354.

33. Newton, W.A. and Snell, E.E. (1964) Catalytic properties of tryptophanase, a multifunctional pyridoxal phosphate enzyme. Proc. Natl. Acad. Sci. USA. 51, 382-389.

34. Awano, N. *et al.* (2005) Identification and functional analysis of *Escherichia coli* cysteine desulfhydrases. Appl. Environ. Microbiol. 71 (7), 4149-4152.

35. Chan, R.K. *et al.* (1972) Specialized transduction of tetracycline resistance by phage P22 in *Salmonella typhimurium*. II Properties of a high transducing lysate. Virology 50, 883-898.

36. Niehaus, T.D. *et al.* (2015) Genomic and experimental evidence for multiple metabolic functions in the RidA/YjgF/YER057c/UK114 (Rid) protein family. BMC Genomics 16, 382.

37. Flavin, M. and Slaughter, C. (1964) An intermediate trapped by maleimides in a pyridoxal-phosphate potentiated enzymatic elimination reaction. Biochemistry 3, 885-93.

38. Lambrecht, J.A. *et al.* (2013) RidA proteins prevent metabolic damage inflicted by PLP-dependent dehydratases in all domains of life. mBio 4 (1), e00033-13.

39. Lambrecht, J.A. and Downs, D.M. (2013) Anthranilate phosphoribosyl transferase (TrpD) generates phosphoribosylamine for thiamine synthesis from enamines and phosphoribosyl pyrophosphate. ACS Chem. Biol. 8 (1), 242-248.

40. Flynn, J.M. *et al.* (2013) Decreased coenzyme A levels in *ridA* mutant strains of *Salmonella enterica* result from inactivated serine hydroxymethyltransferase. Mol. Microbiol. 89 (4), 751-759.

41. Flynn, J.M. and Downs, D.M. (2013) In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5'-phosphate cofactor. J. Bacteriol. 195 (16), 3603-3609.

42. Watanabe, A. *et al.* (2002) Reaction mechanism of alanine racemase from *Bacillus stearothermophilus*: x-ray crystallographic studies of the enzyme bound with N-(5'-phosphopyridoxyl)alanine. J. Biol. Chem. 277 (21), 19166-19172.

43. Phillips, R.S. (2015) Chemistry and diversity of pyridoxal-5'-phosphate dependent enzymes. Biochim. Biophys. Acta 1854 (9), 1167-1174.

44. Steffen-Munsberg, F. *et al.* (2015) Bioinformatic analysis of a PLP-dependent enzyme superfamily suitable for biocatalytic applications. Biotechnol. Adv. 33 (5), 566-604.

45. Toney, M.D. (2011) Controlling reaction specificity in pyridoxal phosphate enzymes. BBA-Proteins Proteom. 1814 (1), 1407-1418.

46. Casasnovas, R. *et al.* (2012) C-H activation in pyridoxal-5'-phosphate Schiff bases: the role of the imine nitrogen. A combined experimental and computational study. J. Phys. Chem. B 116 (35), 10665-10675.

47. Ernst, D.C. and Downs, D.M. (2016) 2-aminoacrylate stress induces a context-dependent glycine requirement in *ridA* strains of *Salmonella enterica*. J. Bacteriol. 198 (3), 536-543.

48. Griswold, W.R. and Toney, M.D. (2011) Role of the pyridine nitrogen in pyridoxal 5'-phosphate catalysis: activity of three classes of PLP enzymes reconstituted with deazapyridoxal 5'-phosphate. J. Am. Chem. Soc. 133 (37), 14823-14830.

49. Singh, R. *et al.* (2013) Chemogenomics of pyridoxal 5 '-phosphate dependent enzymes. J. Enzym. Inhib. Med. Ch. 28 (1), 183-194.

50. Leitner-Dagan, Y. *et al.* (2006) CHRD, a plant member of the evolutionarily conserved YjgF family, influences photosynthesis and chromoplastogenesis. Planta 225 (1), 89-102.

51. Parsons, L. *et al.* (2003) Solution structure and functional ligand screening of HI0719, a highly conserved protein from bacteria to humans in the YjgF/YER057c/UK114 family. Biochemistry 42 (1), 80-89.

52. Burman, J.D. *et al.* (2007) The crystal structure of *Escherichia coli* TdcF, a member of the highly conserved YjgF/YER057c/UK114 family. BMC Struct. Biol. 7, 30.

53. Volz, K. (1999) A test case for structure-based functional assignment: the 1.2 A crystal structure of the *yjgF* gene product from *Escherichia coli*. Prot. Sci. 8 (11), 2428-2437.

54. Enos-Berlage, J.L. *et al.* (1998) Complex metabolic phenotypes caused by a mutation in *yjgF*, encoding a member of the highly conserved YER057c/YjgF family of proteins. J. Bacteriol. 180 (24), 6519-6528.

55. Schmitz, G. and Downs, D.M. (2004) Reduced transaminase B (IlvE) activity caused by the lack of *yjgF* is dependent on the status of threonine deaminase (IlvA) in *Salmonella enterica* Serovar Typhimurium. J. Bacteriol. 186 (3), 803-810.

56. Christopherson, M.R. *et al.* (2008) YjgF is required for isoleucine biosynthesis when *Salmonella enterica* is grown on pyruvate medium. J. Bacteriol. 190 (8), 3057-3062.

57. Christopherson, M.R. *et al.* (2012) Suppressor analyses identify threonine as a modulator of *ridA* mutant phenotypes in *Salmonella enterica*. PLoS One 7 (8), e43082.

58. Borchert, A.J. and Downs, D.M. (2018) Analyses of variants of the Ser/Thr dehydratase IlvA provide insight into 2-aminoacrylate metabolism in *Salmonella enterica*. J. Biol. Chem. 293 (50), 19240-19249.

59. Borchert, A.J. and Downs, D.M. (2017) Endogenously generated 2-aminoacrylate inhibits motility in *Salmonella enterica*. Sci. Rep. 7 (1), 12971.

60. Ernst, D.C. and Downs, D.M. (2018) Mmf1p couples amino acid metabolism to mitochondrial DNA maintenance in *Saccharomyces cerevisiae*. mBio 9 (1), e00084-18.

61. Walsh, C. (1982) Suicide substrates: mechanism-based enzyme inactivators. Tetrahedron 38, 871-909.

62. Likos, J.J. *et al.* (1982) A novel reaction of the coenzyme of glutamate decarboxylase with L-serine O-sulfate. Biochemistry 21 (18), 4377-4386.

63. Ueno, H. *et al.* (1982) Chemistry of the inactivation of cytosolic aspartate aminotransferase by serine O-sulfate. Biochemistry 21 (18), 4387-4393.

64. Esaki, N. and Walsh, C.T. (1986) Biosynthetic alanine racemase of *Salmonella typhimurium*: purification and characterization of the enzyme encoded by the *alr* gene. Biochemistry 25 (11), 3261-3267.

65. Eisenstein, E. (1995) Allosteric regulation of biosynthetic threonine deaminase from *Escherichia coli*: effects of isoleucine and valine on active-site ligand binding and catalysis. Arch. Biochem. Biophys. 316 (1), 311-8.

66. Irons, J. *et al.* (2018) PA5339, a RidA homolog, is required for full growth in *Pseudomonas aeruginosa*. J. Bacteriol. 200 (22).

67. Borchert, A.J. and Downs, D.M. (2017) The response to 2-aminoacrylate differs in *Escherichia coli* and *Salmonella enterica*, despite shared metabolic components. J. Bacteriol. 199 (14).

68. Arfin, S.M. and Koziell, D.A. (1971) Inhibition of growth of *Salmonella typhimurium* and of threonine deaminase and transaminase B by beta-chloroalanine. J. Bacteriol. 105 (2), 519-522.

69. Ahmed, S.A. *et al.* (1991) Mechanism of mutual activation of the tryptophan synthase-alpha and beta-subunits - analysis of the reaction specificity and substrate-induced inactivation of active-site and tunnel mutants of the beta-subunit. J. Biol. Chem. 266 (32), 21548-21557.

70. Schleicher, E. *et al.* (1976) Letter: Stereochemistry and mechanism of reactions catalyzed by tryptophanase and tryptophan synthetase. J. Am. Chem. Soc. 98 (4), 1043-1044.

71. Jhee, K.H. *et al.* (1998) Mutation of an active site residue of tryptophan synthase (beta-serine 377) alters cofactor chemistry. J. Biol. Chem. 273 (19), 11417-11422.

72. Jhee, K.H. *et al.* (1998) Tryptophan synthase mutations that alter cofactor chemistry lead to mechanism-based inactivation. Biochemistry 37 (41), 14591-14604.

73. Percudani, R. and Peracchi, A. (2003) A genomic overview of pyridoxal-phosphate-dependent enzymes. EMBO Rep. 4 (9), 850-854.

74. Schneider, G. *et al.* (2000) The manifold of vitamin B-6 dependent enzymes. Struct. Fold. Des. 8 (1), R1-R6.

75. Karp, P.D. *et al.* (2017) The BioCyc collection of microbial genomes and metabolic pathways. Brief. Bioinform.

76. Hodge-Hanson, K.M. *et al.* (2018) Expression of PLP-independent racemases can reduce 2-aminoacrylate stress in *Salmonella enterica*. J. Bacteriol. 200 (9), e00751-17.

77. Cotter, P.D. *et al.* (2005) Bacterial lantibiotics: strategies to improve therapeutic potential. Curr. Protein Pept. Sci. 6 (1), 61-75.

78. Cavera, V.L. *et al.* (2015) Bacteriocins and their position in the next wave of conventional antibiotics. Int. J. Antimicrob. Agents 46 (5), 494-501.

79. Brandange, S. and Lindblom, L. (1979) The enzyme "aldehyde oxidase" is an iminium oxidase. Reaction with nicotine delta 1'(5') iminium ion. Biochem. Biophys. Res. Commun. 91 (3), 991-996.

80. Isenberg, S. and Newman, E.B. (1974) Studies on L-serine deaminase in *Escherichia coli* K-12. J. Bacteriol. 118 (1), 53-58.

81. Fisher, H.F. *et al.* (1982) Glutamate dehydrogenase catalyzes the reduction of a Schiff base (delta 1-pyrroline-2-carboxylic acid) by NADPH. J. Biol. Chem. 257 (22), 13208-13210.

82. Bazurto, J.V. *et al.* (2017) Untargeted metabolomics confirms and extends the understanding of the impact of aminoimidazole carboxamide ribotide (AICAR) in the metabolic network of *Salmonella enterica*. Microb. Cell 5 (2), 74-87.

83. Percudani, R. and Peracchi, A. (2009) The B6 database: a tool for the description and classification of vitamin B6-dependent enzymatic activities and of the corresponding protein families. BMC Bioinformatics 10, 273.

84. Merigliano, C. *et al.* (2018) The relationship between vitamin B6, diabetes and cancer. Front. Genet. 9, 388.

TABLE 2.1. Characterized generators and targets of 2-aminoacrylate stress in *Salmonella enterica*

A

Endogenous 2AA generators	Gene Product	Stress[a]	Reference
Anabolic Ser/Thr dehydratase [EC: 4.3.1.19]	IlvA	Serine	[57]
Catabolic Ser/Thr dehydratase [EC: 4.3.1.19]	TdcB[b]		
Cysteine desulfhydratase [EC: 2.5.1.47]	CdsH	Cysteine	[15]
Diaminopropionate NH_4 lyase [EC: 4.3.1.15]	DAPAL	Diaminopropionate	[16]

B

Endogenous 2AA targets	Gene Product	Phenotype[c]	Reference
Serine hydroxymethyl transferase [EC: 2.1.2.1]	GlyA	Gly limitation in minimal medium	[40]
BCAA Aminotransferase [EC: 2.6.1.42]	IlvE	Ile requirement on pyruvate carbon source	[57]
Alanine racemase [EC: 5.1.1.1]	Alr	Ala is a poor nitrogen source	[50]
Alanine racemase [EC: 5.1.1.1]	DadX	Ala is a poor nitrogen source	[50]

[a]In the absence of RidA, if the listed metabolite is added to minimal glucose growth medium, the relevant generator enzyme uses it to produce 2AA and prevent cell growth.

[b]TdcB is only produced anaerobically in response to threonine and does not contribute to 2AA stress under standard aerobic growth conditions.

[c]2AA-dependent damage of the relevant enzyme is responsible for the listed growth phenotype in a *S. enterica ridA* mutant. Abbreviations: Gly:glycine, Ile:isoleucine, Ala:alanine

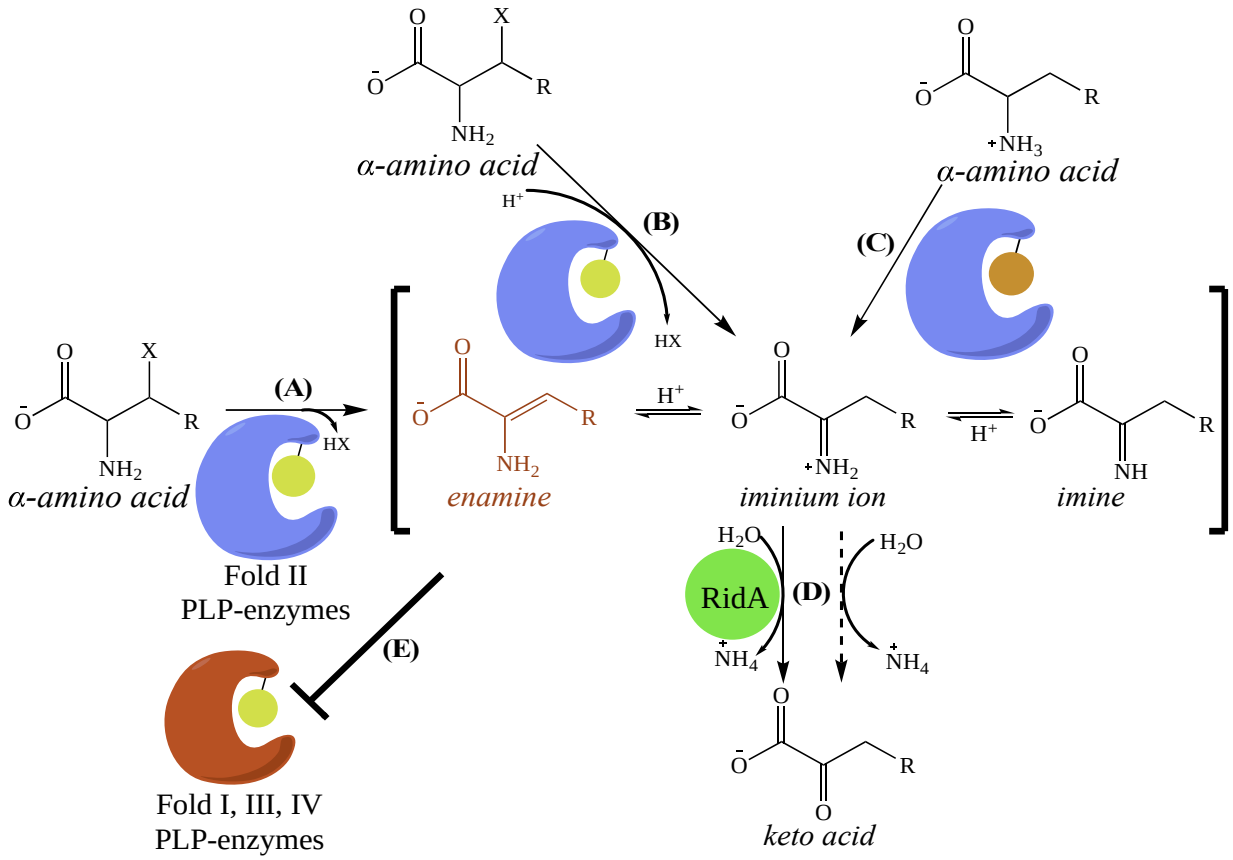


Figure 2.1. Enzymatic production of free enamines and imines. Various fold-type II PLP-dependent enzymes produce and release a reactive enamine intermediate following α,β-elimination of an amino acid precursor (**A**). The enamine is protonated and tautomerizes to form an iminium ion intermediate. A subset of PLP-dependent α,β-eliminases enzymatically facilitate iminium ion formation prior to its release from the active site (**B**). FAD- dependent enzymes can also produce an iminium ion directly (**C**). Equilibrium between enamine, imine, and iminium ion forms is largely dependent upon pH. *In vitro*, the iminium ion can react with free water, generating a stable keto acid product (**D**). RidA serves an important role in catalyzing the deamination of enamines and imines *in vivo*. Elimination of RidA leads to accumulation of free enamines, most notably 2-aminoacrylate (2AA) (R-group = H). Free 2AA can covalently modify and inactivate a number of PLP-dependent enzymes belonging to fold-type I, III, and IV (**E**). Enamine/imine generator and target enzymes are denoted using blue and red, respectively. Relevant coenzymes are denoted as bound circles, where PLP is colored yellow and FAD is orange.

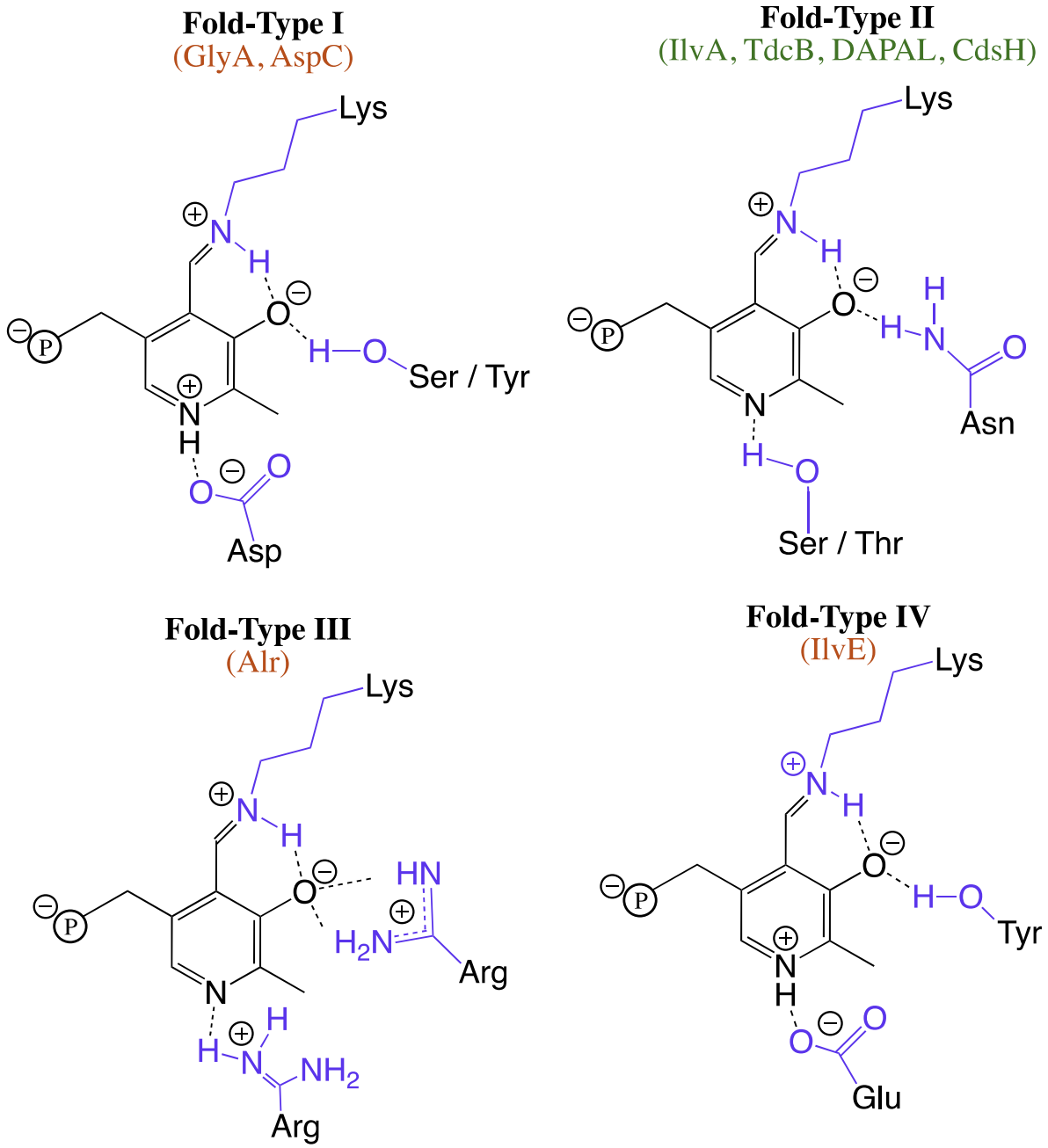


Figure 2.2. Generalized scheme showing coordination of pyridine N and 3'O from pyridoxal 5'-phosphate (black) for the four main fold types of PLP-dependent enzymes. Since fold types V-VII contain members belonging to only one reaction type, these were excluded from the scheme. Enzyme active site residues are shown in blue and the phosphate group of PLP is shown with '℗'. Beneath the fold type label, generators of free 2AA are highlighted in green and enzymes damaged by free 2AA are shown in red. Note, the identified residues are provided using the listed enzymes as templates; for example, fold type III decarboxylases coordinate the pyridine N with a glutamate residue replacing the listed arginine residue. A comprehensive list of PLP coordinating residues is provided by Singh et al. [49].

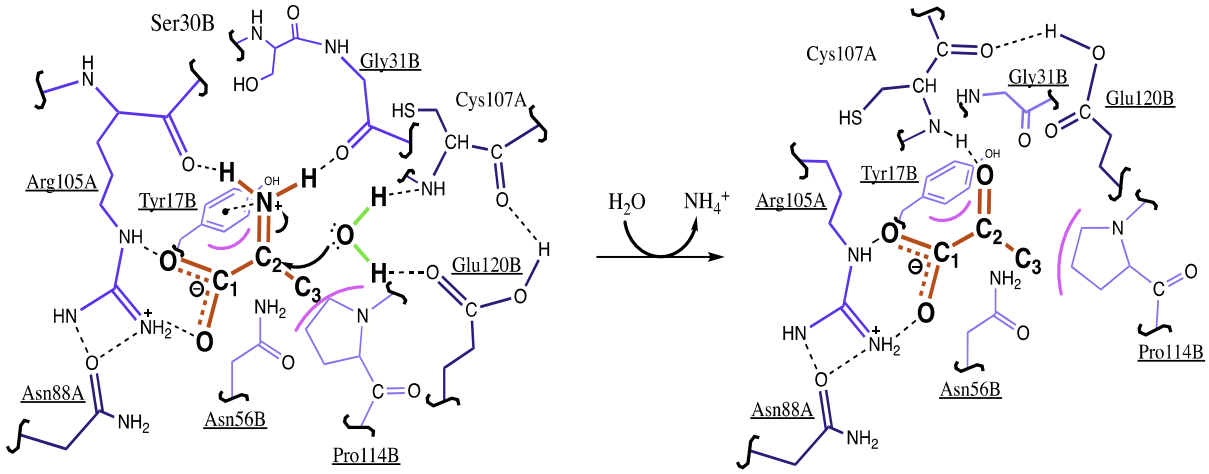


Figure 2.3. Proposed reaction mechanism for RidA-mediated hydrolysis of 2-iminiopropionate (2IP). A schematic for the proposed reaction mechanism of RidA using data extrapolated from the crystal structure of the RidA homolog, TdcF, with bound serine (2UYK) and 2-ketobutyrate (2UYN) [46]. The seven residues conserved among Rid family members are underlined. The substrate/product are shown in red, water is depicted in green, and RidA residues are shown in blue. The active site is at the interface of two subunits and residues from one subunit are labeled with an "A" while those from the other are labeled "B". Darker blue colors indicate residues in the foreground and paler blue colors are used for those in the background. The carboxyl group of 2IP forms a salt bridge with the R105 side-chain, while the protonated imine is stabilized through hydrogen bonding with the backbone carbonyl groups from R105 and G31 (left). The imine is also stabilized through a cation-pi interaction with the phenyl group from Y17. The secondary amine from the C107 backbone and the carbonyl group from the E120 side chain activate water for its nucleophilic attack on C_2 of the bound iminium ion. Following a rearrangement, pyruvate is produced in the active site, before it is released along with ammonium (right).

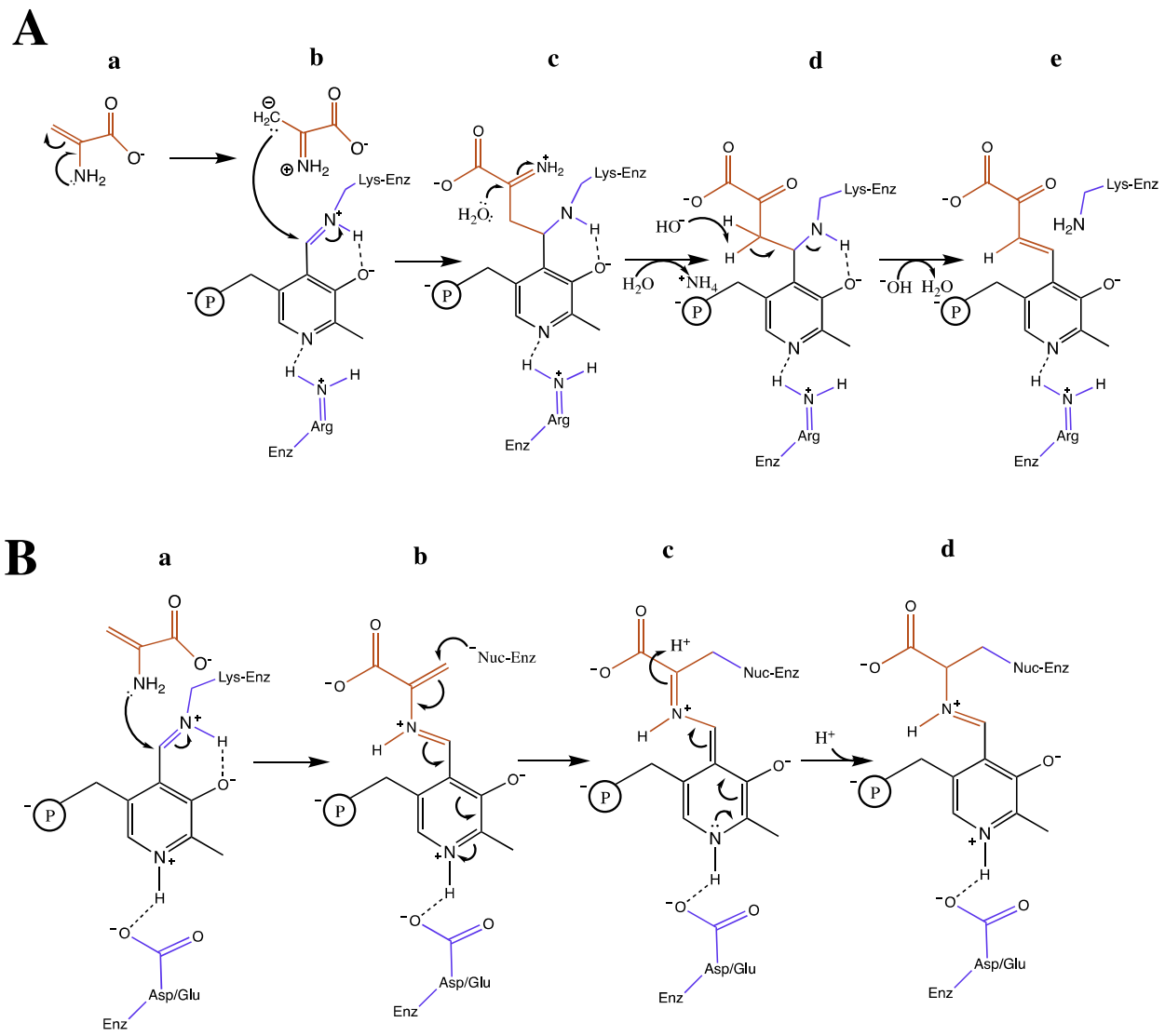


Figure 2.4. Free 2-aminoacrylate inhibits PLP-dependent enzymes by two mechanisms. 2AA is highlighted in red and enzyme active site residues are shown in blue. The phosphate group of PLP is shown with 'Ⓟ'. **(A)** A PLP/pyruvate adduct is formed following addition of 2AA via a C-C linkage. (a) Tautomeric rearrangement of 2AA generates a carbanion intermediate that (b) acts as a nucleophile to attack the 4'C of PLP. (c) Hydrolysis of the imine releases ammonium and creates a PLP/pyruvate adduct covalently bound to the active site lysine residue. (d) The adduct is susceptible to base-catalyzed formation of (e) a free PLP/pyruvate adduct. **(B)** 2AA/PLP covalently modifies a nucleophilic active site residue. (a) 2AA forms an external aldimine with PLP. (b) The alkene from 2AA participates in the conjugated π-bond system with PLP, strengthening the electrophilic nature of Cβ and allowing nucleophilic attack by an active site residue. (c) This forms a quinonoid intermediate, which can be rearranged and protonated to produce (d) a 2AA/PLP external aldimine covalently bound to a nucleophilic active site residue.

CHAPTER 3

THE RESPONSE TO 2-AMINOACRYLATE DIFFERS IN *ESCHERICHIA COLI* AND *SALMONELLA ENTERICA*, DESPITE SHARED METABOLIC COMPONTENTS[1]

[1]Borchert AJ, Downs DM. 2017. *Journal of Bacteriology* 199(14):e00140-17

3.1 ABSTRACT

The metabolic network of an organism includes the sum total of the biochemical reactions present. In microbes, this network has an impeccable ability to sense and respond to perturbations caused by internal or external stimuli. The metabolic potential (i.e, network structure) of an organism is often drawn from the genome sequence, based on the presence of enzymes deemed to indicate specific pathways. *Escherichia coli* and *Salmonella enterica* are members of the Enterobacteriaceae family of Gram-negative bacteria that share the majority of their metabolic components and regulatory machinery as the "core genome". In *S. enterica,* the ability of the enamine intermediate, 2-aminoacrylate (2AA) to inactivate a number of pyridoxal 5'-phosphate (PLP)-dependent enzymes has been established *in vivo*. In this study, 2AA metabolism and the consequences of its accumulation were investigated in *E. coli*. The data showed that despite the conservation of all relevant enzymes, *S. enterica* and *E.* coli differed in both the generation and detrimental consequences of 2AA. In total, the findings suggest that the structure of the metabolic network surrounding the generation and response to endogenous 2AA stress differs between *S. enterica* and *E. coli*.

Importance. This work compared the metabolic network surrounding the endogenous stressor, 2-aminoacrylate in two closely related members of the Enterobacteriaceae. The data showed that despite the conservation of all relevant enzymes in this metabolic node, the two closely related organisms diverged in their metabolic network structures. This work highlights how a set of conserved components can generate distinct network architectures and how this can impact the physiology of an organism. The work defines a model to expand our understanding of the 2-aminoacrylate stress response and the differences in metabolic structure and cellular milieu between *S. enterica* and *E. coli*.

3.2 BACKGROUND

The metabolic networks of microbes are characterized by robustness and redundancy that insure their ability of to absorb perturbations caused by internal and external stress [1-4]. Our ability to predict the metabolic architecture of an organism and how it will respond to perturbations is poorly developed. Despite the fact that metabolic components encoded by the genome can be defined, it is not yet clear how accurately metabolic capacity and robustness of an organism is predicted by the mere presence of these metabolic components. In some cases, the status of central and linear pathways that are predicted from genome annotation can be tested by simple growth phenotypes. For instance, lack of a single histidine biosynthetic enzyme (i.e., ATP phosphoribosyltransferase, HisG; EC 2.4.2.17), predicts there will be a requirement for exogenous histidine. In contrast, growth phenotypes resulting from differential regulation can prove more difficult to predict using genomic sequence [5]. Similarly, phenotypes caused by network perturbations involving flux diversion and/or pathway recruitment are not often evident through genome analysis. This later point is illustrated by the differences in thiamine biosynthesis that exist between *S. enterica* and *E. coli,* despite the similarity of all relevant enzyme components [6]. In addition to modulating the structure of single pathways, microbes often reconfigure their global metabolic networks to absorb stress and preserve metabolic function [7-9]. The capacity to modify network structure and the mechanism(s) used to do so introduces another layer of complexity that is difficult to define based on genome sequence. Historically, the phenotypic analysis of model organisms has been instrumental in identifying and dissecting these complexities.

Systems biology approaches to metabolism often depend on genome annotation to generate metabolic models of relevant organism(s) [10]. The models typically describe the metabolic potential of an organism by assuming that networks with shared component parts are configured

and respond to perturbations similarly. In this scenario, differences in metabolic potential indicate a different catalog of component parts. Examples of such metabolic divergence abound. For instance, the absence of the *lac* operon in *Salmonella enterica* relative to *Escherichia coli*, explains their distinct metabolic capacities in this area [11]. Organisms can also modulate the expression of conserved orthologs to generate distinct network architecture and accommodate differences in lifestyle [12]. In some instances, sequence analysis alone can predict these distinct expression patterns, based on knowledge of regulators and binding sites in combination with global analyses. Recent work outlining the distinct transcriptional responses *E. coli* and *S. enterica* produce during triclosan exposure provides an example of this scenario [13].

The study described here addressed the response of two organisms with similar metabolic components to the endogenously generated reactive metabolite 2-aminoacrylate (2AA). Pyridoxal 5'-phosphate (PLP)-dependent enzymes catalyzing β-elimination reactions proceed through reactive enamine intermediates, one of which is 2AA [14, 15]. A subset of 2AA generating enzymes release this reaction intermediate from the active site into the cellular milieu [16-18]. The RidA protein was demonstrated to be an enamine/imine deaminase, conserved in all domains of life. Work in *S. enterica* showed that in the absence of *ridA*, 2AA accumulates endogenously and inactivates enzymes by covalently modifying the PLP coordinated in the active site [16, 17, 19, 20]. In total, the resulting damage led to metabolic disruptions that were defined through biochemical and phenotypic analysis in *S. enterica.* The analysis revealed that the consequence of 2AA accumulation in *S. enterica* is partial inactivation of at least the PLP-dependent enzymes serine hydroxymethyltransferase (GlyA; EC 2.1.2.1), branched-chain amino acid aminotransferase (IlvE; EC 2.6.1.42) and alanine racemase (Alr/DadX; EC 5.1.1.1). A summary of the RidA paradigm of 2AA stress in *S. enterica* is depicted in Figure 1.

Collectively, the data from *S. enterica* showed that the primary source of free 2AA was serine/threonine dehydratase (IlvA, EC 4.3.1.19), and the critical target that resulted in growth defects was GlyA [19, 21]. The elimination of RidA homologs from organisms including *Lycopersicon esculentum, Arabidopsis thaliana, Zea mays,* and *Saccharomyces cerevisiae* has been linked to a range of metabolic and growth perturbations [22-25]. The growing list of metabolic defects and growth phenotypes associated with organisms lacking RidA function is consistent with an ancient stress system that is modulated by RidA proteins. Growth defects caused by the accumulation of 2AA reflect the response of the metabolic network to multiple perturbations and provide a means to understand metabolic network architecture. The enzymes of the 2AA stress paradigm in *S. enterica*, both the generators and targets, are present in *E. coli.* Despite the conservation of components, the results here showed the metabolic networks of the two organisms have dramatically different abilities to generate and respond to 2AA.

3.3 MATERIALS AND METHODS

Bacterial Strains, Media, and Chemicals. Strains used in this work are derivatives of *Salmonella enterica* serovar Typhimurium LT2 or *Escherichia coli* K12 unless otherwise noted, and are listed in Table 1. Minimal medium was no-carbon E (NCE) supplemented with 1 mM MgSO4 [26], trace minerals [27], and 11 mM D-glucose, 50 mM pyruvate, or 20mM glycerol as the sole carbon source. Luria-Bertani (LB) broth and Difco nutrient broth (NB) (8g/L) containing NaCL (5 g/L) was used as rich medium. Difco BiTek agar (15g/L) was added for solid medium. Antibiotics were used at the following concentrations for rich (or minimal) medium: kanamycin, 50 (12.5) μg/mL; chloramphenicol, 20 (5) μg/mL; and ampicillin 150 (7.5) μg/mL. All chemicals were obtained from Sigma (St. Louis, MO).

Genetic techniques. Transductional crosses in *E. coli* were performed with bacteriophage P1*vir* and used described protocols [6, 28]. In general, *E. coli* strains were constructed by transducing the appropriate mutations from the Keio collection into the appropriate strain background, followed by phenotypic confirmation [29]. The *tdcF* mutation was made using the λ-Red recombinase system described by Datsenko and Wanner [30] and transduced into relevant genetic backgrounds by selecting kanamycin resistance. When multiple mutations were present in the same genetic background, the kanamycin cassette was resolved using a previously described method [30].

Construction of chromosomal construct for inducing 2AA formation. The *ilvA219* allele, encoding feedback-resistant variant $IlvA_{L447F}$, was amplified from *S. enterica*, by PCR with Q5 High-fidelity DNA polymerase (New England Biolabs) using primers AB1 and AB2. The amplification product was purified, digested with *EcoR*I and *Pst*I (New England Biolabs) and ligated into predigested pBAD24 [31] to generate pAB1. The *cat* gene, encoding chloramphenicol acetyltransferase, including the promoter, was amplified from pSU18 [32] by PCR using Q5 High-fidelity DNA polymerase (New England Biolabs) and primers AB3 and AB4. The product was digested with *Hind*III (New England Biolabs) and ligated into predigested pAB1. The ligation mixture was transformed into *E. coli* DH5α and chloramphenicol resistant colonies were selected. Plasmids from the Cm^r clones were isolated, the insert confirmed by PCR and the resulting plasmid was pAB2. The P_{BAD}-*ilvA219 cat* insert from pAB2 was amplified using primers AB5/AB6 or AB5/AB7 and replaced the *araBAD* genes from *E. coli* K12 and *S. enterica* LT2 using the λ-Red recombinase system [30]. The chromosomal construct was confirmed using a combination of PCR amplification and phenotypic analysis. The latter demanded chloramphenicol resistant colonies

that were able to grow on minimal media containing arabinose, but not on minimal media lacking arabinose in a strain where the construct encoded the only source of IlvA in the cell.

Quantification of Growth. A 2 µL aliquot of cells from an overnight LB culture was used to inoculate 198 µL of growth medium. 96-well plates were incubated at 37°C in a microplate reader (model EL808; Bio-Tek Instruments) with shaking and growth was monitored as the change in optical density at 650 nm (OD_{650}) over time. Unless stated, growth experiments were performed with biological triplicates. Results were plotted using GraphPad Prism 6.0b to generate curves (log_{10}-format) represented as a composite of the averages and standard deviation of the replicates.

Threonine dehydratase (IlvA) assays. Overnight LB broth cultures were inoculated into 200 mL minimal medium (with stated additions) and grown at 37°C to late logarithmic phase (OD650 ~ 0.8) before being stored as pellets at -80°C until use. Cell pellets were re-suspended and pelleted twice, using 5mL of 50mM KPO_4 (pH 7.2), before being re-suspended in 5mL of 50mM KPO_4 (pH 7.2) with 0.4mM dithiothreitol. Cell extracts were prepared with a French pressure cell (1,500 psi) and clarified by centrifugation (1 hr @ 40,000 × *g* at 4°C). Threonine dehydratase assays were carried out as described previously [22, 33] and reported as nanomoles 2-ketobutyrate formed per milligram protein per minute. Protein concentrations were estimated by the method of Bradford [34]. Each experiment included three biological replicates and the average with standard deviation is reported.

Branched-chain amino acid aminotransferase (IlvE) assays. Three overnight LB cultures were used to inoculate (1:100) 5ml of minimal glycerol medium with indicated additions. Following growth, cells were pelleted and re-suspended twice using NCE medium, pelleted and frozen at -20 °C until use. Cell pellets were thawed and assayed as described [22]. Briefly, cells were permeabilized in 50 mM potassium phosphate, pH 8 containing 50 µM PLP, 10 mM 2-

ketoglutarate and 10 % PopCulture Reagent (Novagen). The reaction was initiated by addition of 20 mM L-Isoleucine. Product 2-ketomethylvalerate (2-KMV) was derivatized by 2,4-dinitrophenylhydrazine (DNPH) to yield a chromophore absorbing at 540nm. 2-KMV was quantified with a standard curve generated from known quantities of 2-KMV similarly derivatized with DNPH, and normalized to total protein content estimated by the method of Bradford [34]. The significance of IlvE activity differences were determined by analyzing the data using Two-way ANOVA within GraphPad Prism 6.0b and demanding a p-value less than 0.01.

Aspartate aminotransferase (AspC) assays. 100 mL cultures were grown in minimal glucose medium (with indicated additions) to early stationary phase (OD_{650}, ~1.2) and stored as cell pellets at −80°C. Cell pellets were re-suspended in 7 mL of 25 mM KPO_4 (pH 7.5), 0.1 mM EDTA, 0.2 mM pyridoxal 5′-phosphate, 0.2 mM dithiothreitol, and 5% glycerol. Cell extracts were prepared from the re-suspended cell pellets using a French pressure cell (1,500 psi) and clarified by centrifugation (1 hr @ 40,000 × *g* at 4°C). Aminotransferase activity was measured in the reverse direction using a coupled assay in which the oxaloacetate formed following aspartate addition was reduced by malic dehydrogenase to malate using NADH as an electron acceptor. NADH oxidation was monitored at 340 nm (ε_{340} = 6200 $M^{-1}cm^{-1}$) [35]. Protein concentrations were determined using the bicinchoninic acid (BCA) assay (Pierce). AspC activity is reported as nanomoles of NADH oxidized per milligram of protein per minute. The data are presented as the average from three independent experiments, with standard deviations. Statistical significance ($p < 0.01$) was analyzed by performing one-way ANOVA and Tukey's post-hoc test within GraphPad Prism 6.0b.

3.4 RESULTS

***S. enterica* and *E. coli* respond differently to the loss of *ridA*.** A number of nutritional and biochemical phenotypes caused by the accumulation of 2-aminoacrylate (2AA) have been described for mutants of *S. enterica* lacking RidA. None of these growth defects were manifest in an *E. coli* K12 strain lacking RidA. Besides RidA (b4243), *E. coli* contains two additional Rid superfamily members, which are not present in *S. enterica*, TdcF (b3113) and RutC (b1010). An *E. coli* strain lacking these three RidA family members, (*ridA tdcF rutC*) similarly failed to show any growth defects associated with 2AA accumulation (Figure 2). In addition to growth phenotypes, activity of branched-chain amino acid aminotransferase (IlvE, EC 2.6.1.42) is used as a proxy for 2AA accumulation *in vivo,* since 2AA inactivates this enzyme [17, 22, 36]. A *ridA* mutation decreased the activity of IlvE four-fold in *S. enterica* (73 ± 9 nmol 2KMV min^{-1} mg^{-1} protein vs. 18 ± 5 nmol 2KMV min^{-1} mg^{-1} protein, respectively; [36]), yet IlvE activity in an *E. coli* triple mutant (*ridA tdcF rutC*) was indistinguishable from the wild-type strain (180 ± 7 nmol 2KMV min^{-1} mg^{-1} protein vs. 179 ± 2 nmol 2KMV min^{-1} mg^{-1} protein) when grown in minimal glucose medium. Together these results showed there was a significant difference in 2AA metabolism between *S. enterica* and *E. coli.*

There are two simple possibilities for the phenotypic difference between *S. enterica* and *E. coli* lacking RidA proteins: differences in metabolic network architecture or different component composition. To distinguish between these possibilities, a hybrid genome was constructed. The *glyA* and *ilvE* loci in *S. enterica* were replaced with the respective ORFs from *E. coli* using linear recombination techniques [30]. *S. enterica* strains carrying each of the *E. coli* loci were generated and confirmed. A *ridA* mutation was then introduced into each of the strains, generating two pairs of strains isogenic at the *ridA* locus. Without exception, the *ridA* derivative strains had the 2AA-

associated phenotypes characterized for *S. enterica* [19, 37]. Specifically, they were unable to grow in the presence of serine or on pyruvate as a carbon source unless glycine or isoleucine were provided (data not shown). Furthermore, the activity of IlvE was assayed in the *S. enterica* strains carrying the *E. coli ilvE* locus. The strain lacking *ridA* had four-fold decrease in IlvE activity relative to wild-type (11 ± 1 vs. 47 ± 9 nmol 2KMV/min mg, respectively), as is typical of *ridA* mutants of *S. enterica* with a complete complement of *Salmonella* genes [36]. These data supported a role for metabolic networks, not component composition, in determining the phenotypic outcomes in each organism.

An inducible serine/threonine dehydratase system modulates 2-aminoacrylate production. A simple explanation for the lack of RidA-like phenotype in *E. coli* was that less 2AA was produced and/or free in the cell. A construct that allowed inducible production of 2AA was constructed to test this possibility. The *ilvA219* allele was cloned under the control of the *araBAD* promoter (P_{BAD}) and inserted in the chromosome replacing the *araBAD* locus in relevant strains of both *S. enterica* and *E. coli*. The *ilvA219* allele of *S. enterica* encodes a variant ($IlvA_{L447F}$) that is insensitive to inhibition by isoleucine [38]. This construct (*ABcc1*) allowed manipulation of both the timing and relative amount of 2AA production *in vivo*. Control experiments with *S. enterica* showed that without induction of the construct, strains lacking RidA had a minor growth defect characteristic of *ridA* mutants. This growth defect was eliminated when exogenous isoleucine was provided, due to the well-characterized allosteric inhibition of the native IlvA. As expected, induction of *ilvA219* expression resulted in a growth defect that persisted, despite the addition of isoleucine (Figure 3). Importantly, glycine restored full growth to the *ridA* strain inducing expression of *ilvA219*, confirming that growth limitation was caused by increased accumulation of 2AA (Fig 5A) [19].

The increase in threonine dehydratase activity caused by inducing its expression from the *ABcc1* in *E. coli* was quantified. The addition of 0.2% (w/v) arabinose to the growth medium of DM14931 (*ABcc1*) increased the IlvA enzyme activity in crude extract more than five-fold (1071 ± 10 vs. 184 ± 57 nmol 2KB min-1 mg-1, respectively). Additionally, when the cells were grown without arabinose, 100% of the activity was sensitive to isoleucine, consistent with the source of activity being native IlvA. In contrast, with arabinose in the growth medium, greater than 75% of the activity was retained in the presence of isoleucine, consistent with the majority of the activity being from the $IlvA_{L447F}$ variant encoded in *ABcc1*. These data confirmed that induction of *ilvA219* increased the serine/threonine dehydratase activity in *E. coli* and allowed the conclusion that 2AA was similarly increased.

RidA or TdcF is sufficient to protect *E. coli* from 2AA damage. *E. coli* strains carrying the *ABcc1* construct and either lacking the three RidA homologs (*ridA tdcF rutC*), or wild-type at these loci, were assessed for growth phenotypes in minimal glycerol medium. In all cases, the expression of *ilvA219* was induced with arabinose. The data in Figure 4 showed that induction of *ilvA219*, and the resulting increase in 2AA, compromised growth of DM14994 (*ridA tdcF rutC*). Exogenous addition of serine (5 mM) further compromised growth (Fig. 4), while addition of L-threonine (1 mM) restored growth to that of the un-induced control (data not shown). This growth pattern was reminiscent of a *ridA* mutant of *S. enterica* [39, 40], and suggested that when sufficient 2AA was present, *E. coli* was susceptible to growth inhibition, as was *S. enterica.*

Strains containing the inducible *ilvA219* allele (*ABcc1*) and all combinations of *ridA*, *tdcF*, and *rutC* single and double gene deletions were generated and assessed in glycerol medium. Strain DM15055, the *ridA tdcF* double mutant, had a growth defect similar to that of strain DM14944, the *ridA tdcF rutC* triple mutant (Fig 4). Neither the presence of the single mutations in *ridA*, *rutC,*

or *tdcF*, nor double mutations *ridA rutC* or *tdcF rutC*, had any impact on the growth of the strain with the *ABcc1* construct (data not shown). Expression of either *ridA* or *tdcF in trans* eliminated the growth defect of the *ridA tdcF* mutant DM15055 (data not shown). These data allowed the conclusion that RidA and TdcF were at least partially redundant under the growth conditions tested.

Branched-chain amino acid aminotransferase (IlvE) activity was measured in strains with the *ABcc1* construct as a quantitative proxy for the accumulation of 2AA. When the cells reached mid-logarithmic growth (OD_{650} ~0.6), L-arabinose was added and the cells were grown an additional hour before IlvE activity was assayed. The data showed that IlvE activity in the *ridA*, *tdcF,* and wild-type strains was not significantly different ($p < 0.05$) (189 ± 8, 183 ± 20 or 212 ± 1 nmol 2KMV min-1 mg-1, respectively). However, in the *ridA tdcF* double mutant, aminotransferase activity was decreased by ~35% (140 ±18 vs 212 ± 1, nmol 2KMV min-1 mg-1 respectively). In the absence of arabinose, this decrease in activity did not occur (data not shown). Consistent with the growth results in Figure 4, the assay results indicated 2AA had metabolic consequences in *E. coli* only when both RidA and TdcF were absent. The lack of 2AA accumulation in the absence of the inducible *ilvA219* construct suggested *E. coli* might naturally have lower serine/threonine dehydratase activity than *S. enterica*. However, when wild-type strains of *S. enterica* and *E. coli* were grown in minimal glucose medium, serine/threonine dehydratase activity measured in crude extract was not significantly different (300 ± 70 vs. 290 ± 39 nmol 2-ketobutyrate/min/mg, respectively). Based on these data, a more complex explanation for the lack of 2AA damage in *E. coli tdcF ridA* strains was warranted.

2AA generates distinct phenotypic signatures in *S. enterica* and *E. coli*. To better understand the metabolic effect of 2AA accumulation on *E. coli*, metabolites were screened for

the ability to rescue growth of DM15055 (*ABcc1 ridA tdcF*). The data in Figure 5 summarize the results. In contrast to the results with *S. enterica ridA* mutants [19], glycine had no effect on the growth defect caused by 2AA in the *E. coli ABcc1 ridA tdcF* strain. Exogenous aspartate (1mM) restored growth of the *E. coli* strain, but had no effect on growth of the *S. enterica ridA* strain.

Aspartate aminotransferase (AspC; EC 2.6.1.1) is a PLP-dependent enzyme that catalyzes the formation of aspartate from oxaloacetate. AspC shares characteristics of the three PLP-enzymes that are targeted by free 2AA in *S. enterica*, GlyA, IlvE, and Alr [21, 36, 41]. Specifically, each of these four enzymes have reduced specific activity when challenged *in vitro* with 3-chloroalanine (3CA) [42-44], due to generation of 2AA in the active site which covalently modifies the PLP cofactor in the active site and renders the enzyme inactive [45]. The effect of aspartate on the growth phenotype of the *ridA tdcF* strain expressing *ilvA219* suggested the cells could be limited for aspartate. AspC activity was measured in crude extracts of DM15035 (*S. enterica; ABcc1*), DM15036 (*S. enterica; ridA ABcc1*), DM15055 (*E. coli; ridA tdcF ABcc1*) and DM14931 (*E. coli; ABcc1*) (Fig 6). AspC activity was not statistically different ($p < 0.01$) between the *S. enterica ridA* mutant and wild-type. However, when $IlvA_{L447F}$ expression was induced with arabinose in the growth medium, AspC activity in DM15036 (*ridA ABcc1*) was significantly decreased compared to the isogenic wild-type strain (327 ± 21 vs. 493 ± 25 nmol NAD^+ min^{-1} mg^{-1}, respectively) (Fig 6A). A similar result was observed with *E. coli* strains DM14931 (*ABcc1*) and DM15055 (*ridA tdcF ABcc1*) where the AspC activity in the two strains was indistinguishable in the absence of arabinose. As with the *S. enterica* strains, when arabinose was present in the growth medium the *ridA tdcF* mutant strain (DM15055) had significantly less AspC activity than the wild-type parental strain (DM14931) (877 ± 43 vs. 1385 ± 40, nmol NAD^+ min^{-1} mg^{-1}respectively) (Fig 6B).

Collectively, the above data indicated that accumulation 2AA resulted in damage to AspC *in vivo*. Two results suggested that low AspC activity was not responsible for the reduced growth of an *E. coli* strain lacking RidA and TdcF. First, addition of tyrosine did not exacerbate this growth defect (data not shown). The products of both *aspC* and *tyrB* catalyze the conversion of oxaloacetate to aspartate, and thus aspartate aminotransferase activity would be further reduced by tyrosine due to allosteric inhibition of TyrB [46]. Second, while *aspC* expressed *in trans* restored growth of an *E. coli aspC* mutant, it failed to restore growth of the *ridA tdcF* mutant (data not shown).

Aspartate and adenine ameliorate the consequences, not the source of 2AA stress. It was found serendipitously that exogenous addition of purines adenine (0.4 mM), guanine (0.133 mM), or hypoxanthine (0.4 mM) restored growth to an *E. coli ridA tdcF ABcc1* mutant strain under 2AA stress (Figure 7; data not shown). Addition of pyrimidines failed to rescue growth (data not shown). The ability of adenine, and other purines, to restore growth was dependent on a functional PurR (Fig7). Significantly, a *purR* mutation had no effect on the growth stimulation by aspartate (data not shown). PurR is the main transcriptional regulator involved in *de novo* purine synthesis, and also participates in regulation of pyrimidine biosynthesis and one-carbon metabolism [47]. The data are consistent with a general model in which a gene in the PurR regulon interferes with growth of the *ridA tdcF ABcc1* mutant strain by a direct or indirect mechanism. As with aspartate, exogenous purines did not affect growth of the S. *enterica ridA* mutant a further indication of the difference in the 2AA metabolic network between the two organisms (data not shown).

In principle, restored growth of strains compromised by 2AA stress can reflect decreased production of 2AA or a metabolic restructuring that ameliorates the consequences of 2AA damage. IlvE activity was used as a proxy for 2AA accumulation to distinguish these possibilities. IlvE

activity was measured in strains DM14931 (*ABcc1)* and DM15055 *(ABcc1 ridA tdcF)* after growth in media containing 0.2% arabinose alone, or with the additions threonine, aspartate, or adenine. The data are shown in Table 3. As expected, a strain lacking *ridA* and *tdcF* had a ~33% reduction in IlvE activity when grown in minimal medium with arabinose. Addition of threonine, aspartate, or adenine could restore growth, but had disparate effects on the activity of IlvE. In the case of threonine, IlvE activity was restored. As characterized in *S. enterica,* this effect is due to threonine competing with serine for the active site of IlvA and reducing 2AA production. Significantly, IlvE activity was not increased by the presence of either aspartate or adenine in the growth medium. If anything, these additions appeared to increase the discrepancy between the strains with and without RidA activity. These data indicated that the latter two supplements were restoring growth by circumventing the consequences of 2AA damage, not by reducing the damage. In total, these data suggested that purines (via PurR) and aspartate supported metabolic architectures that allow growth in the presence of 2AA.

3.5 DISCUSSION.

A comparison of the *S. enterica* and *E. coli* K12 genomes showed that ~80% of protein-coding genes in *S. enterica* have a highly conserved (>90% amino acid identity) orthologous partner in *E. coli* and defined their shared core genome [48, 49]. Based on this genomic similarity, and supported by numerous studies over decades, these organisms were presumed to share significant metabolic potential and organization. Recent phenotypic studies have highlighted the need to consider this assumption more carefully, potentially on a case-by-case basis [5, 6, 12, 50-52]. This study was initiated to evaluate the conservation of metabolic network structure surrounding 2AA metabolism (Fig 1) between *S. enterica* and *E. coli.* Importantly, each of the enzymatic components in the 2AA metabolic network in *S. enterica* had an orthologous partner in

E. coli that shared >85% identity at the amino acid level (Table 4). The paradigm defined in *S. enterica* that articulated the causes and consequences of 2AA stress provided a means to assess the conservation of a higher order metabolic network architecture that was impacted by and responded to stress generated by endogenous metabolites.

The results of this study showed that *S. enterica* and *E. coli* have distinct metabolic networks associated with 2AA, despite sharing component enzymes. Specifically, the networks differ in, i) generation and/or accumulation of 2AA, and ii) the consequence that 2AA accumulation has on the function of cellular metabolism. The data showed that an *E. coli* strain, made to resemble an *S. enterica ridA* mutant through elimination of RidA, TdcF, and RutC, had no growth defects that paralleled those defined in the *S. enterica ridA* mutant. This result allowed the conclusion that the native network of *E. coli* did not normally generate damaging levels of 2AA, even in the presence of the 2AA precursor serine. However, when the *E. coli* strain was manipulated to increase 2AA production 5-fold, a significant growth defect and damage to isoleucine transaminase (IlvE) were generated. Despite being qualitatively similar, the growth defects in *S. enterica* and *E. coli* were shown to be mechanistically distinct.

Significantly, generation of a 2AA-dependent growth defect in *E. coli* required not only increased IlvA activity, but the absence of both RidA and TdcF. The latter result showed the role of these two proteins in the cell is at least partially redundant. A TdcF-dependent phenotype was unexpected since the *tdc* operon was shown to be transcriptionally silent in the growth conditions used [53, 54]. This result suggests there may be an unappreciated role for this operon in the cell.

There are several possible explanations for the difference in free 2AA generated by the native metabolic networks of *E. coli* and *S. enterica.* Like many PLP-dependent enzymes, IlvA is promiscuous and can use either threonine or serine as a substrate. The former is the "true"

substrate, and leads to isoleucine synthesis with aminocrotonate as the reaction intermediate. Aminocrotonate is not as reactive as the serine derived 2AA and has not been shown to be detrimental *in vivo* or *in vitro*. In contrast, 2AA can damage at least 4 cellular enzymes *in vivo* in the absence of RidA [21, 36, 41]. The IlvA proteins from *E.* coli and *S. enterica* are 95.5% identical, differing by 23/514 amino acids, making it unlikely that the specificity of one is dramatically different than the other. Rather we favor a scenario where a property of the cellular milieu differs between the organisms. Possibilities that could account for different generation/accumulation of 2AA include the ratio of serine/threonine, increased internal pH leading to a longer half-life of 2AA, among others. For instance, increased relative levels of threonine in *E. coli* could minimize occupation of the active site by serine, leading to less 2AA production. Such a possibility is consistent with previous work by Bazurto *et al*. These authors showed that when comparing *S. enterica* and *E. coli*, the latter had a metabolic capability that suggested an increased level of aminocrotonate in *E. coli* compared to *S. enterica* [6].

The distinct growth response of *E. coli* and *S. enterica* to 2AA accumulation raises additional provocative questions about the cellular environment and how metabolic effects are modulated. The enzymes targeted by 2AA are >90% identical between the organisms and their function *in vitro* is often indistinguishable. Interestingly, even when the impact on these enzymes was made equal, the consequences on the growth and function of the metabolic network remained different. Using IlvE as a proxy for 2AA, the general level of 2AA can be equated in the two organisms, however the metabolic additions that overcome the growth defect were distinct.

Past work showed that the critical target of 2AA for causing growth defects in *S. enterica* was serine hydroxymethyltransferase (GlyA) [19, 21], while the work here showed GlyA is not the critical 2AA target in *E. coli*. Furthermore, the data indicate AspC from both organisms can

be damaged by 2AA. In a catalog of PLP-dependent enzymes, AspC and GlyA were the only EC-classified PLP-dependent enzymes found in the genomes of all free-living organisms [55]. Rid family members are also conserved across all domains of life [56]. Collectively, the biochemical results herein from *S. enterica* and *E. coli* showed the most widely conserved EC-classified activities (GlyA and AspC) could both serve as important targets of free 2AA, consistent with the idea that RidA family members are part of an ancient stress response system.

In addition to the initial description of the RidA stress response in *E. coli,* this work highlighted the subtleties involved in defining the metabolic network based on the presence of component enzymes. Data herein suggests that the 2AA stress system provides a model to explore the causes and consequences of subtle differences in the cellular environment. It also provides a system that is amenable to mathematical modeling to define the impact small changes in substrate concentration and/or enzyme kinetics can have on a defined complex system. Ultimately this work contributes to our ability to access the next level of information from genome sequence and predict network structure and flexibility.

Acknowledgements. The authors thank Jannell Bazurto for helpful discussions. This work was supported by the competitive grants program at the NIH (GM095837) to DMD.

3.6 REFERENCES

1. Koenigsknecht, M.J. and Downs, D.M. (2010) Thiamine biosynthesis can be used to dissect metabolic integration. Trends Microbiol. 18 (6), 240-247.

2. Koenigsknecht, M.J. *et al.* (2012) Perturbations in histidine biosynthesis uncover robustness in the metabolic network of *Salmonella enterica*. PLoS ONE 7 (10), e48207.

3. Albert, R. *et al.* (2000) Error and attack tolerance of complex networks. Nature 406, 378-382.

4. Bazurto, J.V. and Downs, D.M. (2011) Plasticity in the purine-thiamine metabolic network of *Salmonella*. Genetics 187 (2), 623-631.

5. Winfield, M.D. and Groisman, E.A. (2003) Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl. Environ. Microbiol. 69 (7), 3687-3694.

6. Bazurto, J.V. *et al.* (2016) An unexpected route to an essential cofactor: *Escherichia coli* relies on threonine for thiamine biosynthesis. mBio 7 (1), e01840-15.

7. Ishii, N. *et al.* (2007) Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. Science 316 (5824), 593-597.

8. Ibarra, R. *et al.* (2002) *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. Nature 420 (6912), 186-189.

9. Hartman, J.L. *et al.* (2001) Principles for the buffering of genetic variation. Science 291 (5506), 1001-1004.

10. AbuOun, M. *et al.* (2009) Genome scale reconstruction of a *Salmonella* metabolic model: comparison of similarity and differences with a commensal *Escherichia coli* strain. J. Biol. Chem. 284 (43), 29480-29488.

11. Ewing, W.H. (1973) Differentiation of Enterobacteriaceae by biochemical reactions. Center for Disease Control.

12. Meysman, P. *et al.* (2013) Expression divergence between *Escherichia coli* and *Salmonella enterica* serovar Typhimurium reflects their lifestyles. Mol. Biol. Evol. 30 (6), 1302-1314.

13. Bailey, A.M. *et al.* (2009) Exposure of *Escherichia coli* and *Salmonella enterica* serovar Typhimurium to triclosan induces a species-specific response, including drug detoxification. J. Antimicrob. Chemother. 64 (5), 973-985.

14. Chargaff, E. and Sprinson, D.B. (1943) Studies on the mechanism of deamination of serine and threonine in biological systems. J. Biol. Chem. 151 (1), 273-280.

15. Phillips, A.T. and Wood, W.A. (1965) The mechanism of action of 5'-adenylic acid-activated threonine dehydrase. J. Biol. Chem 240 (240), 4703-4709.

16. Lambrecht, J.A. *et al.* (2012) Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. J. Biol. Chem. 287 (5), 3454-3461.

17. Ernst, D.C. *et al.* (2014) Endogenous synthesis of 2-aminoacrylate contributes to cysteine sensitivity in *Salmonella enterica*. J. Bacteriol. 196 (18), 3335-3342.

18. Ernst, D.C. *et al.* (2016) L-2,3-diaminopropionate generates diverse metabolic stresses in *Salmonella enterica*: targets of diaminopropionate in *Salmonella*. Mol. Microbiol. 101, 210-223.

19. Ernst, D.C. and Downs, D.M. (2016) 2-aminoacrylate stress induces a context-dependent glycine requirement in *ridA* strains of *Salmonella enterica*. J. Bacteriol. 198 (3), 536-543.

20. Walsh, C. (1982) Suicide substrates: mechanism-based enzyme inactivators. Tetrahedron 38 (7), 871-909.

21. Flynn, J.M. *et al.* (2013) Decreased coenzyme A levels in *ridA* mutant strains of *Salmonella enterica* result from inactivated serine hydroxymethyltransferase: *ridA* mutants are deficient in one carbon metabolism. Mol. Microbiol. 89 (4), 751-759.

22. Schmitz, G. and Downs, D.M. (2004) Reduced transaminase B (IlvE) activity caused by the lack of *yjgF* is dependent on the status of threonine deaminase (IlvA) in *Salmonella enterica* Serovar Typhimurium. J. Bacteriol. 186 (3), 803-810.

23. Leitner-Dagan, Y. *et al.* (2006) CHRD, a plant member of the evolutionarily conserved YjgF family, influences photosynthesis and chromoplastogenesis. Planta 225 (1), 89-102.

24. Niehaus, T.D. *et al.* (2014) *Arabidopsis* and maize RidA proteins preempt reactive enamine/imine damage to branched-chain amino acid biosynthesis in plastids. Plant Cell 26 (7), 3010-3022.

25. Kim, J.-M. *et al.* (2001) A member of the YER057c/yjgf/Uk114 family links isoleucine biosynthesis and intact mitochondria maintenance in Saccharomyces cerevisiae. Genes Cells 6 (6), 507-517.

26. Vogel, H.J. and Bonner, D.M. (1956) Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. 218 (1), 97-106.

27. Balch, W.E. and Wolfe, R.S. (1976) New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. Appl. Environ. Microbiol. 32, 781-791.

28. Willetts, N.S. *et al.* (1969) Genetic location of certain mutations conferring recombination deficiency in *Escherichia coli*. J. Bacteriol. 97 (1), 244-249.

29. Baba, T. *et al.* (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2.

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

31. Guzman, L.-M. *et al.* (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. 177 (14), 4121-4130.

32. Bartolomé, B. *et al.* (1991) Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. Gene 102 (1), 75-8.

33. Burns, R.O. (1971) L-threonine deaminase — biosynthetic (*Salmonella typhimurium*). Methods Enzymol. 17B, 555-560.

34. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-54.

35. Arnold, P.M. and Parslow, G.R. (1995) Designing a coupled assay system for aspartate aminotransferase. Biochem. Educ. 23 (1), 40-41.

36. Lambrecht, J.A. *et al.* (2013) RidA proteins prevent metabolic damage inflicted by PLP-dependent dehydratases in all domains of life. mBio 4 (1), e00033-13-e00033-13.

37. Christopherson, M.R. *et al.* (2008) YjgF is required for isoleucine biosynthesis when *Salmonella enterica* is grown on pyruvate medium. J. Bacteriol. 190 (8), 3057-3062.

38. LaRossa, R.A. and Van Dyk, T.K. (1987) Metabolic mayhem caused by 2-ketoacid imbalances. Bioessays 7, 125-130.

39. Enos-Berlage, J.L. *et al.* (1998) Complex metabolic phenotypes caused by a mutation in *yjgF*, encoding a member of the highly conserved YER057c/YjgF family of proteins. J. Bacteriol. 180 (24), 6519-6528.

40. Christopherson, M.R. *et al.* (2012) Suppressor analyses identify threonine as a modulator of *ridA* mutant phenotypes in *Salmonella enterica*. PLoS ONE 7 (8), e43082.

41. Flynn, J.M. and Downs, D.M. (2013) In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5'-phosphate cofactor. J. Bacteriol. 195 (16), 3603-3609.

42. Rando, R.R. (1974) Irreversible inhibition of aspartate aminotransferase by 2-amino-3-butenoic acid. Biochemistry 13 (19), 3859-3863.

43. Silverman, R.B. and Abeles, R.H. (1976) Inactivation of pyridoxal phosphate dependent enzymes by mono-and polyhaloalanines. Biochemistry 15 (21), 4718-4723.

44. Ueno, H. *et al.* (1982) Chemistry of the inactivation of cytosolic aspartate aminotransferase by serine O-sulfate. Biochemistry 21 (18), 4387-4393.

45. Walsh, C.T. (1984) Suicide substrates, mechanism-based enzyme inactivators: recent developments. Annu. Rev. Biochem. 53 (1), 493-535.

46. Van Dyk, T.K. and LaRossa, R.A. (1986) Sensitivity of a *Salmonella* typhimurium *aspC* mutant to sulfometuron methyl, a potent inhibitor of acetolactate synthase II. J. Bacteriol. 165 (2), 386-392.

47. Cho, B.K. *et al.* (2011) The PurR regulon in *Escherichia coli* K-12 MG1655. Nucleic Acids Res. 39 (15), 6456-6464.

48. Anjum, M.F. *et al.* (2005) Identification of core and variable components of the *Salmonella enterica s*ubspecies I genome by microarray. Infect. Immun. 73 (12), 7894-7905.

49. Parkhill, J. *et al.* (2001) Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature 413 (6858), 848-852.

50. Dekel, E. and Alon, U. (2005) Optimality and evolutionary tuning of the expression level of a protein. Nature 436 (7050), 588-592.

51. Dekel, E. *et al.* (2005) Environmental selection of the feed-forward loop circuit in gene-regulation networks. Phys. Biol. 2 (2), 81-88.

52. Quan, S. *et al.* (2012) Adaptive evolution of the lactose utilization network in experimentally evolved populations of *Escherichia coli*. PLoS Genet. 8 (1), e1002444.

53. Sawers, G. (1998) The anaerobic degradation of L-serine and L-threonine in enterobacteria: networks of pathways and regulatory signals. Arch. Microbiol. 171 (1), 1-5.

54. Chattopadhyay, S. *et al.* (1997) Involvement of Fnr and ArcA in anaerobic expression of the *tdc* operon of *Escherichia coli*. J. Bacteriol. 179 (15), 4868-4873.

55. Percudani, R. and Peracchi, A. (2003) A genomic overview of pyridoxal-phosphate-dependent enzymes. EMBO reports 4 (9), 850-854.

56. Niehaus, T.D. *et al.* (2015) Genomic and experimental evidence for multiple metabolic functions in the RidA/YjgF/YER057c/UK114 (Rid) protein family. BMC Genomics 16 (1).

57. Castilho, B.A. *et al.* (1984) Plasmid insertion mutagenesis and *lac* gene fusion with mini-mu bacteriophage transposons. J. Bacteriol. 158 (2), 488-495.

58. Kanehisa, M. and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28 (1), 27-30.

59. Sievers, F. *et al.* (2014) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7 (1), 539-539.

TABLE 3.1. Bacterial strains and plasmids

Strain	Genotype	Source
DM14520	*E. coli* wild-type (K12 MG1655 $F^-\lambda^-$ rph-1)	Lab collection
DM14581	*E. coli ilvA723*::*kan*	This study
DM14597	*E. coli ridA790*::*kan*	This study
DM14601	*E. coli* DH5α/pAB1	This study
DM14705	*E. coli* Δ*ridA890 tdcF13*::*cat rutC754*::*kan*	This study
DM14712	*E. coli* DH5α/pAB2	This study
DM14930	*E. coli ilvA723*::*kan ara*::P_{BAD}-*ilvA219 cat*[b]	This study
DM14931	*E. coli ara*::P_{BAD}-*ilvA219 cat*	This study
DM14949	*E. coli* Δ*ridA890 ara*::P_{BAD}-*ilvA219 cat*	This study
DM14994	*E. coli* Δ*ridA890* Δ*tdcF14* Δ*rutC854 ara*::P_{BAD}-*ilvA219 cat*	This study
DM15049	*E. coli rutC754*::*kan ara*::P_{BAD}-*ilvA219 cat*	This study
DM15050	*E. coli* Δ*tdcF14* Δ*rutC854 ara*::P_{BAD}-*ilvA219 cat*	This study
DM15055	*E. coli* Δ*ridA890* Δ*tdcF14 ara*::P_{BAD}-*ilvA219 cat*	This study
DM15077	*E. coli tdcF12*::*kan ara*::P_{BAD}-*ilvA219 cat*	This study
DM15078	*E. coli* ΔridA890 rutC754::kan *ara*::P_{BAD}-*ilvA219 cat*	This study
DM15832	*E. coli* Δ*ridA890* Δ*tdcF14 ara*::P_{BAD}-*ilvA219 cat purR746*::*kan*	This study
DM9404	*S. enterica* LT2 (wild-type)	Lab collection
DM15035	*S. enterica ara*::P_{BAD}-*ilvA219 ca*t	This study
DM3480	*S. enterica ridA3*::MudJ[a]	Lab collection
DM15036	*S. enterica ridA3*::MudJ *ara*::P_{BAD}-*ilvA219 cat*	This study

Plasmid	Description	Source
pAB1	pBAD24-IlvA^{L447F}	This study
pAB2	pBAD24-IlvA^{L447F} Cat	This study

[a]MudJ refers to the MudJ1734 transposon [57]. Plasmids constructs used pBAD24 [31] and are described in Materials and Methods.

[b] For simplicity this construct is designated *ABcc1* in the text.

TABLE 3.2. Primers

Primer	Sequence
AB1	gagagaattcatggcggaatctcaacctct
AB2	gagactgcagttaacccgccagaaag
AB3	tagaagcttgatcggcacgtaagagg
AB4	tagaagcttacgccccgccctgcca
AB5	agccatgacaaaaacgcgtaac
AB6	agcctggtttcgtttgattggctgtggttttatacagtcaccgccaaaacagccaagct
AB7	ttcatcaacgcgcccccatgggacgtttttagaggcaccgccaaaacagccaagct

TABLE 3.3. Aspartate and adenine restore growth without reducing 2AA accumulation.

Supplement	Strain	Relevant genotype*	IlvE Activity (nmol 2KMV min^{-1} mg^{-1})	Rid^-/Rid^+
None	DM14931	wild-type	148 ± 23	0.70
	DM15055	*ridA tdcF*	102 ± 5	
Threonine	DM14931	wild-type	279 ± 26	0.96
	DM15055	*ridA tdcF*	269 ± 13	
Aspartate	DM14931	wild-type	174 ± 18	0.63
	DM15055	*ridA tdcF*	109 ± 2	
Adenine	DM14931	wild-type	183± 18	0.54
	DM15055	*ridA tdcF*	98 ±8	

* All strains include the chromosomal construct *ABcc1* that have the *ilvA219* allele under control of the arabinose inducible pBAD promoter.

The indicated strains were grown in minimal glycerol medium, containing indicated supplements at 1 mM concentration, to an OD_{650} ~ 0.6. Arabinose was added to a final concentration of 0.2% and cultures were grown for an hour before pelleting cells. Branched-chain amino acid aminotransferase activity (IlvE) was measured in crude extracts. Data represent the mean and standard deviation of results from experiments performed in biological triplicate measuring the rate of formation of 2-ketomethylvalerate (2-KMV). The Rid^-/Rid^+ ratio reflects the ratio of IlvE activity in stains without RidA and TdcF to the wild-type parent and is a measure of 2AA accumulation.

TABLE 3.4. The 2AA subnetwork components are conserved between *S. enterica* and *E. coli*

Function	Protein	% Amino Acid Identity*
2AA Generator	IlvA	95.5
2AA Quencher	RidA	93.8
RidA Homologs (absent in *S. enterica*)	TdcF	73.5 (to *S. enterica* RidA)
	RutC	28.8 (to *S. enterica* RidA)
2AA Targets	IlvE	97.1
	GlyA	93.3
	Alr	91.4
	DadX	85.7

*Amino Acid sequences obtained from KEGG [58], protein alignments and % Amino Acid Identity calculations using EMBL-EBI Clustal Omega online analysis tool [59]

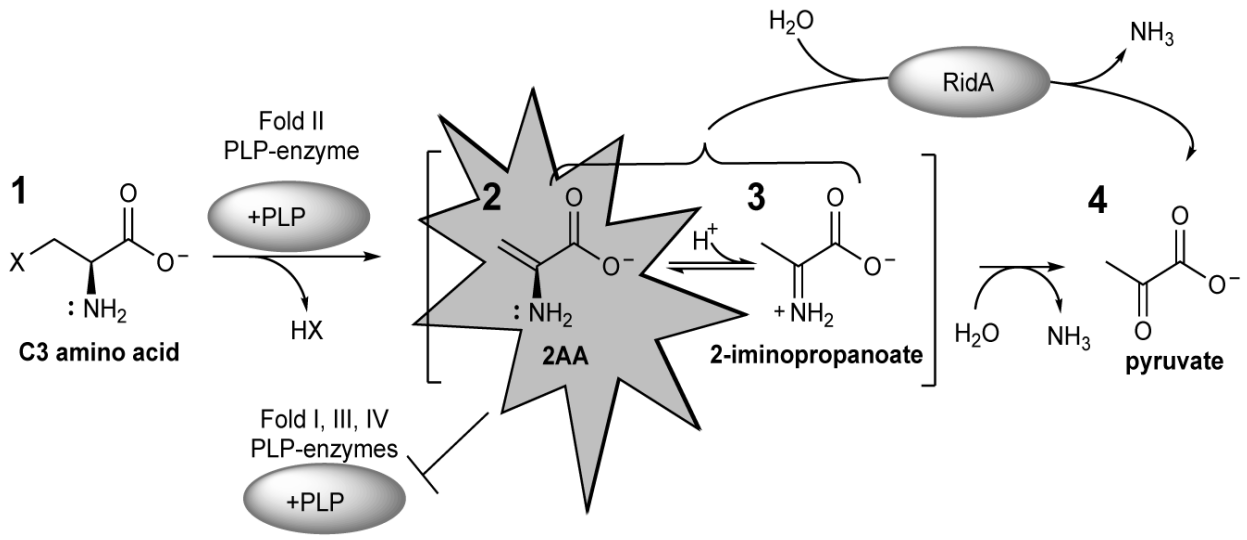


FIGURE 3.1. RidA paradigm in *S. enterica*. Various fold-type II Pyridoxal 5'-phosphate (PLP)-dependent enzymes catalyze β-elimination reactions of (1) 3 carbon (C3) alpha amino acids with favorable leaving groups on the beta-carbon. These reactions proceed through the reactive enamine intermediate (2) 2-aminoacrylate (2AA). 2AA can tautomerize to its imine form, producing (3) 2-iminiopropanoate, which then is non-enzymatically hydrolyzed into the more stable keto-acid (4) pyruvate. A subset of these enzymes release 2AA from the active site. RidA catalyzes the deamination of 2AA and other enamines *in vivo*. It has been demonstrated that elimination of RidA leads to accumulated free 2AA, which can covalently modify and inactivate a number of PLP-dependent enzymes that belong to the fold-type I, III, and IV families.

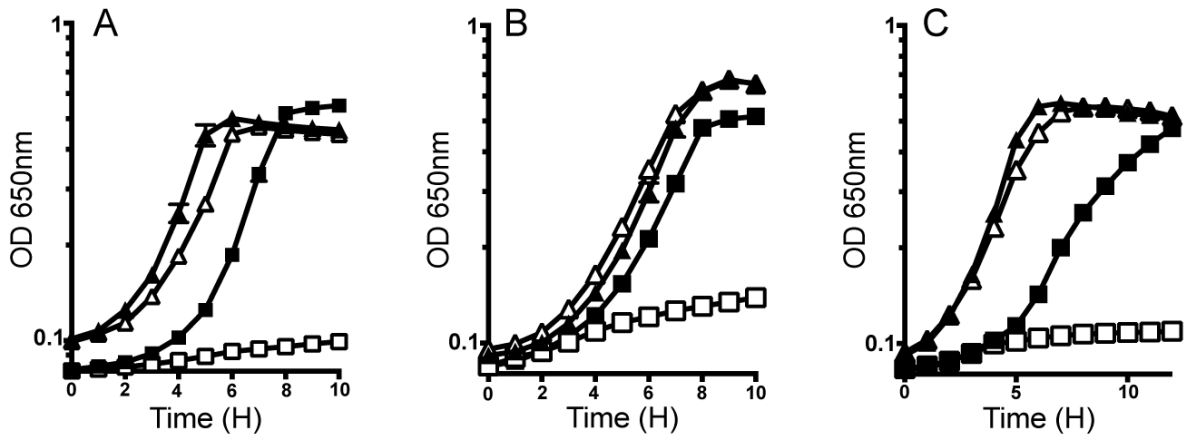


FIGURE 3.2. *S. enterica* and *E. coli* respond distinctly to the loss of RidA proteins. Growth of *S. enterica* (squares) and *E. coli* (triangles) strains is shown. Strains are wild-type (shaded), and *ridA* (*S. enterica*) or *ridA tdcB rutC* (*E. coli*) (open) strains grown in; (**A**) glucose (11 mM) medium with serine (5 mM), (**B**) pyruvate (50 mM) medium, and (**C**) glucose medium with L-alanine (15 mM) as sole nitrogen source. Media in panels A and B had ammonia as nitrogen source. All strains were grown at 37°C and data are the mean of three biological replicates.

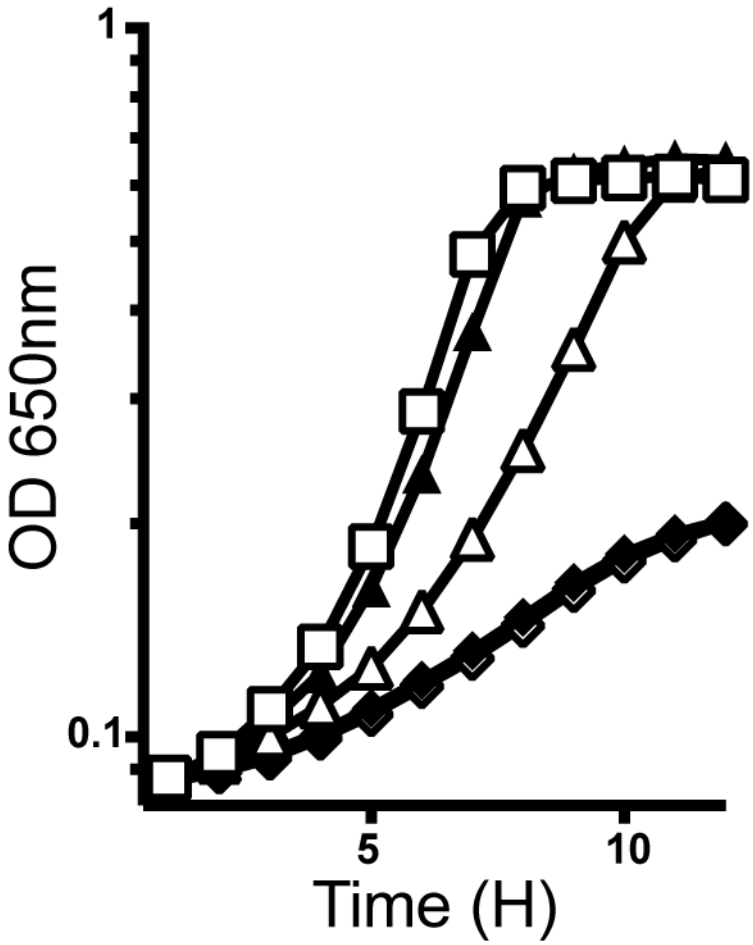


FIGURE 3.3. Endogenously generated 2AA is increased with inducible IlvA system. *S. enterica* strain DM15035 (*ABcc1*) and DM15036 (*ridA ABcc1*) strains (open boxes and open triangles, respectively), were grown in 20 mM glycerol. Strain DM15036 was also grown in in 20 mM glycerol plus 1 mM isoleucine (closed triangles), plus 0.2% L-arabinose (open diamonds) and plus 0.2% L-arabinose and 1 mM isoleucine (closed diamonds). Data are the mean of experiments performed in biological triplicate.

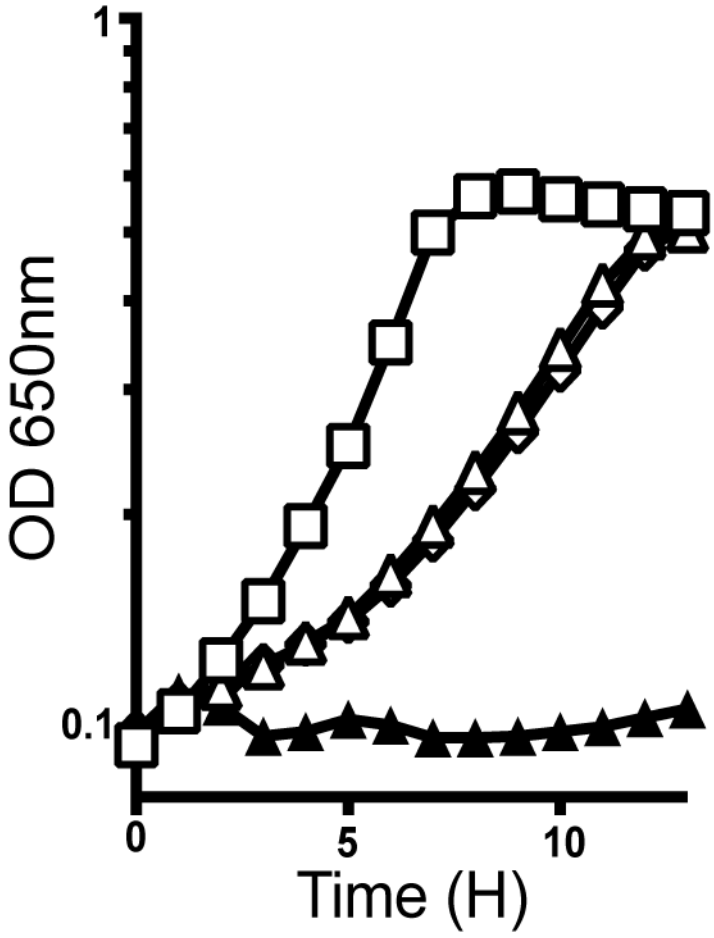


FIGURE 3.4. RidA and TdcF can quench 2AA in *E. coli*. *E. coli* strains DM14931 (*ABcc1*) (open boxes), DM15055 (*ridA tdcF ABcc1*) (open triangles), and DM14994 (*ridA tdcF rutC ABcc1*) (open diamonds) strains were grown in 20 mM glycerol plus 0.2% L-arabinose. Strain DM14994 (*ridA tdcF rutC ABcc1*) was also grown in 20 mM glycerol plus 0.2% L-arabinose and 5 mM serine (filled triangles). Data are the mean of experiments performed in biological quadruplicate.

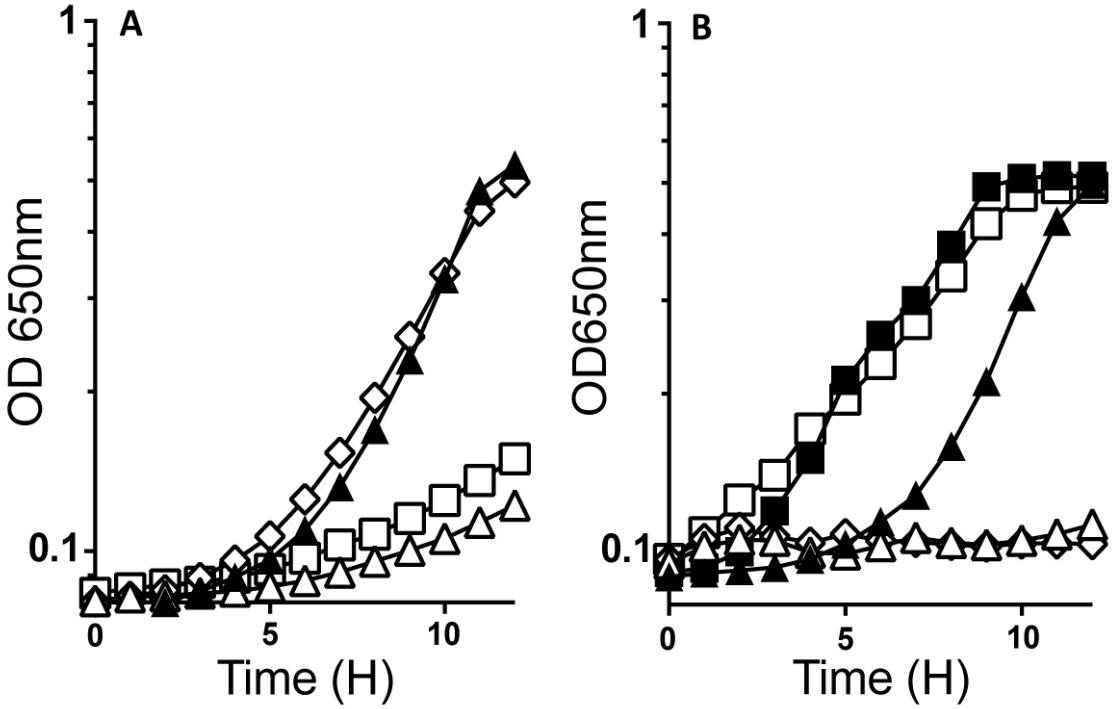


FIGURE 3.5. Aspartate addition overcomes 2AA-dependent growth defect in *E. coli*. (A) Growth of *S. enterica* strains DM15035 (*ABcc1*)(shaded) and DM15036 (*ridA ABcc1*)(open) in minimal glycerol medium containing 0.2% L-arabinose (triangles), 0.2% L-arabinose and 1 mM aspartate (squares), or 0.2% L-arabinose and 1 mM glycine (diamonds). (B) Growth of *E. coli* strains DM14931 (*ABcc1*) (shaded) and DM15055 (*ridA tdcF ABcc1*)(open) in minimal glycerol media with 0.2% L-arabinose. Media was supplemented with 5 mM serine (triangles), 5 mM serine and 1 mM glycine (diamonds), or 5 mM serine and 1 mM aspartate (squares).

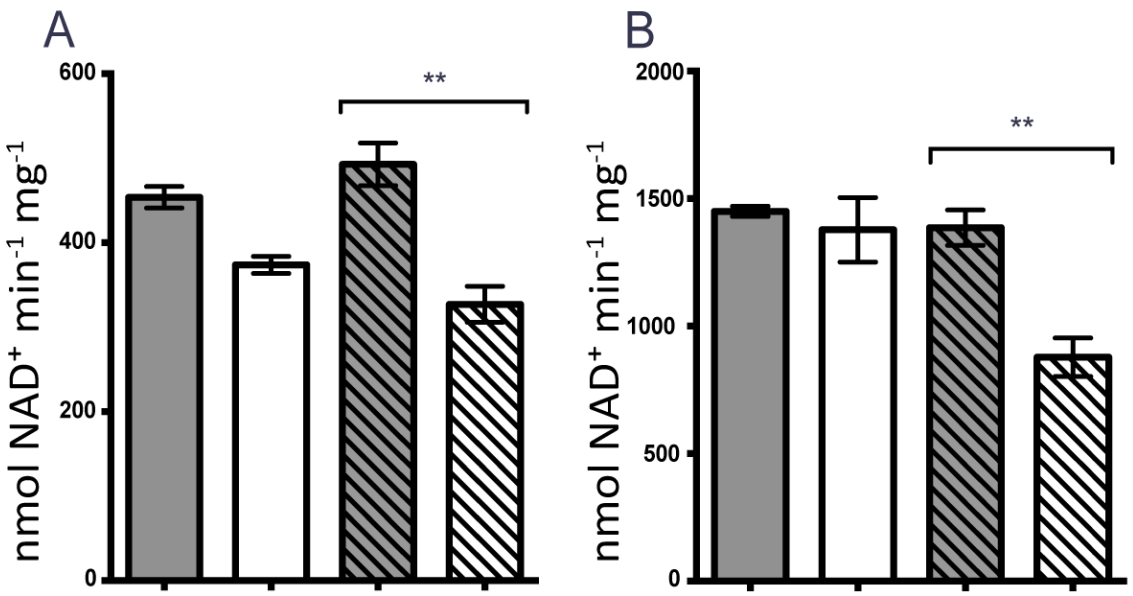


FIGURE 3.6. 2AA lowers the activity of asparate aminotransferase *in vivo*. (A) Aspartate transaminase (AspC) activity was measured following growth of *S. enterica* strains DM15035 (*ABcc1;* gray fill) and DM 15036 (*ridA ABcc1;* white fill) to late-logarithmic phase (OD_{650} 0.9-1.0) in minimal glycerol (20 mM) medium containing 0.67 mM glycine (solid bars) or minimal glycerol medium containing glycine and 0.2% L-arabinose (hatched bars). (B) AspC activity was also measured following growth of *E. coli* strains DM14931 (*ABcc1;* gray fill) and DM15055 (*ridA ABcc1;* white fill) to late-logarithmic phase (OD_{650} 0.9-1.0) in minimal glycerol (20 mM) medium (solid bars) or minimal glycerol medium containing 0.2% L-arabinose (hatched bars). In all cases, aminotransferase activity was measured in crude extract by coupling oxaloacetate formation to malate dehydrogenase activity and monitoring NADH oxidation at 340 nm from three biological replicates. Error bars represent the standard error of the means of the replicates; asterisks (**) denote statistically significant ($p < 0.01$) variation between samples.

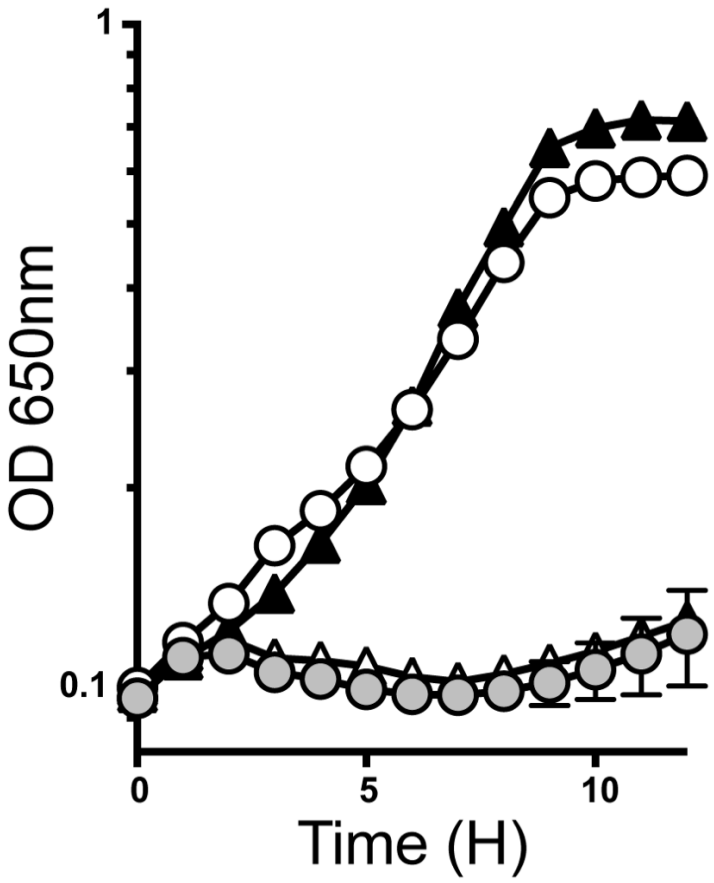


FIGURE 3.7. Purines restore growth via a PurR-dependent mechanism. *E. coli* strains DM14931 (*ABcc1*)(black), DM15055 (*ridA tdcF ABcc1*)(open), and DM15832 (*ridA tdcF ABcc1 purR*) (gray) in minimal glycerol media with 0.2% L-arabinose. Media was supplemented with 5 mM serine (triangles), or 5 mM serine and 0.4 mM adenine (circles). All strains were grown at 37°C and data are presented as the mean of experiments performed in biological triplicate.

CHAPTER 4

ANALYSIS OF VARIANTS OF THE SER/THR DEHYDRATASE, ILVA, PROVIDE INSIGHT INTO 2-AMINOACRYLATE METABOLISM IN *SALMONELLA ENTERICA*[1]

[1]Borchert AJ, Downs DM. 2018. *Journal of Biological Chemistry* 293(50):19240-19249
Reprinted here with permission of the publisher.

4.1 ABSTRACT

RidA is a conserved and broadly distributed protein that has enamine deaminase activity. In a variety of organisms tested thus far, lack of the RidA(s) result in the accumulation of the reactive metabolite, 2-aminoacrylate (2AA). 2AA is an obligate intermediate in the catalytic mechanism of several PLP-dependent enzymes. This study reports the characterization of variants of the biosynthetic serine/threonine dehydratase (EC 4.3.1.19; IlvA), which is a significant generator of 2AA in at least *S. enterica*, *E. coli*, *Pseudomonas aeruginosa* and *Saccharomyces cerevisiae*. The previously isolated *ilvA3210* and *ilvA3211* alleles suppressed phenotypic growth consequences of 2AA accumulation. Characterization of the relevant protein variants suggested they affected 2AA metabolism *in vivo* by two different catalytic mechanisms, both leading to an overall reduction in serine dehydratase activity. A simple mathematical model describing the affect these catalytic deficiencies placed on 2AA production was generally supported by the data. However, caveats arose when kinetic parameters, determined *in vitro,* were used to predict formation of the isoleucine precursor 2-ketobutyrate and model *in vivo* (growth) behaviors. Altogether, the data support the value a holistic approach to modeling metabolism, accounting for the robust and often subtle responses of the metabolic network to perturbations.

4.2 BACKGROUND

Microbial metabolic networks are robust systems able to absorb perturbations caused by internal and external stress [1-4]. Metabolic robustness provides an evolutionary advantage to an organism exposed to sudden and frequent environmental shifts, but the integration of pathways can complicate the ability to predict metabolic potential or model flux distribution based solely upon genomic cataloging [5-7]. The extensive integration of metabolic pathways within the

network can lead to unanticipated solutions to metabolic perturbations that are often uncovered by genetic analyses [8-10]. This point is exemplified by the observable differences between *S. enterica* and *E. coli* with respect to 2-aminoacrylate (2AA) stress and thiamine biosynthesis, despite the conservation of all the component enzymes in each [8, 11]. The observation that organisms containing the same components may alter the configuration of those parts indicates the need for genetic and biochemical approaches to supplement genomic cataloging in efforts to model cellular metabolism [5, 12].

A number of pyridoxal 5'-phosphate (PLP)-dependent β-eliminase enzymes produce and release the reactive molecule 2-aminoacrylate (2AA). 2AA then inactivates target PLP-enzymes through covalent modification of active site bound PLP [13-16]. RidA subfamily members, belonging to the broader Rid (YjgF/YER057c/UK114) protein family, are enamine/imine deaminases that reduce 2AA accumulation [13, 14, 17, 18]. Strains of *S. enterica* lacking RidA do not grow on minimal pyruvate medium or minimal glucose medium containing serine. These growth defects are due to 2AA generated by the biosynthetic serine/threonine dehydratase (IlvA; EC 4.3.1.19) from serine. *S. enterica ridA* mutants show a mild growth defect in minimal glucose medium, attributable to 2AA generated from endogenously generated serine. Inclusion of serine increases endogenous serine pools and results in greater 2AA production. A *ridA* mutant fails to grow as a result of the increased 2AA-dependent damage to multiple enzymes [19-22]. This global stress is best prevented by mechanisms that reduce 2AA production (e.g. allosteric inhibition of IlvA by isoleucine) [14, 19, 23-26]. Lack of growth on minimal pyruvate medium has been attributed to 2AA-dependent damage of a single enzyme, the branched-chain amino acid aminotransferase (IlvE; EC 2.6.1.42) [21]. Wild-type *S. enterica*, grown on pyruvate, experiences a flux bias away from isoleucine biosynthesis and toward valine, though sufficient isoleucine is

still produced for growth [27]. However, *ridA* mutants undergo 2AA-dependent damage of IlvE, the penultimate enzyme in the isoleucine synthesis pathway, and experience an isoleucine limitation [14, 21]. An overview of the role of RidA in the context of branched chain amino acid biosynthesis is depicted in Fig 1.

Suppressor mutations that restored growth to a *ridA* mutant on minimal pyruvate mapped to *ilvA* and encoded variant IlvA proteins [25]. The suppressor alleles of *ilvA, ilvA3210* and *ilvA3211*, restored wild-type levels of IlvE activity and allowed isoleucine biosynthesis sufficient for growth. Isolation and initial analysis of these suppressors was prior to the elucidation of RidA function and the role of 2AA in causing the phenotypes of a *ridA* mutant. Notably, the suppressor mutations were unable to restore growth of a *ridA* mutant on medium containing serine, indicating the IlvA variants still produce 2AA when serine levels are elevated.

Based on the current understanding of the RidA paradigm, two scenarios could explain how the IlvA variants restored growth of a *ridA* mutant on pyruvate medium, but not in the presence of serine. First, the IlvA variants could have decreased affinity for the 2AA precursor, L-serine. Second, the catalytic turnover of the variants could be reduced, maintaining sufficient 2-aminocrotonate production for isoleucine biosynthesis, but minimizing 2AA formation. In either scenario, 2AA production by the variants would be reduced, but not eliminated. The goal of this work was to biochemically characterize the IlvA variants and better understand the synthesis and impact of 2AA formation by IlvA in the metabolic context of branched chain amino acid synthesis. We further addressed the feasibility of using kinetic and metabolite concentration data to model the factors controlling IlvA-dependent 2AA metabolism and its impact *in vivo*.

4.3 MATERIALS AND METHODS

Bacterial Strains, Media, and Chemicals. Strains used in this work are provided in Table 3 and are all derivatives of *Salmonella enterica* serovar Typhimurium LT2. Minimal medium was no-carbon E (NCE) supplemented with 1 mM MgSO4 [28], trace minerals [29], and 11 mM D-glucose or 50 mM pyruvate as the sole carbon source. Difco nutrient broth (NB) (8g/L) containing NaCL (5 g/L) was used as rich medium. Difco BiTek agar (15g/L) was added for solid medium. Superbroth (32 g tyrptone, 20 g yeast extract, 5 g sodium chloride, and 5 mL 1 N sodium hydroxide/liter) was used for purifying proteins. Antibiotics were used at the following concentrations for rich (or minimal) medium: kanamycin, 50 (12.5) μg/mL; chloramphenicol, 20 (5) μg/mL; and ampicillin 150 (7.5) μg/mL. All chemicals were obtained from Sigma (St. Louis, MO).

Genetic techniques. Strains were constructed via transductional crosses using the high-frequency general transducing mutant of bacteriophage P22 (HT105/1, *int*-201)[30]. Methods for performing transductions, isolation of cells from phage contamination, and identifying phage-free recombinants were previously described [31, 32]. Briefly, recipient cells (~10^8 CFU) and transducing phage (~10^9 PFU) were preincubated 1 h at 37 °C prior to plating on selective media. Transductants were purified by isolation using non-selective green indicator plates and putative phage-free clones were identified as light colonies. These clones were checked for phage sensitivity by cross-streaking with phage P22. *E. coli* BL21-AI *ΔilvA823* was constructed by transducing the *ilvA723*::*kan* mutation from the Keio collection into the BL21-AI strain background [33]. The kanamycin cassette was then resolved using FLP recombinase, as previously described by Datsenko and Wanner [34], to create *E. coli* BL21-AI *ΔilvA823*.

Molecular techniques. The *ilvA* locus was amplified from LT2 strains DM7608

(*ilvA3211*) and DM7610 (*ilvA3210*) by PCR with Q5 high-fidelity DNA polymerase (New England BioLabs) using primers IlvA_BspQI_LT2_F and IlvA_BspQI_LT2_R (Table 3). The PCR product was purified, digested with BspQI, and ligated into the modified pET28b, pSAPKO-CH, as described by Galloway et al. [35]. The resulting plasmids were named pDM1461 and pDM1543, which encoded $IlvA_{A142T}$ and $IlvA_{G191S}$ fused to a C-terminal penta-histidine tag, respectively. *ilvA219* was amplified from LT2 strain DM6947 by PCR with Q5 high-fidelity DNA polymerase (New England BioLabs) using primers IlvA_NdeI_LT2_F and IlvA_XhoI_LT2_R (Table 1). The PCR product was purified, digested with NdeI and XhoI (New England BioLabs), and ligated into predigested pET20b (Novagen), digested with the same enzymes. The resulting plasmid was named pDM1420 and encoded $IlvA_{L447F}$ fused to a C-terminal hexa-histidine tag. Constructs were transformed into *Escherichia coli* DH5α and transformants were screened for vectors containing the correct insert. Inserts were confirmed through sequence analysis, performed by Eton Biosciences.

Protein Purification. The three plasmids generated above and pDM1578, which encoded *S. enterica* IlvA fused to a C-terminal hexa-histidine tag and was generated previously [13], were each inserted into an *E. coli* BL21AI strain (Invitrogen, Madison, WI) lacking the native *ilvA* locus. The resulting strains were inoculated into 2 ml of LB containing either ampicillin (pET20b-*ilvA* and pET20b-*ilvA219*) or kanamycin (pSapKOCH-*ilvA3210* and pSapKOCH-*ilvA3211*) and allowed to reach full density. These were subcultured into 50ml superbroth and grown for 3 h before inoculating 1.5 L of superbroth, containing the appropriate drug. Cultures were grown at 30 °C shaking (180 rpm) to A_{650} of 0.6, and expression was induced by the addition of 0.2% arabinose (pET20b-*ilvA* and pET20b-*ilvA219*) or 0.2% arabinose and 0.2 mM IPTG (pSapKOCH-*ilvA3210* and pSapKOCH-*ilvA3211*) and allowed to grow an additional 16 hours. Cells were

harvested at 4 °C by centrifugation (7,000 x g) and pellets were frozen at -80 °C until purification. During purification, cells were resuspended in 2.5 mL buffer (100 mM KPO_4, pH 8, 100 mM NaCl, 20 mM imidazole, and 10% glycerol) for every 1 g cell pellet. Resuspension was incubated on ice with 1 mg/mL lysozyme and 0.125 mg/mL DNase 30 min prior to cell disruption using a One Shot Cell Disruptor at 18,000 psi (Constant Systems). After disruption, phenylmethylsulfonyl fluoride was added to 1 mM. Lysates were clarified by centrifugation (48,000 x g), and filtered through a 0.45 μm filter. Filtered lysates were bound to a 5 mL HisTrap™ HP column and purified using the manufacturer's protocol (GE Healthcare). The column was washed using resuspension buffer (100 mM KPO_4, pH 8, 100 mM NaCl, 20 mM imidazole, and 10% glycerol), and then eluted with elution buffer (100 mM KPO_4, pH 8, 100 mM NaCl, 500 mM imidazole, and 10% glycerol), in a linear gradient over the course of 10 column volumes. The fractions were assessed for purity, pooled, and concentrated using a 30,000 molecular weight cut-off Ultra-15 Centrifugal filter (Amicon). Samples were desalted using a PD-10 Desalting column (GE Healthcare), and eluted in storage buffer (100 mM KPO_4, pH 8.0, 100 mM NaCl, 10 μM PLP, 10% glycerol). Protein concentration was determined by the BCA assay (Pierce). Protein was frozen in liquid nitrogen and stored at -80 °C.

Quantification of Growth. A 2 μL aliquot of cells from an overnight NB culture was used to inoculate 198 μL of growth medium. 96-well plates were incubated at 37°C in a microplate reader (model EL808; Bio-Tek Instruments) with low shaking and growth was monitored as the change in optical density at 650 nm (OD_{650}) over time. Unless stated, growth experiments were performed in biological triplicate. Results were plotted using GraphPad Prism 7.0d, with curves (log_{10}-format) representing averages and standard error of the means for the replicates.

Dehydratase (IlvA) assays. *Crude assay.* Full-density NB cultures were inoculated

(1:100) into 100 mL minimal medium containing 0.67 mM glycine and grown at 37°C shaking to stationary phase (12 h) before being pelleted (10,000 × *g*, 10 min) and stored as pellets at -80°C. Cell pellets were resuspended in 2 mL of buffer (50 mM KPO_4, pH 8.0, 0.4 mM dithiothreitol). Lysates were prepared using one pass through a One Shot cell disrupter (Constant Systems) at 18,000 psi, and clarified by centrifugation (20 min @ 16,000 × *g* at 4°C). Threonine dehydratase assays were carried out as described previously [26, 36] and reported as micromoles 2-ketobutyrate formed per milligram protein per minute. Protein concentrations were estimated by the method from Bradford [37]. The experiment contained three biological replicates for each strain tested and the average with standard deviation is reported.

Purified enzyme assay. Pyruvate and 2-ketobutyrate formation was assayed by tracking the absorbance at 230 nm [13, 38] at room temperature (25 °C) for two minutes. Assays were carried out in quadruplicate, using 100 mM potassium phosphate buffer (pH 8.0). To obtain values within the range of detection for the instrument, reactions used 100 nM of IlvA, $IlvA_{L447F}$, and $IlvA_{A142T}$ for threonine assays and 200 nM of IlvA, $IlvA_{L447F}$, and $IlvA_{A142T}$ for serine assays; 1 μM $IlvA_{G191S}$ was used for both assays. The indicated concentrations of serine and threonine were used to start each 200 μL reaction, which were assayed continuously in a quartz 96 well plate at room temperature by a SpectraMax Plus (Molecular Devices) supporting SoftMax Pro 7.0 software.

Stability Assays. Pyruvate formation with purified protein was assayed as described above. Assays were carried out every two hours, for a total of 8 hours with 200 mM L-serine used to start each reaction. Each measurement was performed in duplicate, using 200 nM of IlvA, $IlvA_{L447F}$, and $IlvA_{A142T}$ and 1 μM $IlvA_{G191S}$. Enzyme preparations were kept on ice between measurements.

Branched-chain amino acid aminotransferase (IlvE) assays. Full-density NB cultures were inoculated (1:100) into 5 mL minimal medium containing 0.67 mM glycine and grown at

37°C shaking to stationary phase (12 h) before being pelleted (10,000 × *g, 2 min)* and stored as pellets at -20°C. IlvE activity was assayed according to previously described methods [26]. Cells were permeabilized in buffer (50 mM KPO_4, pH 8, 50 μM PLP, 10 mM 2-ketoglutarate) containing 10% PopCulture Reagent (Novagen). The reaction was started with 20 mM L-Isoleucine. The product, 2-keto-3-methylvalerate (KMV), was derivatized by 2,4-dinitrophenylhydrazine (DNPH), producing a chromophore with absorbance of 540nm. KMV was quantified using a standard curve from known quantities of KMV derivatized with DNPH, and normalized to total protein content [37].

Kinetic data analysis. Initial velocity calculations were estimated over the course of 2 min reactions tracking $\Delta mOD_{230}\ min^{-1}$ for various concentrations of L-threonine or L-serine. Standard curves for absorbance at 230 nm for known amounts of 2-ketobutyrate and pyruvate were used to transform $\Delta mOD_{230}\ min^{-1}$ to μmol 2-KB min^{-1} and μmol pyruvate min^{-1}. A non-linear regression fit (Equation 2) to a graph of the substrate concentration vs. initial velocity was used to fit the data. In the absence of cooperative effects ($h = 1$), this reduced to the equation by Michaelis and Menton (Equation 3), from which each IlvA variant's Michaelis constant (K_M) and maximal velocity (V_{max}) values for each substrate could be determined. Calculations were performed using GraphPad Prism 7.0d.

$$v = \frac{V_{max}[S]^h}{K_{1/2}{}^h + [S]^h} \quad (2)$$

$$v = \frac{V_{max}[S]}{K_M + [S]} \quad (3)$$

The turnover number (k_{cat}) was obtained using the active site concentration (E_t) (Equation 4). Since IlvA is a dimer of dimers, containing two active sites per enzyme, E_t is equal to half of the enzyme concentration used in the assay [39, 40].

$$k_{cat} = \frac{V_{max}}{[E_t]} \quad (4)$$

4.4 RESULTS

***ilvA3210* and *ilvA3211* alleles suppress one 2AA-dependent growth defect.** Growth of strains with wild-type *ilvA*, *ilvA3210,* and *ilvA3211* was quantified and the data are shown in Figure 2. The *ridA* mutant failed to grow with pyruvate as a sole carbon source (Fig. 2A, B) or when minimal glucose medium included serine (Fig. 2C, D). As previously reported, the *ilvA3210* and *ilvA3211* alleles significantly improved growth of the *ridA* mutant on pyruvate, but not in minimal glucose with serine [21]. On pyruvate medium, i) growth of the *ridA ilvA3210* and *ridA ilvA3211* mutants was not fully restored to wildtype levels, and ii) the *ilvA3210* allele caused a growth defect in an otherwise wild-type strain. Both of these phenotypes were corrected by the addition of isoleucine (Fig 2). In glucose medium containing 5 mM serine, the *ridA ilvA3210* and *ridA ilvA3211* strains grew only slightly better than a *ridA* mutant (Fig. 2 C&D). In total these data were consistent with the hypothesis that i) $IlvA_{A142T}$, and $IlvA_{G191S}$ have reduced dehydratase activity, ii) the reduced serine dehydratase activity of the IlvA variants decreases 2AA available to damage cellular targets, iii) *ilvA3210* decreases flux enough to cause a starvation for isoleucine on pyruvate, even in the absence of a *ridA* mutation.

***ilvA* alleles restore branched-chain amino acid aminotransferase activity.** Threonine dehydratase (IlvA) and branched-chain amino acid aminotransferase (IlvE) activity was assayed in the crude extracts of strains containing four different *ilvA* alleles (Table 1, white fill) and stains

carrying the same IlvA alleles in addition to a *ridA* mutation (Table 1, gray fill). Analysis of the strains with a wild-type *ridA* provides insight into the impact of the IlvA variants on the biosynthetic pathway. Each of the *ilvA* alleles caused a decrease in total threonine dehydratase activity, with *ilvA3210* ($IlvA_{A142T}$, DM15898) and *ilvA3211* ($IlvA_{G191S}$, DM15936) causing a ~10-fold decrease. Despite the decrease in threonine dehydratase activity, the strains had wild-type growth level on minimal glucose medium, emphasizing the excess capacity the cells have to synthesize isoleucine. IlvA and IlvE are encoded in an operon (*ilvGMEDA*) and the ratio of their activity suggests the status of isoleucine synthesis and/or starvation in the cell [41]. IlvE activity was significantly increased in the *ilvA3210* strain, suggesting that the decreased activity of $IlvA_{A142T}$ caused a limitation of isoleucine, which resulted in induction of the operon. Consistent with this model, previous work showed by Western Blot that greater levels of IlvE protein accumulate in the *ilvA3210* strain, as compared to wild-type LT2 [21]. These data correlate with the growth defect of DM15898 (*ilvA3210*) on pyruvate, where flux through the pathway is less than on glucose medium. While measured threonine dehydratase activity in crude lysates was slightly higher for $IlvA_{A142T}$ than $IlvA_{G191S}$, assuming a similar induction across the *ilv* operon, the data suggest the specific activity of $IlvA_{A142T}$ *in vivo* is no greater than that of $IlvA_{G191S}$. In total, the results show that the capacity of the biosynthetic pathway exceeds what is needed to satisfy growth and supports a hypothesis that allosteric regulation is a major driver controlling metabolic efficiency.

The impact of eliminating RidA in the strains described above is shown by the data in gray in Table 2. IlvE is irreversibly damaged by 2AA and can be used as a proxy for the amount of this metabolite that accumulates *in vivo* [14]. A *ridA* mutant (DM3480) accumulates 2AA, which results in the ~50% reduction in IlvE activity, compared to wild-type (DM9404) (Table 1). The *S.*

enterica $IlvA_{L447F}$ variant, encoded by the *ilvA219* allele, is insensitive to inhibition by isoleucine [42] and despite having decreased threonine dehydratase activity, increases 2AA production *in vivo* [14, 26]. This results in a severe decrease in IlvE activity (~75%) in the *ridA ilvA219* mutant (DM6946). In contrast, while the *ridA* mutant strains expressing the $IlvA_{A142T}$ or $IlvA_{G191S}$ variants had lower IlvA activity, IlvE activity was similar, or slightly higher, than the parental strains (Table 1). This latter result suggested that unlike the $IlvA_{L447F}$ variant, the two suppressor variants of IlvA generated less 2AA *in vivo.*

The data presented in Table 1 and Figure 2 suggested a simple correlation between IlvE activity and growth on pyruvate. IlvE activity of ~100 nmol KMV min^{-1} mg^{-1} (measured on glucose medium) supports adequate isoleucine synthesis for growth in minimal pyruvate medium. By this scenario, *ridA* derivatives of wild-type or *ilvA219* strains have lower IlvE activity, which constricts isoleucine synthesis and prevents growth. The restored activity of IlvE in the *ridA ilvA3210* and *ridA ilvA3211* mutants overcomes the growth defect. Notably, this simple model does not explain why *ilvA3210*, but not *ilvA3211*, resulted in an isoleucine requirement in minimal pyruvate media.

Characterization of purified IlvA variants. Since IlvA uses serine and threonine as competing substrates, it was possible that differential substrate specificities influenced the *in vivo* effects described above. Kinetic parameters were determined from purified preparations of the four IlvA protein variants, using either serine or threonine substrate (Fig 3). The threonine saturation curves for all four proteins were fit using the equation by Michaelis and Menton. The saturation curves using serine substrate showed mild positive cooperativity, with Hill coefficients (*h*) greater than 1. Since the fit of the curves was close using the equation from Michaelis and Menton, the slight cooperative effects were not considered in determining the kinetic values of the

proteins. This tactic simplified the comparisons between the results with serine and threonine substrates.

Considering wild-type protein, the values obtained for the Michaelis constant (K_M) and catalyst rate constant (k_{cat}) with serine or threonine were consistent with values previously observed for *S. enterica* IlvA [43]. The $IlvA_{L447F}$ variant had a lower K_M for serine and higher k_{cat} with both threonine and serine substrates. Compared to wild-type enzyme, the $IlvA_{A142T}$ variant had significantly higher K_M and lower V_{max} values for both substrates, while the $IlvA_{G191S}$ variant had significantly lower K_M and V_{max} values for both substrates. The specificity constant (k_{cat}/K_M) of an enzyme can be used to assess preference for different substrates. The k_{cat}/K_M values showed that all four IlvA proteins were more effective at using threonine than serine (Fig 3). However, the ratio of specificity constants (threonine k_{cat}/K_M versus serine k_{cat}/K_M) increases from 5 for IlvA to 8 and 9 for $IlvA_{A142T}$ and $IlvA_{G191S}$, respectively. This suggests that $IlvA_{A142T}$ and $IlvA_{G191S}$ have a slightly stronger preference for threonine over serine than wild-type enzyme.

To assess the relative stability of each enzyme variant, enzyme activity was measured in the presence of 200 mM serine periodically for a total of eight hours. Under the test conditions, $IlvA_{A142T}$ was least stable, showing an 80 % reduction from initial activity after eight hours (40 nmol pyruvate min^{-1} vs. 184 nmol pyruvate min^{-1}). This was 8-fold greater than wild-type IlvA (716 nmol pyruvate min^{-1} vs. 777 nmol pyruvate min^{-1}) or $IlvA_{L447F}$ (1,043 nmol pyruvate min^{-1} vs. 1,162 nmol pyruvate min^{-1}), which only showed a 10% loss in activity. $IlvA_{G191S}$ showed moderate susceptibility to degradation, with a 40% reduction from initial activity after 8 hours (18 nmol pyruvate min^{-1} vs. 30 nmol pyruvate min^{-1}).

Modeling the competing synthesis of 2AA and 2AC *in vivo*. Use of threonine and serine by IlvA generates two significant metabolites, 2-aminocrotonate (2AC) and 2-aminoacrylate,

respectively. 2AC is deaminated to 2-ketobutyrate, a precursor of isoleucine which is not known to be detrimental if it accumulates. In contrast, 2AA is detrimental if it accumulates before its conversion to pyruvate, as described above. Thus, the two products of IlvA have opposing impacts on the growth phenotypes and therefore differences in how the IlvA variants impacted production of both 2AC and 2AA *in vivo* was significant. The kinetic information of each of the four IlvA enzymes (Fig 3) was used to explore the relationship between formation of 2AA and 2AC in cells with different *ilvA* alleles. Reported concentrations of endogenous metabolites in *E. coli* were used to estimate the concentration of L-serine (67 μM) and L-threonine (180 μM) [44]. The concentration of IlvA (20 μM) was based on the concentration of other PLP-dependent enzymes in *E. coli* [45]. For simplicity, known cooperative and allosteric behaviors of IlvA were excluded from the calculations [46, 47]. Equation 1 predicts product formation for a specified substrate (serine or threonine), accounting for competitive inhibition by the alternative substrate [48]. V_{Prod} is the rate of product production (μM/s), $k_{cat(Sub)}$ is the turnover number for the substrate (s^{-1}), [E_t] is IlvA active-site concentration (μM), [Sub] is substrate concentration (μM), [CompSub] is competing substrate concentration (μM), $K_{M(sub)}$ is the Michaelis constant for the substrate (μM), and $K_{M(CompSub)}$ is the Michaelis constant for the substrate (μM).

$$V_{Prod} = \frac{(\boldsymbol{k_{cat(Sub)}} \times [\boldsymbol{E_t}])[\boldsymbol{Sub}]}{\boldsymbol{K_{M(Sub)}}\left(\mathbf{1} + \frac{[\boldsymbol{CompSub}]}{\boldsymbol{K_{M(CompSub)}}}\right) + [\boldsymbol{Sub}]} \quad (1)$$

Using Equation 1, the amount of 2AC and 2AA generated during steady state growth was estimated for each of the IlvA variants (Table 2). Consistent with the conclusion from the data in Figure 3, Equation 1 predicts each IlvA enzyme generates more 2AC than 2AA. Further, the variants that suppressed the *ridA* mutant phenotype on pyruvate produced dramatically less 2AA

and 2AC than the wildtype protein. The growth of strains with these variants on minimal medium indicated that these alleles produced sufficient 2AC to satisfy the isoleucine requirement of the strain. These data suggest that despite being simplified, Equation 1 does an adequate job of accounting for the effect of the IlvA variants *in vivo* and suggests the mechanism they use to suppress the *ridA* phenotype on pyruvate. The *ilvA* alleles that restored growth to a *ridA* mutant failed to restore growth when exogenous serine was added, indicating detrimental levels of 2AA still accumulated under these conditions. Equation 1 was used to calculate the impact of 5mM serine on the products of IlvA. Assuming endogenous serine matched the external concentration, the data showed a significant increase in 2AA production in all strains (Table 2). The strains carrying *ilvA3210* and *ilvA3211* were calculated to have 2AA levels well above what accumulated in a wild-type strain without serine addition (85 μM/s and 29 μM/s for *ilvA3210* and *ilvA3211*, respectively). Interestingly, the amount of 2AC was essentially unchanged in this condition.

A hallmark of the RidA paradigm is that the effects are prevented by the addition of isoleucine, which inhibits IlvA and prevents formation of the toxic 2AA. Addition of threonine was similarly shown to reduce the formation of 2AA, a result that was previously suggested to be due to competitive inhibition of serine catalysis by IlvA. However, for threonine to cause a 10-fold decrease in 2AA production by competitive inhibition, 55 mM L-threonine would be required (Equation 1). In contrast, addition of 1 mM L-threonine eliminated the growth defect of a *ridA* mutant on minimal glucose medium (Fig. 4A) [25]. Modeling of this condition with Equation 1 showed that 1 mM threonine did not affect the level of 2AA, but rather significantly increased the production of 2AC. Based on this analysis, we hypothesized that the exogenous threonine increased isoleucine biosynthesis, which allosterically inhibited IlvA, decreasing 2AA production (Figure 1). A *ridA* mutant expressing the isoleucine feedback-resistant variant has a growth defect

when grown in minimal glucose medium due to the increased 2AA formed. In this strain, L-threonine did not restore growth on minimal medium, a result that is consistent with threonine acting to increase isoleucine synthesis and resulting in the allosteric inhibition of IlvA to allow growth (Figure 4).

4.5 DISCUSSION

Strains lacking *ridA* in *S. enterica* have several metabolic defects that result from the accumulation of 2AA in the absence of RidA deaminase activity. This study sought to use the biochemical characteristics of IlvA variants to better understand the mechanism by which they reversed one defect, growth with pyruvate as a carbon source. Analyzing the role of IlvA in *ridA* phenotypes is complicated by, among other things, the fact that the enzyme generates both 2AA (from serine) and 2AC (from threonine). The former is potentially damaging and the latter is a metabolite necessary for synthesis of isoleucine. Therefore, variants that minimize 2AA formation cannot decrease 2AC so much that isoleucine synthesis is compromised. Kinetic analysis of four IlvA proteins, two of which suppressed a *ridA* growth defect, provided data to address the impact of 2AA and 2AC on growth. Based on these data and a simplified equation to characterize product formation *in vivo*, several conclusions were made.

IlvA variants able to suppress *ridA* phenotypes decrease 2AA production. Strains with $IlvA_{A142T}$ or $IlvA_{G191S}$ had significantly less threonine deaminase activity than wildtype in crude extracts. Equation 1 used data from Figure 3 to estimate that 2AA production from either $IlvA_{A142T}$ or $IlvA_{G191S}$ was reduced tenfold from wild-type IlvA (Table 2). IlvE activity in *ridA ilvA3210* and *ridA ilvA3211* strains was not statistically different from the isogenic parents (*ilvA3210* and *ilvA3211*, respectively), allowing the conclusion that a tenfold reduction in 2AA eliminated detectable damage to the target IlvE. Exogenous serine (5mM) in the growth media could still lead

to significant 2AA production (Table 2), which then eliminates growth (Figure 2). The fact that the IlvA variants with reduced 2AA production also had decreased 2AC production highlights the difficulty in altering enzymes to impact a single substrate.

IlvA$_{A142T}$ and IlvA$_{G191S}$ were compromised catalytically and had increased preference for threonine substrate. An expanded two-state model accounting for cooperative interactions within *E. coli* IlvA has been described previously [46, 47]. Since the saturation curves using threonine or serine substrate could be fit using the hyperbolic equation by Michaelis and Menton, the majority of enzyme active sites (in the absence of isoleucine) exist in the active R-state [47]. This simplified determination and comparison of K_M and k_{cat} values for each enzyme. Increased K_M and lower k_{cat} values suggested that IlvA$_{A142T}$ was unable to bind either substrate as well as wild-type enzyme and generated both products more slowly. Alternatively, IlvA$_{G191S}$ bound both substrates more readily (lower K_M), but was compromised in dehydratase activity (lower k_{cat}). Thus, both enzyme variants were defective in 2KB and pyruvate generation, but by different mechanisms. This highlights two distinct modifications that may be made to known 2AA generators, which could act as an alternative means for controlling 2AA production and stress in different organisms. Further, the ratio of k_{cat}/K_M values for threonine vs. serine indicated that IlvA$_{A142T}$ and IlvA$_{G191S}$ had a greater preference for threonine over serine than wild-type enzyme. PLP enzymes, such as IlvA, are well known for their substrate and reaction promiscuity [49-51]. This work shows how changes that result in greater specificity toward a single substrate can be undesirable since they can also hamper the maximal rate of catalysis.

Exogenous threonine increases flux toward isoleucine, indirectly controlling 2AA production. A previous report found that mutations increasing threonine biosynthesis could suppress the growth defect of a *ridA* mutant in the presence of serine [25]. Restoration of IlvE

activity suggested that threonine competed with serine for the active site of IlvA and thus reduced the amount of 2AA produced [25]. Implementation of Equation 1 suggested this was an unlikely explanation for the mechanism of threonine, since the amount of threonine required to reduce 2AA generation ten-fold would be unrealistically high (55 mM). The data supported a model in which isoleucine synthesis was increased when threonine was provided (i.e., increased AC production, no change in 2AA). Analysis of the impact of threonine on the growth of an *ilvA219 ridA* mutant supported this model. The *ilvA219* allele encodes a variant of IlvA ($IlvA_{L447F}$) insensitive to inhibition by isoleucine. If threonine addition reduced 2AA production via competitive inhibition, the *ilvA219 ridA* mutant should still be rescued by threonine supplementation. The finding that the *ilvA219 ridA* strain was not rescued by threonine supported the model implicating isoleucine biosynthesis (and allosteric inhibition of the 2AA generator, IlvA) as the relevant consequence of threonine addition.

Kinetic models can complement *in vivo* analyses. Assays of IlvA activity in crude extracts were the starting point for understanding how the IlvA variants impacted *ridA* phenotypes. Unexpectedly, there were several instances where the growth data, crude extract assays, and activity measurements using pure protein seemed inconsistent. First, purified $IlvA_{L447F}$ was not significantly different in catalytic capacity from the wild-type protein, yet in crude extract, it produced half the activity. Second, data from crude extracts confirmed the growth observation that strains expressing the $IlvA_{A142T}$ variant were starving for isoleucine, based on the apparent induction of the *ilv* operon (i.e, increased IlvE activity). However, the total activity of $IlvA_{A142T}$ in crude extracts was no lower than that of $IlvA_{G191S}$, despite strains with $IlvA_{G191S}$ having sufficient isoleucine synthesis for growth. Although the specific activity of $IlvA_{A142T}$ versus $IlvA_{G191S}$ is decreased when correcting for apparent induction of the operon, this could not explain how strains

expressing the IlvA$_{A142T}$ variant, but not IlvA$_{G191S}$, are starving for isoleucine. The isoleucine limitation disparity between *ilvA3210* and *ilvA3211* strains was somewhat reconciled by *in vitro* assays with purified protein, which identified that IlvA$_{A142T}$ had much lower affinity for threonine. Generally, this difference could explain how IlvA$_{A142T}$ was unable to produce sufficient 2AC *in vivo*, where threonine concentration is expected to be two orders of magnitude lower than that used in crude assays (20 mM). Nonetheless, when the specific K_M and k_{cat} values were modeled into Equation 1, strains encoding IlvA$_{A142T}$ were predicted to produce more 2AC than those encoding IlvA$_{G191S}$. This work provides some evidence that IlvA$_{A142T}$ could be less stable than IlvA$_{G191S}$. If IlvA$_{A142T}$ is more susceptible proteolysis *in vivo*, this may place a considerable energy burden upon the cell, as it must continually re-synthesize IlvA$_{A142T}$ to meet 2-ketobutyrate demands. However, whether increased IlvA$_{A142T}$ turnover occurs *in vivo* remains unexplored. Therefore, while thorough biochemical and genetic characterization proved useful in uncovering subtle metabolic discrepancies between similar systems, an inherent difficulty still exists in using *in vitro* determinations to precisely calculate metabolic flux and model *in vivo* behaviors. This finding may speak to the technical limitations, where aspects such as oligomerization, protein stability, small molecule binding affinity, and enzymatic activity are not effectively recapitulated *in vitro* [52]. The discrepancies between *in vitro* assays and *in vivo* behavior may also be explained, in large part, by transcriptional and post-transcriptional systems of regulation, and serves as a reminder that changes in flux caused by the enzyme variants may lead to regulatory shifts, as the cells respond to the perturbation. Therefore, regulatory mechanisms must be identified and incorporated into comprehensive models describing metabolic flux. Altogether, the limited effectiveness in using *in vitro* data to correctly predict growth behavior differences between wild-type, *ilvA3211*, and *ilvA3210* strains underscores technical limitations and physiological complexity, and so care

must be taken when using genomic and biochemical determinations to design *in silico* models predicting metabolic flux. Despite its simplicity, Equation 1 proved useful in predicting the physiological consequence of threonine addition on 2AA metabolism, and invalidated a previous model. In total, the data and discussion herein emphasize the need for a holistic combined approach to understanding and modeling metabolism, as we seek to define subtle and integrated responses of the cell to metabolic stress and cellular perturbation.

Acknowledgements. The authors thank Kelsey Hodge-Hanson for constructing plasmid pDM1461. An award from the competitive grants program at the NIH (GM095837) to DMD supported this work.

4.6 REFERENCES

1. Koenigsknecht, M.J. and Downs, D.M. (2010) Thiamine biosynthesis can be used to dissect metabolic integration. Trends Microbiol. 18 (6), 240-247.

2. Koenigsknecht, M.J. *et al.* (2012) Perturbations in histidine biosynthesis uncover robustness in the metabolic network of *Salmonella enterica*. PLoS ONE 7 (10), e48207.

3. Albert, R. *et al.* (2000) Error and attack tolerance of complex networks. Nature 406, 378-382.

4. Bazurto, J.V. and Downs, D.M. (2011) Plasticity in the purine-thiamine metabolic network of *Salmonella*. Genetics 187 (2), 623-631.

5. Bazurto, J.V. and Downs, D.M. (2016) Metabolic network structure and function in bacteria goes beyond conserved enzyme components. Microbial cell 3 (6), 260-262.

6. Cardinale, S. and Arkin, A.P. (2012) Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. Biotechnol. J. 7 (7), 856-866.

7. Kittleson, J.T. *et al.* (2012) Successes and failures in modular genetic engineering. Curr. Opin. Chem. Biol. 16 (3-4), 329-336.

8. Bazurto, J.V. *et al.* (2016) An unexpected route to an essential cofactor: *Escherichia coli* relies on threonine for thiamine biosynthesis. mBio 7 (1), e01840-15.

9. Mehta, V. *et al.* (2016) Efficiently functionalized oxacalix[4]arenes: synthesis, characterization and exploration of their biological profile as novel HDAC inhibitors. Bioorg. Med. Chem. Lett. 26 (3), 1005-1010.

10. Kim, J. *et al.* (2010) Three serendipitous pathways in *E. coli* can bypass a block in pyridoxal-5′-phosphate synthesis. Mol. Syst. Biol. 6.

11. Borchert, A.J. and Downs, D.M. (2017) The response to 2-aminoacrylate differs in *Escherichia coli* and *Salmonella enterica*, despite shared metabolic components. J. Bacteriol. 199 (14), e00140-17.

12. Downs, D. *et al.* (2018) The three-legged stool of understanding metabolism: integrating metabolomics with biochemical genetics and computational modeling. AIMS Microbiology 4 (2), 289-303.

13. Lambrecht, J.A. *et al.* (2012) Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. J. Biol. Chem. 287 (5), 3454-3461.

14. Lambrecht, J.A. *et al.* (2013) RidA proteins prevent metabolic damage inflicted by PLP-dependent dehydratases in all domains of life. mBio 4 (1), e00033-13-e00033-13.

15. Flynn, J.M. *et al.* (2013) Decreased coenzyme A levels in *ridA* mutant strains of *Salmonella enterica* result from inactivated serine hydroxymethyltransferase: *ridA* mutants are deficient in one carbon metabolism. Mol. Microbiol. 89 (4), 751-759.

16. Flynn, J.M. and Downs, D.M. (2013) In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5'-phosphate cofactor. J. Bacteriol. 195 (16), 3603-3609.

17. Hodge-Hanson, K.M. and Downs, D.M. (2017) Members of the Rid protein family have broad imine deaminase activity and can accelerate the *Pseudomonas aeruginosa* D-arginine dehydrogenase (DauA) reaction *in vitro*. PloS one 12 (9), e0185544.

18. Downs, D.M. and Ernst, D.C. (2015) From microbiology to cancer biology: the Rid protein family prevents cellular damage caused by endogenously generated reactive nitrogen species: RidA stress response. Mol. Microbiol. 96 (2), 211-219.

19. Ernst, D.C. and Downs, D.M. (2016) 2-aminoacrylate stress induces a context-dependent glycine requirement in *ridA* strains of *Salmonella enterica*. J. Bacteriol. 198 (3), 536-543.

20. Lambrecht, J.A. *et al.* (2010) Members of the YjgF/YER057c/UK114 family of proteins inhibit phosphoribosylamine synthesis *in vitro*. J. Biol. Chem. 285 (45), 34401-34407.

21. Christopherson, M.R. *et al.* (2008) YjgF is required for isoleucine biosynthesis when *Salmonella enterica* is grown on pyruvate medium. J. Bacteriol. 190 (8), 3057-3062.

22. Enos-Berlage, J.L. *et al.* (1998) Complex metabolic phenotypes caused by a mutation in *yjgF*, encoding a member of the highly conserved YER057c/YjgF family of proteins. J. Bacteriol. 180 (24), 6519-6528.

23. Ernst, D.C. and Downs, D.M. (2018) Mmf1p couples amino acid metabolism to mitochondrial DNA maintenance in *Saccharomyces cerevisiae*. mBio 9 (1).

24. Hodge-Hanson, K.M. *et al.* (2018) Expression of PLP-independent racemases can reduce 2-aminoacrylate stress in *Salmonella enterica*. J. Bacteriol., JB.00751-17.

25. Christopherson, M.R. *et al.* (2012) Suppressor analyses identify threonine as a modulator of *ridA* mutant phenotypes in *Salmonella enterica*. PLoS ONE 7 (8), e43082.

26. Schmitz, G. and Downs, D.M. (2004) Reduced transaminase B (IlvE) activity caused by the lack of *yjgF* is dependent on the status of threonine deaminase (IlvA) in *Salmonella enterica* Serovar Typhimurium. J. Bacteriol. 186 (3), 803-810.

27. Barak, Z. *et al.* (1987) Physiological implications of the specificity of acetohydroxy acid synthase isozymes of enteric bacteria. J. Bacteriol. 169, 3750-3756.

28. Vogel, H.J. and Bonner, D.M. (1956) Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. 218 (1), 97-106.

29. Balch, W.E. and Wolfe, R.S. (1976) New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. Appl. Environ. Microbiol. 32, 781-791.

30. Schmieger, H. (1972) Phage P22-mutants with increased or decreased transduction abilities. Mol. Gen. Genet. 119 (1), 75-88.

31. Chan, R.K. *et al.* (1972) Specialized transduction of tetracycline resistance by phage P22 in *Salmonella typhimurium*. II Properties of a high transducing lysate. Virology 50, 883-898.

32. Downs, D.M. and Petersen, L. (1994) *apbA*, a new genetic locus involved in thiamine biosynthesis in *Salmonella* typhimurium. J. Bacteriol. 176 (16), 4858-64.

33. Baba, T. *et al.* (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2.

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

35. Galloway, N.R. *et al.* (2013) Rapid cloning for protein crystallography using type IIS restriction enzymes. Crst. Growth Des. 13 (7), 2833-2839.

36. Burns, R.O. (1971) L-threonine deaminase — biosynthetic (*Salmonella typhimurium*). Methods Enzymol. 17B, 555-560.

37. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-54.

38. Davis, L. (1965) A spectrophotometric method for the assay of threonine dehydratase. Anal. Biochem. 12 (1), 36-40.

39. Burns, R.O. and Zarlengo, M.H. (1968) Threonine deaminase from *Salmonella* typhimurium. I. Purification and properties. J. Biol. Chem. 243 (1), 178-85.

40. Zarlengo, M.H. *et al.* (1968) Threonine deaminase from *Salmonella* typhimurium. II. The subunit structure. J. Biol. Chem. 243 (1), 186-91.

41. Chen, J.W. *et al.* (1991) Specificity of attenuation control in the *ilvGMEDA* operon of *Escherichia coli* K-12. J. Bacteriol. 173 (7), 2328-40.

42. LaRossa, R.A. and Van Dyk, T.K. (1987) Metabolic mayhem caused by 2-ketoacid imbalances. Bioessays 7, 125-130.

43. Hofler, J.G. and Burns, R.O. (1978) Threonine deaminase from *Salmonella typhimurium*. Effect of regulatory ligands on the binding of substrates and substrate analogues to the active sites and the differentiation of the activator and inhibitor sites from the active sites. J. Biol. Chem. 253 (4), 1245-51.

44. Bennett, B.D. *et al.* (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. Nat. Chem. Biol. 5 (8), 593-599.

45. Albe, K.R. *et al.* (1990) Cellular concentrations of enzymes and their substrates. J. Theor. Biol. 143 (2), 163-195.

46. Eisenstein, E. *et al.* (1995) An expanded two-state model accounts for homotropic cooperativity in biosynthetic threonine deaminase from Escherichia coli. Biochemistry 34 (29), 9403-9412.

47. Gallagher, D.T. *et al.* (2004) Local and global control mechanisms in allosteric threonine deaminase. Methods Enzymol. 380, 85-106.

48. Segel, I.H. and Segel, L.D. (1992) An alternative substrate is not the same as a dead end inhibitor. Biochem. Mol. Biol. Edu. 20 (3), 155-157.

49. Lal, P.B. *et al.* (2014) The redundant aminotransferases in lysine and arginine synthesis and the extent of aminotransferase redundancy in *Escherichia col*i: Aminotransferase redundancy. Mol. Microbiol. 94 (4), 843-856.

50. Toney, M.D. (2005) Reaction specificity in pyridoxal phosphate enzymes. Arch. Biochem. Biophys. 433 (1), 279-287.

51. Toney, M.D. (2011) Controlling reaction specificity in pyridoxal phosphate enzymes. Biochim. Biophys. Acta 1814 (11), 1407-1418.

52. Kuznetsova, I.M. *et al.* (2014) What macromolecular crowding can do to a protein. Int. J. Mol. Sci. 15 (12), 23090-140.

TABLE 4.1. IlvA activity and 2AA accumulation differs among *ilvA* mutants[a]

Genotype	Strain	IlvA Activity[b] ± SD	% IlvA Activity	IlvE Activity[c] ± SD	*ridA* effect
wild-type	DM9404	67 ± 14	100	101 ± 3	
ridA	DM3480	51 ± 3	76	53 ± 1	52
ilvA219 ($IlvA_{L447F}$)	DM6947	28 ± 1	42	103 ± 2	
ilvA219 ridA	DM6946	20 ± 3	30	27 ± 1	26
ilvA3210 ($IlvA_{A142T}$)	DM15898	7 ± 2	10	162 ± 16	
ilvA3210 ridA	DM15899	3 ± 1	4	136 ± 7	84
ilvA3211 ($IlvA_{G191S}$)	DM15936	4 ± 3	6	97 ± 7	
ilvA3211 ridA	DM15937	14 ± 2	21	114 ± 4	118

[a]Indicated strains were grown to stationary phase in minimal glucose medium containing 0.67 mM glycine.

[b]Threonine Deaminase (IlvA) acitivty was measure from crude lysate extracts. Data represent the mean and standard deviation (SD) from three biological replicates by measuring 2-ketobutyrate (KB) formation (nmol KB min^{-1} mg^{-1}). % IlvA activity represents the ratio of IlvA activity for each strain relative to wild-type (DM9404).

[c]Branched-chain amino acid aminotransferase (IlvE) activity was measured from crude lysate extracts. Data represent the mean and standard deviation from three biological replicates by measuring 2-keto-3-methylvalerate (KMV) formation (nmol KMV min^{-1} mg^{-1}). *ridA* effect represents the ratio of IlvE activity in strains without RidA compared to their $RidA^+$ parent and is a measure of relative 2AA accumulation.

TABLE 4.2. Estimated endogenous product synthesis rates for IlvA variants[a]

180 μM Threonine and 68 μM Serine				
Product	**IlvA**	**IlvA$_{L447F}$**	**IlvA$_{A142T}$**	**IlvA$_{G191S}$**
2-aminoacrylate	11 μM/s	23 μM/s	1 μM/s	1 μM/s
2-aminocrotonate	149 μM/s	210 μM/s	28 μM/s	16 μM/s
180 μM Threonine and 5000 μM Serine				
Product	**IlvA**	**IlvA$_{L447F}$**	**IlvA$_{A142T}$**	**IlvA$_{G191S}$**
2-aminoacrylate	736 μM/s	1467 μM/s	85 μM/s	29 μM/s
2-aminocrotonate	134 μM/s	183 μM/s	27 μM/s	8 μM/s
1000 μM Threonine and 68 μM Serine				
Product	**IlvA**	**IlvA$_{L447F}$**	**IlvA$_{A142T}$**	**IlvA$_{G191S}$**
2-aminoacrylate	10 μM/s	20 μM/s	1 μM/s	1 μM/s
2-aminocrotonate	730 μM/s	1030 μM/s	151 μM/s	69 μM/s

[a]Rates of 2-aminoacrylate and 2-aminocrotonate synthesis were estimated using the competing substrate concentrations shown and Equation 1. During calculation, IlvA active-site concentrations for all variants were provided as 10 μM [45]. Appropriate K_M and k_{cat} values for each variant enzyme were taken from Fig 3.

TABLE 4.3. Bacterial strains, plasmids, and primers

Strain	Genotype
DM9404	*Salmonella enterica* LT2 (Wild-type)
DM3480	*ridA3*::MudJ1734
DM6946	*ilvA219* ($IlvA_{L447F}$) *ridA3*::MudJ1734
DM6947	*ilvA219* ($IlvA_{L447F}$)
DM15898	*ilvA3210* ($IlvA_{A142T}$)
DM15899	*ilvA3210* ($IlvA_{A142T}$) *ridA3*::MudJ1734
DM15936	*ilvA3211* ($IlvA_{G191S}$)
DM15937	*ilvA3211* ($IlvA_{G191S}$) *ridA3*::MudJ1734
DM16371	*E. coli* BL21-AI *ΔilvA823*
DM16399	*E. coli* BL21-AI *ΔilvA823*/pDM1578
DM16406	*E. coli* BL21-AI *ΔilvA823*/pDM1420
DM16407	*E. coli* BL21-AI *ΔilvA823*/pDM1461
DM16408	*E. coli* BL21-AI *ΔilvA823*/pDM1543
Plasmid	**Description**
pDM1578	pET20b-*ilvA*
pDM1420	pET20b-$ilvA_{L447F}$
pDM1461	pSAPKO-CH-$ilvA_{A142T}$
pDM1543	pSAPKO-CH-$ilvA_{G191S}$
Primer	**Sequence**
IlvA_BspQI_LT2_F	5’-NNGCTCTTCNATGGCGGAATCTCAACCTCT-3’
IlvA_BspQI_LT2_R	5’- NNGCTCTTCNGTGACCCGCCAGAAAGAACC-3’
IlvA_NdeI_LT2_F	5’-GAGACATATGGCGGAATCTCAACCTCTGTC-3’
IlvA_XhoI_LT2_R	5’-GAGACTCGAGACCCGCCAGAAAGAAC-3’

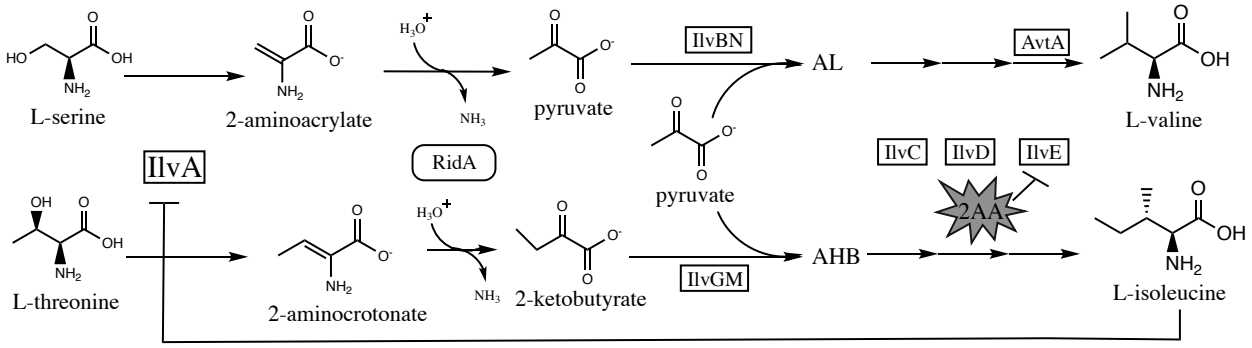


Fig 4.1. 2-aminoacrylate in the context of branched-chain amino acid metabolism. The pathways for branched-chain amino acid biosynthesis and 2-aminoacrylate hydrolysis are shown schematically. Enzymes are indicated next to the step they catalyze. IlvA catalyzes the first committed step in isoleucine biosynthesis using threonine as a substrate. IlvA can also dehydrate serine to form 2-aminoacrylate. RidA activity increases the rate of hydrolysis for both enamine products of IlvA. Elimination of RidA causes accumulation of 2-aminoacrylate, which can damage IlvE. (2AC would also accumulate, but is not known to have deleterious effects) Importantly, due to AvtA, absence of active IlvE prevents isoleucine but not valine biosynthesis. IlvA is sensitive to feedback inhibition by isoleucine, as indicated in the schematic. Intermediates in isoleucine and valine biosynthesis compete for enzymes IlvC, IlvD, and IlvE. Abbreviations: AL, 2-acetolactate; AHB, 2-aceto-3-hydroxybutyrate.

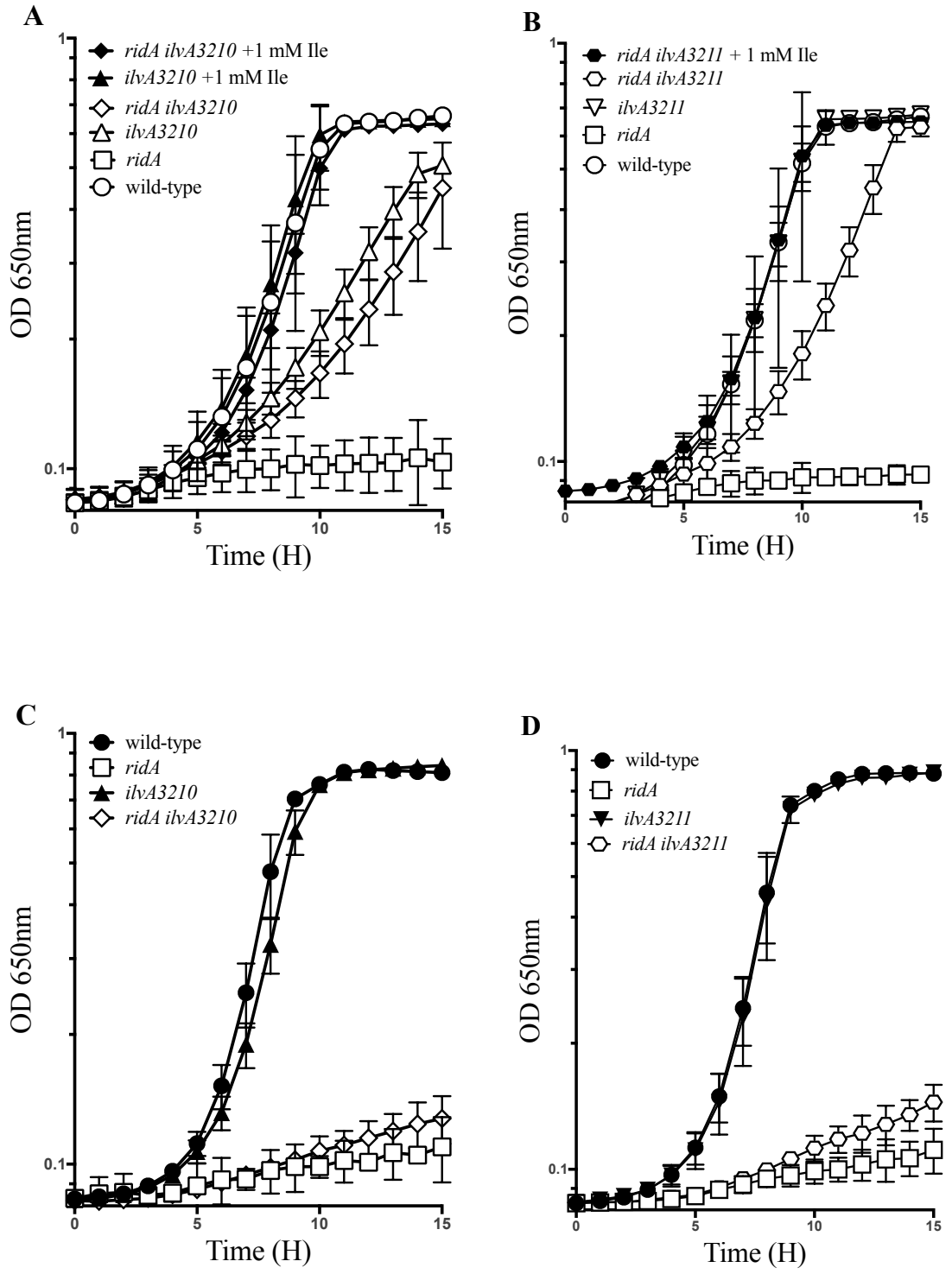
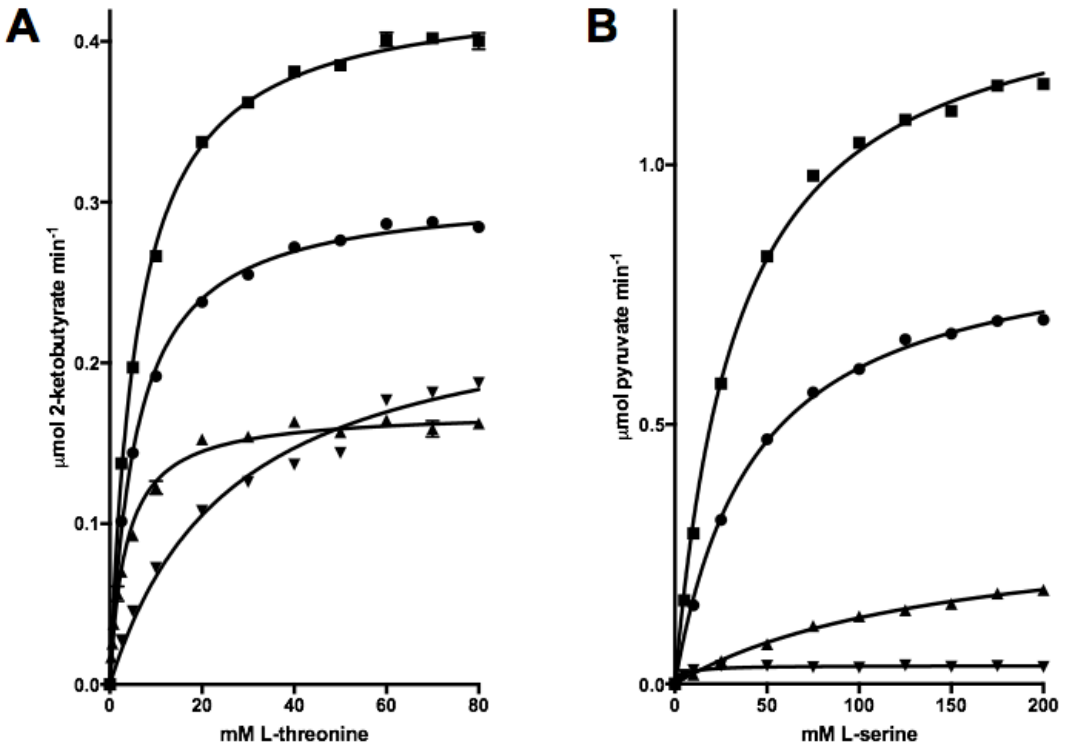


Fig 4.2. Mutations in *ilvA* restore growth to *ridA* mutants under some conditions. Strains were grown at 37°C in minimal medium containing 50 mM pyruvate as the sole carbon source (A, B), or with 11 mM glucose as the sole carbon source and supplemented with 5 mM L-serine (C, D). Growth is shown for wild-type (DM9404), *ridA* (DM3480), *ilvA3210* (DM15899), *ridA ilvA3210* (DM15898), *ilvA3211* (DM15936), and *ridA ilvA3211* (DM15937) as indicted in the legends. Data are the mean and error bars represent the 95% confidence interval of three biological replicates.



Variant	K_M (mM) *Threonine*	K_M (mM) *Serine*	k_{cat} (s^{-1}) *Threonine*	k_{cat} (s^{-1}) *Serine*	k_{cat}/K_M (s^{-1} mM^{-1}) *Threonine*	k_{cat}/K_M (s^{-1} mM^{-1}) *Serine*
IlvA	6 ± 1	43 ± 1	512 ± 2	726 ± 6	85	17
$IlvA_{L447F}$	6 ± 1	34 ± 1	722 ± 2	1148 ± 8	120	34
$IlvA_{A142T}$	26 ± 2	149 ± 7	407 ± 10	264 ± 7	16	2
$IlvA_{G191S}$	3 ± 1	5 ± 1	28 ± 1	6 ± 1	9	1

Fig 4.3. Kinetic characterization of IlvA variants. Saturation curves for IlvA (circles), $IlvA_{L447F}$ (squares), $IlvA_{A142T}$ (upward-facing triangles), and IlvAG191S (downward-facing triangles) measuring the initial rate of 2-ketobutyrate formation versus L-threonine concentration (A) or the initial rate of pyruvate formation versus L-serine concentration (B). Error bars represent the standard deviation of four replicates. The kinetic parameters used to plot the hyperbolic line describing the data (K_M and k_{cat}) are provided in a table containing the values and their respective standard deviations. The specificity constant (k_{cat}/K_M) was determined for each enzyme with each of the two substrates.

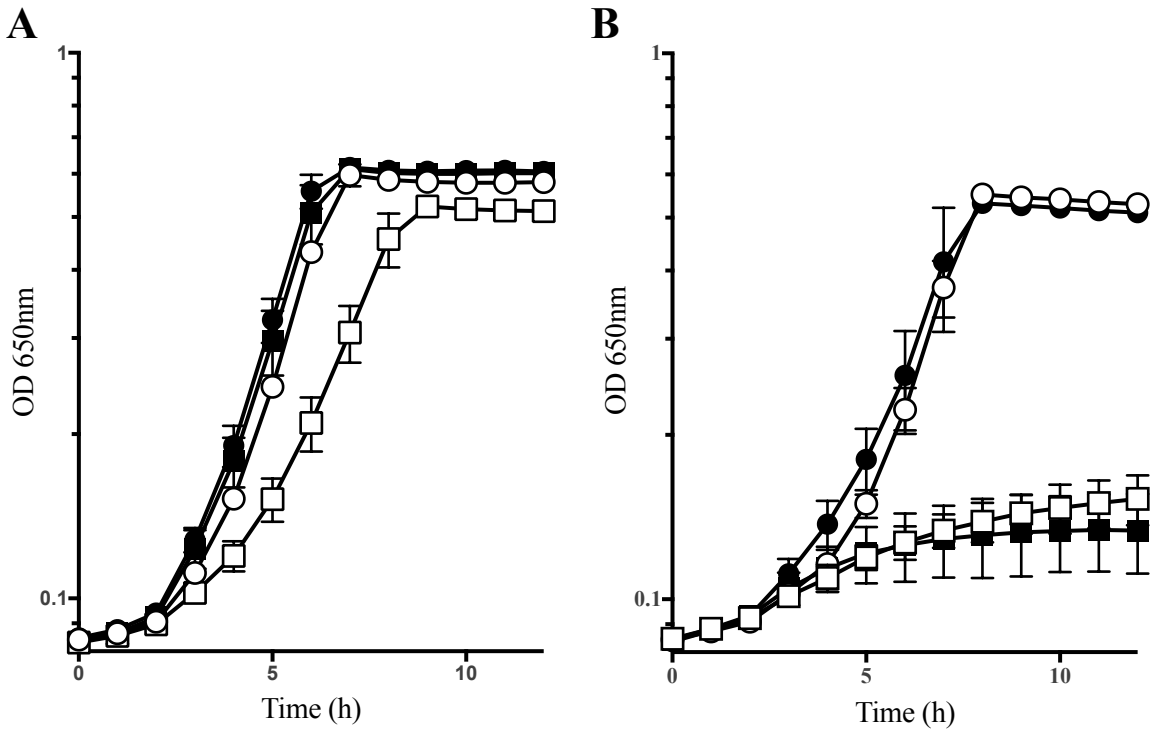


Fig 4.4. An *ilvA219* prevents effect of threonine on a *ridA* mutant growth. (A) Growth is shown for *S. enterica* wild-type (DM9404; circles) and *ridA* (DM3480; squares) strains. (B) Growth is shown for *S. enterica ilvA219* (DM6947; circles), and *ridA ilvA219* (DM6946; squares) strains. Growth analysis was performed at 37°C in minimal medium containing 11 mM glucose as the sole carbon source alone (open symbols), or supplemented with 1 mM L-threonine (open symbols). Data are the mean and error bars represent the 95% confidence interval of three biological replicates.

CHAPTER 5

ENDOGENOUSLY GENERATED 2-AMINOACRYLATE INHIBITS MOTILITY IN *SALMONELLA ENTERICA*[1]

[1]Borchert AJ, Downs DM. 2017. *Scientific Reports* 7(1): 12971. doi: 10.1038/s41598-017-13030-x. Reprinted here with permission of the publisher.

5.1 ABSTRACT

Members of the broadly distributed Rid/YER057c/UK114 protein family have imine/enamine deaminase activity, notably on a 2-aminoacrylate (2AA). Strains of *Salmonella enterica*, and other organisms lacking RidA, have diverse growth phenotypes, attributed to the accumulation of 2AA. In *S. enterica*, 2AA inactivates a number of pyridoxal 5'-phosephate(PLP)-dependent enzymes, some of which have been linked to the growth phenotypes of a *ridA* mutant. This study used transcriptional differences between *S. enterica* wild-type and *ridA* strains to explore the breadth of the cellular consequences that resulted from accumulation of 2AA. Accumulation of endogenously generated 2AA in a *ridA* mutant resulted in lower expression of genes encoding many flagellar assembly components, which led to a motility defect. qRT-PCR results were consistent with the motility phenotype of a *ridA* mutant resulting from a defect in $FlhD_4C_2$ activity. In total, the results of comparative transcriptomics correctly predicted a 2AA-dependent motility defect and identified additional areas of metabolism impacted by the metabolic stress of 2AA in *Salmonella enterica*. Further, the data emphasized the value of integrating global approaches with biochemical genetic approaches to understand the complex system of microbial metabolism.

5.2 BACKGROUND

Microbes inhabit environments that are often dynamic with respect to nutrient availability and the presence of chemical and/or physical stresses. A hallmark of microbes is their ability to respond to changing conditions by altering their metabolic network and/or physical characteristics to survive and thrive under new conditions[1, 2]. A common response to stresses and/or changing environmental conditions is the coordinated regulation of genes that ameliorate the stress and exploit changing nutrient circumstances[3-5]. Numerous examples of global responses to

environmental stress and nutrient availability that are mediated by transcriptional regulation have been described[1, 2]. Characterization of these responses has contributed to our understanding of not only metabolic strategies but mechanisms of regulation, effector binding, and gene function. Metabolism can be regulated at many levels, but perturbations in the metabolic network are ultimately expected to, directly or indirectly, impact the transcriptome. Thus, the global expression pattern of a microbe can be used to decipher the environment perceived by the cell and understand conditions both internal and external to the organism[6-8].

Metabolic imbalance caused by inhibition of a specific enzyme, for instance by exposure to antimicrobial agents or toxic metabolites, often result in transcriptional changes that are triggered as a consequence of the resultant metabolic imbalance[9-12]. In this scenario, if multiple insults generated similar or overlapping transcriptional profiles, it would indicate the perturbations ultimately affected the same part of metabolism. For instance, in a *Bacillus subtilis* strain harboring a *purA* mutation that diminished flux toward adenine biosynthesis, the cell senses a purine limitation and attempts to overcome it via de-repression of the PurR regulon[13]. Treatment of *Bacillus subtilis* with the dihydrofolate reductase (FolA; EC 1.5.1.3) inhibitor, Trimethoprim, causes a disruption of C_1 unit transfer in the cell, resulting in a bottleneck in purine biosynthesis at the AICAR transformylase (PurH; EC 2.1.2.3) enzyme. In this case, the cell perceives a purine limitation and attempts to overcome it via de-repression of the PurR regulon[14]. Thus, the shared features of the transcriptional response to both perturbations indicated that the perceived stress in each case was purine limitation.

In *Salmonella enterica,* accumulation of the reactive enamine species 2-aminoacrylate (2AA) perturbs the metabolic network by partially inactivating a number of PLP-dependent enzymes, which results in growth phenotypes[15-19]. 2AA is generated as a mechanistic

intermediate of multiple PLP-dependent enzymes in central metabolic pathways (e.g., serine/threonine dehydratases EC 4.3.1.19)[20, 21]. Thus far, the validated targets of 2AA *in vivo* are PLP enzymes where the enamine attacks the internal aldimine and results in the covalent modification of the PLP cofactor. Depending on the mechanism of attack, an inactive pyruvate-PLP adduct is formed in the active site, or the enzyme itself forms an irreversible bond with the modified PLP[17, 19, 22]. In either scenario, the affected enzyme is irreversibly damaged, and flux through the pathway to which it belongs is constrained. A summary of the described RidA paradigm of 2AA stress in *S. enterica* is provided in Figure 5.1.

Members of the RidA subfamily, belonging to the Rid protein family, are enamine deaminases that quench 2AA and thus prevent the accumulation of this metabolite *in vivo*[16, 23]. The combination of *in vitro* and *in vivo* studies in *S. enterica* showed that at least the PLP-dependent enzymes serine hydroxymethyltransferase (GlyA; EC 2.1.2.1), branched-chain amino acid aminotransferase (IlvE; EC 2.6.1.42) and alanine racemase (Alr/DadX; EC 5.1.1.1) were targets of free 2AA *in vivo* and their activity is between 40-80% decreased in strains lacking RidA[17, 19, 22, 24, 25]. The presence of RidA proteins across all domains of life supports a conserved role for this family in quenching reactive metabolites, including 2AA, to prevent damage to metabolic components.

The current understanding of 2AA targets *in vivo* is the result of biochemical genetic studies, primarily in *S. enterica*. Although powerful, this approach demanded that the damage caused by 2AA generated a detectable growth phenotype. It was therefore likely that the impact of free 2AA on the metabolic network was broader than what has been defined by this approach. Additionally, the impact that 2AA damage of known targets has on the global network configuration remains unexplored. For example, while it is known that the growth defect of an *S.*

enterica ridA mutant grown in minimal medium is caused by 2AA damage to serine hydroxymethyl transferase (GlyA; EC 2.1.2.1), it is unclear how the cell perceives the resulting reduction in C_1 unit production[24]. The current study was initiated to generate a global view of metabolic changes that result from the accumulation of 2AA in *S. enterica.* The premise of this work was that a transcriptional profile reflects expression changes that result, directly or indirectly, from perceived changes in the internal or external environment[6, 8, 26].

5.3 MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Primers. Strains and plasmids used in this work were derivatives of *Salmonella enterica* subsp. *enterica* serovar Typhimurium LT2 and are listed in Appendix Table A.2. Transductional crosses were carried out using the high-frequency general transducing mutant of bacteriophage P22 (HT105/1, *int*-201)[27]. New gene disruptions constructed during this study were made using the λ-Red recombinase system described by Datsenko and Wanner[28] using the primers listed in Appendix Table A.3, transduced into relevant genetic backgrounds by selecting chloramphenicol resistance, and verified by PCR. Plasmid pDM1439 was constructed using the high-efficiency cloning method[29] to clone the LT2 *ridA* gene into plasmid pCV1. The resulting construct was confirmed by sequence. pCV1 is a modified pBAD24 plasmid containing BspQ1 sites, conferring ampicillin resistance, and the cloned gene's expression is inducible by L-(+)-arabinose[30]. Gene-specific primers for cloning and qRT-PCR analysis were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA) and are listed in Appendix Table A.3. Primers were designed using Primer 3 software, and evaluated for specificity by melting curve analysis using *S. enterica* LT2 gDNA prior to running the qPCR experiment.

Culture Media and Chemicals. Minimal medium was no-carbon E medium (NCE) supplemented with 1 mM MgSO4[31], trace minerals[32], and 11 mM D-glucose as the sole

carbon source. L-(+)-arabinose was added to a final concentration of 0.2 g/L when needed. Difco nutrient browth (NB) (8g/L) containing NaCL (5 g/L) was used as rich medium. Difco BiTek agar (15g/L) was added for solid medium. Ampicillin was added when required for plasmid maintenance to rich and minimal medium at 150 and 7.5 µg/mL, respectively. Chloramphenicol (20 µg/mL) and Ampicillin (150 µg/mL) were used in rich medium, during strain construction. All chemicals were obtained from Sigma-Aldrich Chemical Company (St. Louis, MO).

RNA isolation. Relevant strains were grown overnight with shaking at 37°C in quadruplicate (RNA-seq) or triplicate (qRT-PCR) in 2 mL NB. Following incubation, strains we diluted 1:100 into 5 mL fresh minimal medium, containing any additional supplements indicated in the text, and allowed to grow with shaking at 37°C to an optical density at 650 nm (OD_{650}) of 0.6. Each 5 mL sample was quickly centrifuged at 16,000 x *g* in 1.5 mL Eppendorf tubes, supernatant removed and pellets flash-frozen in liquid nitrogen and kept on dry ice. Total RNA was extracted using the RNA*snap*™ method[33]. Pellets were resuspended in 110 µl boil solution (95% [vol/vol] RNA-grade formamide, 18 mM EDTA, 0.025% [wt/vol] SDS, 1% 2-mercaptoethanol in UltraPure™ (ThermoFisher) distilled water). Samples were incubated at 95°C for 7 min before centrifuging at 16,000 x *g* for 5 min while still hot and 100 µl supernatant was transferred to a new 1.5 mL Eppendorf tube. 300 µl UltraPure™ water and 40 µl sodium acetate (3M [pH5.2]) was added and mixed into each sample before adding 400 µl ice-cold ethanol (100%) and incubating on ice 15 min. Samples were centrifuged at 16,000 x *g* for 15 min at 4°C and ethanol was decanted. 400 µl ethanol (70% [vol/vol]) was added and pellets were immediately centrifuged at 12,000 x *g* for 10 min at 4°C before decanting ethanol and pellets were allowed to dry 20 min at 25°C. DNA/RNA pellets were resuspended in 90 µl UltraPure™ water at 4°C overnight prior to treatment with RNase-free Turbo DNase (Ambion), and being precipitated once

more by sodium acetate-ethanol treatment. Resulting RNA was resuspended at 4°C for 3 h and stored at -80°C until use.

RNA-Seq expression profiling and data analysis. The total RNA was extracted from four independent replicates of *S. enterica* wild-type and *ridA* mutant strains. In each of the resulting eight samples, rRNA was removed from 5 µg total RNA using the Ribo-Zero rRNA removal kit (Bacteria; Illumina) following manufacturer's recommendations. A cDNA library was created from samples containing < 10% rRNA contamination, as measured by the RNA 6000 pico kit for the Agilent 2100 bioanalyzer, and sequenced by the University of Georgia Genomics Facility (GGF) from these RNA samples. Specifically, one hundred nanograms of the rRNA-free RNA was used to create each sequencing library using the KAPA stranded RNA-seq kit (KAPA Biosystems). The RNA libraries were sequenced on an Illumina NextSeq500 (150 cycles) Mid Output Flowcell in the paired-end mode with a read length of 75bp. Processing and differential expression determination of the sequencing data were performed by the Georgia Advanced Computing Resource Center (GACRC) at the University of Georgia. Adaptor removal and quality trimming (Phred score, ~20) of the raw reads were done using FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/) and Trimmomatic (Usadel). The trimmed reads were mapped to Salmonella enterica subsp. Enterica serovar Typhimurium str. LT2 chromosome, complete genome (RefSeq Accession Number: NC_003197.1) and read counts were generated using the prokaryotic-tailored gene alignment tool EDGE-Pro[34]. Genes differentially expressed (false discovery rate, FDR < 0.05) between the respective wild-type and *ridA* strains of *S. enterica* were identified using DEseq[35-37]. Briefly, an FDR < 0.05 cutoff controls the expected proportion of falsely discovered differentially expressed genes (false positives) to 5% (21 genes) of the identified positive hits meeting this threshold (413 genes). The procedures used

by DEseq software to determine the FDR value for each gene is described by Benjamini and Yekutieli[35].

qRT-PCR. The total RNA from three independent replicates of each *S. enterica* strain tested, as described in the text, was extracted. An aliquot of each sample was sent to GGF for quality control analysis and quantification using the RNA 6000 nano kit for the Agilent 2100 bioanalyzer and only samples with an RNA integrity number (RIN) over 5.0 were used[38]. DNase-treated RNA (600ng) was subjected to first strand cDNA synthesis, using the iScript cDNA synthesis kit (Bio-Rad Laboratories) according to the manufacturer's protocol. PCRs were performed in an Applied Biosystems 7500 Fast real-time (RT) PCR system. Each reaction was carried out in a total volume of 20 µl on a 96-well optical reaction plate (Applied Biosystems) containing 10 µl FastStart Universal SYBR green Master (ROX) mix (Roche Applied Science), 8 ng cDNA, and two gene-specific primers at a final concentration of 0.5 µM each. The real-time cycling conditions were as follows: 95°C for 20s, and 40 cycles of 95°C for 3s and 60°C for 30s. Melting-curve analysis verified that each reaction mixture contained a single PCR product. The threshold cycle (CT) values of *gyrB* and *rpoB* were used as internal controls[39, 40]. Fold changes were calculated using the comparative threshold cycle (ΔΔCT) method[41]. Briefly, the fold-change in gene expression was calculated using the formula: *ridA*/wild-type = $2^{\Delta\Delta CT}$, where $\Delta\Delta CT = \Delta CT_{ridA} - \Delta CT_{\text{wild-type}}$ and $\Delta CT = CT_{\text{target gene}} - CT_{\text{normalization gene (}gyrB\text{)}}$. Gaussian error propagation was used to determine the standard error of the mean (SEM) for ΔΔCT from the ΔCT SEM values, and this value used to calculate the 95% confidence interval for each gene. Welch's two-tailed t-test was performed on $\log_2$ transformed $2^{\Delta\Delta CT}$ values, using GraphPad Prism 6.0b software. To ensure *gyrB* was constant under the conditions tested, the relative fold change for the internal control *rpoB*, whose expression was expected to remain constant across treatments, was calculated

as described above. Only those samples with an *rpoB* fold change between 0.8 and 1.25 were used.

Motility Assays. Motility was monitored on minimal medium motility plates (0.25% agar) using either D-glucose (11 mM) or glycerol (20 mM) as the sole carbon source. The plates were prepared in the morning and allowed to solidify at room temperature un-stacked on the benchtop ~6 h before use in the evening. *S. enterica* cells were grown to full-density in LB and 1 µl of the culture was inoculated onto the plate and incubated at 30°C inside of a sealed 6 quart plastic Sterilite container to maintain constant moisture content. Motility halos were measured after 20 h using a ruler with markings every mm; if the halo diameter fell between mm markings, the recorded halo diameter was rounded to the nearest marking. Technical replicates (i.e., plates) were prepared from each of four biological replicates and the data were averaged and expressed as mean (mm) ± SD.

Data Availability. Sequencing data were deposited at the NCBI Gene Expression Omnibus (GSE103146), and can be found at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE103146. All other datasets generated in the current study are available from the corresponding author on reasonable request.

5.4 RESULTS

Deletion of *ridA* alters the global transcription profile. The transcriptomes of a *S. enterica ridA* mutant and the isogenic wild-type strain were defined after growth on minimal glucose medium to provide a snapshot of the global metabolic state of the cells. The transcriptome data were viewed with two questions in mind; i) are previously characterized metabolic defects caused by elevated 2AA discoverable, and ii) do the data generate new insights about the global impact of 2AA accumulation?

Pair-wise comparison of gene expression identified 413 genes with changes (186 higher expression, and 227 lower expression) in transcript abundance between *ridA* mutant and wild-type strains, using a false discovery rate (FDR) less than 0.05 (Appendix Table A.1)[35]. Of these loci, genes with no known function (STMxxxx and *yxxX* locus tags) and those that showed less than a two-fold difference in expression were excluded to generate the dataset of 84 genes shown in Figure 5.2. There was no evidence that the biosynthetic genes for isoleucine (*ilv*) were de-repressed in a *ridA* mutant, despite the fact that the specific activity of isoleucine transaminase (IlvE, the last step in isoleucine synthesis) is more than 50% reduced in the absence of RidA[17, 18]. These data corroborated the conclusion reached with a single reporter, that lack of RidA did not result in a starvation for isoleucine[18]. Similarly, no differential expression was simply attributable to the decreased activity of serine hydroxymethyl transferase (GlyA) found in *ridA* mutants. The transcription of *dadAX* was increased in the *ridA* mutant and potentially reflected a consequence of the decreased alanine racemase activity of DadX caused by 2AA accumulation[22], since increased accumulation of L-alanine may contribute to the complex transcriptional regulation of *dadAX*[42].

Results from the global RNA-seq experiment were independently validated with real-time qRT-PCR. The relative expression level of six select genes was measured using the comparative threshold cycle ($\Delta\Delta CT$) method[41] (Fig. 5.3). Four of the six genes chosen met the two-fold expression threshold in the RNA-seq experiment (*napF*, *metH*, *dadX*, and *fliI*), one (*thiF*) had a statistically significant difference in expression that was less than two-fold (1.88-fold, Appendix Table A.1), and the last (*sdaC*) showed no statistically significant change in gene expression between mutant and wild-type. For the six genes, a strong positive correlation between results from the qRT-PCR and global RNA-seq approaches was supported by a Pearson's correlation

coefficient of $r^2 = 0.96$. Of the four genes with a greater than two-fold expression difference, only *napF* had a statistically significant (p-value < 0.05) discrepancy between the RNA-seq and qRT-PCR data. However, both approaches detected a greater than ten-fold decrease in the *napF* transcript levels in a *ridA* mutant compared to wild-type. Finally, although RNA-seq transcriptome analysis detected increased *sdaC* expression in a *ridA* strain, the expression difference did not meet the statistical cut off (FDR < 0.05). The qRT-PCR data similarly found this transcript increased in a *ridA* mutant, but the difference was statistically significant in this technique, suggesting that the DEseq algorithm, used to call differentially expressed genes from the RNAseq dataset, performed conservatively[37].

Features of the transcriptome data beyond metabolic gene expression were noted. For instance, multiple genes in the salmonella pathogenicity islands (SPIs) were differentially expressed in the two strains. Genes in SPI1 and SPI5 generally had lower expression and those in SPI2 had higher expression in the *ridA* mutant compared to the wild-type strain, which was reminiscent of the expression pattern during *S. enterica* serovar Typhimurium replication within murine macrophages[7]. This similarity was extended with the decrease in expression of the flagellar biosynthetic operons and several genes involved in chemotaxis in the *ridA* mutant.

Transcription profiling uncovers a role for RidA in motility. Without exception, expression of the genes involved in assembly of flagella was more than two-fold lower in a *ridA* mutant than wild-type strain. Expression of genes involved in chemotaxis, including *cheM*, *cheA* and *cheY*, was also decreased in a *ridA* mutant (2.34, 1.68, 1.78, -fold, respectively). This level of differential expression can generate a detectable motility defect[43], and the swimming motility of the wild-type and *ridA* mutant strains was tested. As predicted by the transcript levels, the *ridA* mutant was less motile than the wild-type strain on minimal medium (Fig. 5.4). Quantification of

the zone of swimming showed motility in *ridA* mutants was ~40% less than wildtype (Table 5.1), and was restored to wild-type by the expression of *ridA in trans* (Table 5.1, ln 5-8). It was formally possible that the slight growth defect of a *ridA* mutant on minimal medium was responsible for the *ridA* decrease in motility. The addition of glycine, which restored wild-type growth to a *ridA* strain[24], failed to restore full motility to the mutant (Table 5.1, ln 3-4). This result supported the conclusion that reduced motility reflected a new consequence of a *ridA* mutation.

Accumulation of 2-aminoacrylate (2AA) generates a motility defect. Metabolic phenotypes of a *ridA* mutant have been attributed to the accumulation of 2-aminoacrylate, which is generated by the serine/threonine dehydratase IlvA, using serine as a substrate[17]. Three lines of evidence supported the role of IlvA generated 2AA in compromising motility of a *ridA* mutant (Table 5.2). First, exogenous threonine restored motility (ln 1-4), presumably by outcompeting serine for the active site of IlvA, as described for other *ridA* mutant phenotypes [16, 44]. Second, addition of isoleucine restored motility (ln 5,6), presumably by allosteric inhibition of IlvA[45], that decreased the formation of 2AA[25]. The latter conclusion was further supported by strains carrying the *ilvA219* allele, which encodes an $IlvA_{L447F}$ variant that is resistant to allosteric inhibition by isoleucine[9]. When the *ilvA219* allele replaced wild-type *ilvA*, data in Table 5.2 (ln 7-10) showed that with or without isoleucine, a *ridA* mutation decreased motility by ~35%. Finally, by increasing the production of 2AA, the motility was further decreased (Table 5.2, ln 11-16). In strains carrying a 2AA generating system (*ABcc1*), there are two copies of *ilvA,* the wild-type allele in the chromosome and the *ilvA219* allele under the control of an inducible P_{BAD} in the chromosome[46]. Isogenic *ridA* and *ridA+* strains carrying the *ABcc1* construct were generated. In this experiment, without induction of *ilvA219,* the *ridA* mutant had 20% less motility than the wild-type strain. When expression of *ilvA219* was induced, the *ridA* mutant had a greater than 40%

defect in relative motility. Addition of isoleucine partially restored motility, consistent with the presence of the wild-type IlvA, which is sensitive to allosteric inhibition. Taken together the data support the conclusion that like previously characterized phenotypes, the effect of the *ridA* mutation on motility is due to the accumulation of 2AA generated by IlvA.

Lowered $FlhD_4C_2$ activity can explain reduced motility but not SPI expression. Biosynthesis of the *S. enterica* flagellar apparatus is controlled by a well-characterized regulatory cascade, and expression/activity of the master regulator $FlhD_4C_2$ is tightly regulated. Figure 5.5 schematically represents the relevant genes and their general position in this cascade. The *ridA* mutant had decreased expression of several operons with class II promoters that are directly, or indirectly, activated by the $FlhD_4C_2$ master regulator. These genes included *napFDAG, flg* and *fli* genes and those in SPI I, as indicated in Figure 5.2[47, 48]. A simple scenario would suggest the function of the $FlhD_4C_2$ activator complex was reduced in a *ridA* mutant, resulting in the decreased expression of its regulon. However, the expression of *flhD* and *flhC* were significantly (FDR < 0.05) higher in the *ridA* strain (1.64, 1.58 –fold, respectively). These two results could be reconciled if the *ridA* mutation generated a post-transcriptional effect that resulted in differential function of $FlhD_4C_2$ in the two strains.

If accumulation of 2AA impacted one of the characterized regulators of *flhDC* translation or $FlhD_4C_2$-dependent promoter binding, elimination of the respective regulatory component would nullify the effect of the *ridA* mutation. Lesions in the small regulatory RNAs *omrAB, micA, arcZ, oxyS,* as well as *rpoS*, each of which post-transcriptionally regulate $FlhD_4C_2$, were introduced to isogenic *ridA* strains. The motility of the resulting strains was determined, and the data are presented in Table 5.3. Mutations in *arcZ* and *rpoS* caused no detectable defect in motility compared to wild-type and did not impact the effect of the *ridA* mutation (Table 5.3, ln 5-6, 13-

14). Mutations in *micA, omrBA,* or *oxyS* decreased motility slightly (~20-25%) with a *ridA* mutation further decreasing the motility by ~30% (Table 5.3, ln 3-4, 7-10). The proteins YdiV and FliZ also regulate $FlhD_4C_2$ activity post-transcriptionally. YdiV acts on $FlhD_4C_2$ to disrupt promoter binding[49] (Fig. 5.5). FliZ can repress *ydiV* expression, indirectly activating $FlhD_4C_2$ function[50]; as well as increase $FlhD_4C_2$ protein levels through an unknown YdiV-independent mechanism[51] (Fig. 5.5). In the absence of FliZ, expression of genes activated by $FlhD_4C_2$ is reduced ~ 2-fold[43], similar to the magnitude resulting in the *ridA* mutant (Fig. 5.2). Deletion of *fliZ* decreased motility of *S. enterica* by ~ 43%. Inclusion of a *ridA* mutation resulted in a strain that had 33% the motility of the parental *fliZ* strain (Table 5.3, ln 15-16) and suggested these two mutations had independent (additive) effects on motility. Deletion of *ydiV* slightly (~7%) increased motility of wild-type *S. enterica* and the impact of a *ridA* mutation introduced into this background was consistent with the ~30% decrease repeatedly observed (Table 5.3, ln 11-12). Taken together these data did not support a role for 2AA acting via a known regulator of $FlhD_4C_2$ translation or promoter binding of the complex.

To investigate the scope of *flhDC* involvement of the differential expression of motility, chemotaxis, and SPI-encoded genes in a *ridA* mutant, qRT-PCR compared the expression of representative loci (*napF*, *fliI*, *fliC*, *cheM*, *invA*, *ssaV*, and *sopB*) between a *ridA flhDC::cat* mutant (DM15818) and *flhDC::cat* wild-type (DM15817) (Fig. 5.6). Welch's two-tailed t-test revealed that elimination of *flhDC* significantly (p-value < 0.05) changed the relative expression (*ridA*:wild-type) of the *napF*, *fliI*, and *cheM* genes. Strikingly, expression of *fliC* was not significantly ($p < 0.05$) different between the *flhDC* and *flhDC ridA* strains. These data supported a model that a single perturbation affecting $FlhD_4C_2$ activity, directly or indirectly, could be responsible for the expression pattern of genes involved in flagellar biosynthesis, nitrate reduction, and chemotaxis.

It is formally possible that 2AA interacts directly with $FlhD_4C_2$ to reduce its activity, but since free 2AA is not known to damage enzymes lacking a PLP cofactor, a more likely model suggests a disruption of the metabolic network to which a post-transcriptional regulator of *flhDC* responds.

Exogenous glycine overrides *ridA* mutant effect on SPI expression. The *flhDC* mutation did not alter the influence a *ridA* mutation had on transcription of genes located within SPI1 (*invA*), SPI2 (*ssaV*), and SPI5 (*sopB*). These data suggested that the accumulation of 2AA, caused by the *ridA* mutation, affected expression of genes in SPI1, SPI2, and SPI5 by an $FlhD_4C_2$-independent mechanism. Exogenous glycine eliminates many growth phenotypes of a *ridA* mutant because it bypasses the need for serine hydroxymethyl transferase (GlyA), which is a primary target of 2AA in *S. enterica*[24]. qRT-PCR compared the expression of relevant loci (*napF*, *fliI*, *fliC*, *cheM*, *invA*, *ssaV*, and *sopB*) between *ridA* (DM3480) and wild-type (DM9404) strains grown on minimal medium supplemented with 1 mM glycine (Fig. 5.6, dark grey bars). Importantly, based on Welch's two-tailed t-test, the relative expression (*ridA*:wild-type) of the *napF*, *fliI*, *cheM*, *invA*, and *sopB* loci was significantly different (p-value < 0.05) when the strains were grown in the presence vs. absence of glycine in the medium. Further, when considered as an isolated experiment, expression of *cheM*, *invA*, *ssaV*, and *sopB* was not significantly (p-value < 0.05) different between the *ridA* and wild-type backgrounds, when the cells were grown in medium with glycine. Taken together, these data suggested that exogenous glycine eliminated the effect of a *ridA* mutation on the expression of *cheM*, *invA*, *ssaV*, *and sopB*, but *napF*, *fliI*, and *fliC* expression was still significantly (p-value < 0.05) reduced in the *ridA* mutant. The latter result was consistent with the finding that glycine did not completely restore motility to the *ridA* mutant (Table 5.1, ln 3-4). The effect a *ridA* mutation had on the transcription of *napF*, *fliI*, and *cheM* was affected by

both an *flhDC* deletion and glycine addition, suggesting both $FlhD_4C_2$ activity and glycine/5,10-methylene-tetrahydrofolate influence the expression of these genes[19, 24].

5.5 DISCUSSION

Differential gene expression was used to probe the global consequences of eliminating the RidA protein in *S. enterica.* The transcription dataset described herein made a few points. First, the biochemical defects previously identified in *ridA* mutants, with the possible exception of DadX, did not alter the transcription of genes in the respective metabolic pathways affected. This result was consistent with genetic data that *ridA* mutants have compromised enzyme activity but, in the absence of further perturbations, maintain a robust and functional metabolism. This observation is consistent with the capacity of microbes to adapt to environment changes by maintaining a robust metabolic network.

Second, the expression pattern of *ridA* vs. wild-type paralleled the expressional changes found with cells replicating in a macrophage, compared to the same strain grown *in vitro*[7]. Similar expressional changes were seen for the genes associated with nitrate reduction, methionine biosynthesis, chemotaxis, flagellar biosynthesis, and genes encoded on SPI1, SPI2, and SPI5 (Fig. 5.2). These results hint at similarities between the metabolic state of *S. enterica* cells accumulating 2AA and what the cells face during infection of a host. This similarity was emphasized by the coordinated response of genes encoded by the three different SPIs, which are all required for Salmonella enteropathogenicity[52, 53]. Serine hydroxymethyl transferase (GlyA) is compromised, (i.e., glycine and C_1 units are reduced) in strains lacking *ridA*[24]. By adding glycine and restoring one-carbon unit abundance in a *ridA* mutant, the differential expression of these SPI loci was minimized. Further work will be needed to extend this correlation and determine whether the *ridA* paradigm can provide a model to extend the understanding of the intracellular

environment and the factors influencing the response of *S. enterica* to the intracellular environment of the macrophage.

A third feature of the data was the reduced expression of flagellar biosynthesis genes, chemotaxis genes, and periplasmic nitrate reductase (*nap*) genes in a *ridA* mutant. Consistent with the expression pattern, *ridA* mutants had decreased motility, when compared to the parental strain. As with previously characterized phenotypes of a *ridA* mutant, the lowered motility was a consequence of 2AA accumulation, although the specific target and resulting mechanism were not clear. The only targets of 2AA characterized thus far are PLP-dependent enzymes, where the reactive 2AA enamine attacks the PLP in an active site to lock the respective enzyme in an inactive state[17, 19, 22]. There are no PLP-dependent enzymes directly involved in motility, suggesting either a new class of targets for 2AA, or an indirect response to a targeted PLP-dependent enzyme(s). Apart from 2AA deamination activity, previous work suggests a role for *E. coli* RidA in acting as a chaperon to protect against reactive oxygen species (ROS) toxicity, specifically hypochlorite[54]. ROS stress is a known signal for repression of the flagellar biosynthesis machinery[55], and so it is plausible that the proposed chaperon role of RidA may be implicated in the motility defect observed in this study. Since increasing or decreasing 2AA production served to exacerbate or eliminate the motility defect of a *ridA* mutant, respectively (Table 5.2), we believe that it is 2AA accumulation and not the elimination of a RidA chaperon activity that lead to the motility defect observed for a *ridA* mutant. The master activator, $FlhD_4C_2$, controls a significant number of the genes that are differentially regulated in a *ridA* mutant, including the periplasmic nitrate reductase (*nap*) associated genes. The qRT-PCR data showed that the transcriptional remodeling in a *ridA* mutant is multifaceted, even when only considering pathogenesis-associated genes (Fig. 5.5). Significantly, the differential expression of the *nap* operon, chemotaxis genes,

and flagellar biosynthesis genes in a *ridA* mutant required an intact *flhDC* locus, while the differences in SPI gene expression were *flhDC*-independent. The *flhDC*-dependent link between the effect of 2AA on the *nap* operon and motility-associated gene expression suggested a need to coordinate nitrate reduction and motility[56]. Further work needs to confirm whether coordination of nitrate reduction and motility is an important physiological trait for the Salmonella lifecycle.

In total, the comparative transcription analyses herein identified a new motility phenotype for mutants lacking *ridA*. This study emphasized the ability of transcriptome analyses to provide global metabolic snapshots that may not result in a detectable growth phenotype. It is likely that some of the genes that failed to meet the arbitrary two-fold threshold will prove to be physiologically relevant. These data provided insights into the endogenous environment resulting from the accumulation of 2AA in Salmonella and uncovered areas of metabolism that may be targeted by metabolic stress caused by 2AA. Identifying additional enzymes that are affected by 2AA will require targeted genetic and biochemical approaches that are driven by global datasets, like the one described here. Importantly, since a 2AA-quenching role has been demonstrated for all RidA subfamily members tested, and RidA is conserved across all domains of life, conclusions drawn from Salmonella, regarding 2AA accumulation, may extend to other Enterobacteriaceae and even a wider set of organisms[16, 23]. Two examples for extension of the RidA paradigm from Salmonella into other organisms have already been seen[46, 57].

Acknowledgements. This work was supported by competitive grant GM095837 from the National Institutes of Health (DMD). The authors gratefully acknowledge the valuable discussions with Magdy Alabady of the Georgia Genome Facility and Walt Lorenz of the Georgia Advanced Computing Resource Center. We thank Jorge Escalante-Semerena for use of the Applied Biosystems 7500 Fast RT-PCR system.

5.6 REFERENCES

1. Bailey, A.M. *et al.* (2009) Exposure of *Escherichia coli* and *Salmonella enterica* serovar Typhimurium to triclosan induces a species-specific response, including drug detoxification. J. Antimicrob. Chemother. 64 (5), 973-985.

2. Shimizu, K. (2015) Metabolic regulation and coordination of the metabolism in bacteria in response to a variety of growth conditions. In Bioreactor Engineering Research and Industrial Applications I (Ye, Q. et al. eds), pp. 1-54, Springer Berlin Heidelberg.

3. Dandekar, T. *et al.* (2015) *Salmonella* — How a metabolic generalist adopts an intracellular lifestyle during infection. Front. Cell. Infect. Mi. 4 (January), 1-11.

4. Blair, J.M.A. *et al.* (2013) Choice of bacterial growth medium alters the transcriptome and phenotype of *Salmonella enterica* serovar Typhimurium. PLoS ONE 8 (5), e63912.

5. Ishii, N. *et al.* (2007) Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. Science 316 (5824), 593-597.

6. Alves, P.C. *et al.* (2016) Transcriptomic and metabolomic profiling of ionic liquid stimuli unveils enhanced secondary metabolism in *Aspergillus nidulans*. BMC Genomics 17 (1).

7. Srikumar, S. *et al.* (2015) RNA-seq brings new insights to the intra-macrophage transcriptome of *Salmonella* Typhimurium. PLoS Pathog. 11 (11), e1005262.

8. Tognetti, V.B. *et al.* (2010) Perturbation of indole-3-butyric acid homeostasis by the UDP-glucosyltransferase UGT74E2 modulates *Arabidopsis* architecture and water stress tolerance. Plant Cell 22 (8), 2660-2679.

9. LaRossa, R.A. *et al.* (1987) Toxic accumulation of alpha-ketobutyrate caused by inhibition of the branched-chain amino acid biosynthetic enzyme acetolactate synthase in *Salmonella typhimurium*. J. Bacteriol. 169 (4), 1372-1378.

10. Hama, H. *et al.* (1991) Inhibition of homserine dehydrogenase I by L-serine in *Escherichia coli*. J. Biochem. 109, 604-608.

11. Zhang, X. *et al.* (2010) Deficiency in L-serine deaminase interferes with one-carbon metabolism and cell wall synthesis in *Escherichia coli* K-12. J. Bacteriol. 192 (20), 5515-5525.

12. Kim, J. *et al.* (2010) Three serendipitous pathways in *E. coli* can bypass a block in pyridoxal-5′ -phosphate synthesis. Mol. Syst. Biol. 6.

13. Saxild, H.H. and Nygaard, P. (1991) Regulation of levels of purine biosynthetic enzymes in *Bacillus subtilis*: effects of changing purine nucleotide pools. Microbiology 137 (10), 2387-2394.

14. Stepanek, J.J. *et al.* (2016) Purine biosynthesis is the bottleneck in trimethoprim-treated *Bacillus subtilis*. Proteom. Clin. Appl. 10 (9-10), 1036-1048.

15. Ernst, D.C. *et al.* (2014) Endogenous synthesis of 2-aminoacrylate contributes to cysteine sensitivity in *Salmonella enterica*. J. Bacteriol. 196 (18), 3335-3342.

16. Lambrecht, J.A. *et al.* (2012) Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. J. Biol. Chem. 287 (5), 3454-3461.

17. Lambrecht, J.A. *et al.* (2013) RidA proteins prevent metabolic damage inflicted by PLP-dependent dehydratases in all domains of life. mBio 4 (1), e00033-13-e00033-13.

18. Schmitz, G. and Downs, D.M. (2004) Reduced transaminase B (IlvE) activity caused by the lack of *yjgF* is dependent on the status of threonine deaminase (IlvA) in *Salmonella enterica* Serovar Typhimurium. J. Bacteriol. 186 (3), 803-810.

19. Flynn, J.M. *et al.* (2013) Decreased coenzyme A levels in *ridA* mutant strains of *Salmonella enterica* result from inactivated serine hydroxymethyltransferase: *ridA* mutants are deficient in one carbon metabolism. Mol. Microbiol. 89 (4), 751-759.

20. Chargaff, E. and Sprinson, D.B. (1943) Studies on the mechanism of deamination of serine and threonine in biological systems. J. Biol. Chem. 151 (1), 273-280.

21. Phillips, A.T. and Wood, W.A. (1965) The mechanism of action of 5'-adenylic acid-activated threonine dehydrase. J. Biol. Chem 240 (240), 4703-4709.

22. Flynn, J.M. and Downs, D.M. (2013) In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5'-phosphate cofactor. J. Bacteriol. 195 (16), 3603-3609.

23. Niehaus, T.D. *et al.* (2015) Genomic and experimental evidence for multiple metabolic functions in the RidA/YjgF/YER057c/UK114 (Rid) protein family. BMC Genomics 16 (1).

24. Ernst, D.C. and Downs, D.M. (2016) 2-aminoacrylate stress induces a context-dependent glycine requirement in *ridA* strains of *Salmonella enterica*. J. Bacteriol. 198 (3), 536-543.

25. Christopherson, M.R. *et al.* (2008) YjgF is required for isoleucine biosynthesis when *Salmonella enterica* is grown on pyruvate medium. J. Bacteriol. 190 (8), 3057-3062.

26. Davletova, S. (2005) The zinc-finger protein Zat12 plays a central role in reactive oxygen and abiotic stress signaling in *Arabidopsis*. Plant Physiol. 139 (2), 847-856.

27. Schmieger, H. (1972) Phage P22-mutants with increased or decreased transduction abilities. Mol. Gen. Genet. 119 (1), 75-88.

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

29. Galloway, N.R. *et al.* (2013) Rapid cloning for protein crystallography using type IIS restriction enzymes. Crst. Growth Des. 13 (7), 2833-2839.

30. VanDrisse, C.M. and Escalante-Semerena, J.C. (2016) New high-cloning-efficiency vectors for complementation studies and recombinant protein overproduction in *Escherichia coli* and *Salmonella enterica*. Plasmid 86, 1-6.

31. Vogel, H.J. and Bonner, D.M. (1956) Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. 218 (1), 97-106.

32. Balch, W.E. and Wolfe, R.S. (1976) New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressureized atmosphere. Appl. Environ. Microbiol. 32 (6), 781-791.

33. Stead, M.B. *et al.* (2012) RNAsnap : a rapid, quantitative and inexpensive, method for isolating total RNA from bacteria. Nucleic Acids Res. 40 (20), e156-e156.

34. Magoc, T. *et al.* (2013) EDGE-pro: Estimated degree of gene expression in prokaryotic genomes. Evolutionary Bioinformatics 9, 127-136.

35. Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. Royal Stat. Soc. 57 (1), 289-300.

36. Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. Genome biol 11 (10), R106.

37. Robles, J.A. *et al.* (2012) Efficient experimental design and analysis strategies for the detection of differential expression using RNA-Sequencing. BMC Genomics 13 (1), 484.

38. Fleige, S. and Pfaffl, M.W. (2006) RNA integrity and the effect on the real-time qRT-PCR performance. Mol. Aspects Med. 27 (2-3), 126-139.

39. Heng, S.S.J. *et al.* (2011) Glucan biosynthesis protein G is a suitable reference gene in *Escherichia coli* K-12. ISRN Microbiology 2011, 1-6.

40. Czapski, T.R. and Trun, N. (2014) Expression of *csp* genes in *E. coli* K-12 in defined rich and defined minimal media during normal growth, and after cold-shock. Gene 547 (1), 91-97.

41. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25 (4), 402-408.

42. Zhi, J. *et al.* (1999) Lrp binds to two regions in the *dadAX* promoter region of *Escherichia coli* to repress and activate transcription directly. Mol. Microbiol. 32 (1), 29-40.

43. Saini, S. *et al.* (2008) FliZ Is a posttranslational activator of FlhD4C2-dependent flagellar gene expression. J. Bacteriol. 190 (14), 4979-4988.

44. Christopherson, M.R. *et al.* (2012) Suppressor analyses identify threonine as a modulator of *ridA* mutant phenotypes in *Salmonella enterica*. PLoS ONE 7 (8), e43082.

45. Umbarger, H.E. and Brown, B. (1957) Threonine deamination in *Escherichia coli* ii.: Evidence for two l-threonine deaminases. J. Bacteriol. 73 (1), 105.

46. Borchert, A.J. and Downs, D.M. (2017) The response to 2-aminoacrylate differs in *Escherichia coli* and *Salmonella enterica*, despite shared metabolic components. J. Bacteriol. 199 (14), e00140-17.

47. Pruss, B.M. *et al.* (2003) FlhD/FlhC is a regulator of anaerobic respiration and the Entner-Doudoroff pathway through induction of the methyl-accepting chemotaxis protein Aer. J. Bacteriol. 185 (2), 534-543.

48. Iyoda, S. *et al.* (2001) A flagellar gene *fliZ* regulates the expression of invasion genes and virulence phenotype in *Salmonella enterica* serovar Typhimurium. Microb. Pathog. 30 (2), 81-90.

49. Wada, T. *et al.* (2011) EAL domain protein YdiV acts as an Anti-FlhD4C2 factor responsible for nutritional control of the flagellar regulon in *Salmonella enterica* serovar Typhimurium. J. Bacteriol. 193 (7), 1600-1611.

50. Wada, T. *et al.* (2011) FliZ Acts as a repressor of the *ydiV* gene, which encodes an anti-FlhD4C2 factor of the flagellar regulon in *Salmonella enterica* Serovar Typhimurium. J. Bacteriol. 193 (19), 5191-5198.

51. Stewart, M.K. and Cookson, B.T. (2014) Mutually repressing repressor functions and multi-layered cellular heterogeneity regulate the bistable *Salmonella fliC* census: Multi-layered heterogeneity shapes *fliC* bistability. Mol. Microbiol. 94 (6), 1272-1284.

52. Rice, C.J. *et al.* (2015) Transcriptional and post-transcriptional modulation of SPI1 and SPI2 expression by ppGpp, RpoS and DksA in *Salmonella enterica* sv Typhimurium. PLoS ONE 10 (6), e0127523.

53. Wood, M.W. *et al.* (1998) Identification of a pathogenicity island required for *Salmonella* enteropathogenicity. Mol. Microbiol. 29 (3), 883-891.

54. Müller, A. *et al.* (2014) Activation of RidA chaperone function by N-chlorination. Nat. Commun. 5, 5804.

55. De Lay, N. and Gottesman, S. (2012) A complex network of small non-coding RNAs regulate motility in *Escherichia coli*: sRNA regulation of motility in *Escherichia coli*. Mol. Microbiol. 86 (3), 524-538.

56. Pruss, B.M. *et al.* (2001) FlhD/FlhC-regulated promoters analyzed by gene array and *lacZ* gene fusions. FEMS Microbiol. Lett. 197 (1), 91-97.

57. Niehaus, T.D. *et al.* (2014) *Arabidopsis* and maize RidA proteins preempt reactive enamine/imine damage to branched-chain amino acid biosynthesis in plastids. Plant Cell 26 (7), 3010-3022.

58. Chubiz, J.E.C. *et al.* (2010) FliZ regulates expression of the *Salmonella* pathogenicity island 1 invasion locus by controlling HilD protein activity in *Salmonella enterica* serovar Typhimurium. J. Bacteriol. 192 (23), 6261-6270.

Table 5.1. Swimming motility is decreased in *ridA* mutant strains.

	Strain	Relevant genotype	Addition	Swim zone (mm)	% motility
1	DM9404	wild-type	None	19 ± 1.7	100
2	DM3480	*ridA*	None	11 ± 1.0	58
3	DM9404	wild-type	Gly	18 ± 1.7	100
4	DM3480	*ridA*	Gly	14 ± 1.3	78
5	DM15418	pCV1	L-ara	20 ± 0.8	100
6	DM15419	*ridA* pCV1	L-ara	14 ± 0.3	70
7	DM15420	pDM1439	L-ara	22 ± 1.3	100
8	DM15421	*ridA* pDM1439	L-ara	21 ± 0.3	95

The diameter of motility halos was determined on minimal glucose medium with 0.25% agar. When indicated, L-glycine (0.67 mM), or L-arabinose (0.2%) was present in the media. Media contained ampicillin (7.5 ug/mL) when plasmid-containing strains were used. Values shown as the average diameter and standard deviation (SD) of the growth zone of four biological replicates incubated for 20 h at 30°C. The data for each biological replicate was the average of two technical replicates. Percent motility is the ratio of the *ridA+/ridA-* isogenic strains in each case.

Table 5.2. Accumulation of 2AA causes a motility defect of *ridA* mutant strains.

	Strain	Relevant genotype	Addition	Swim zone (mm)	% motility
1	DM9404	wild-type	None	19 ± 1.7	100
2	DM3480	*ridA*	None	11 ± 1.0	58
3	DM9404	wild-type	Thr	20 ± 0.7	100
4	DM3480	*ridA*	Thr	20 ± 0.4	100
5	DM9404	wild-type	Ile	21 ± 1.3	100
6	DM3480	*ridA*	Ile	23 ± 0.8	110
7	DM6947	*ilvA219*	Gly	19 ± 0.3	100
8	DM6946	*ridA ilvA219*	Gly	11 ± 0.3	58
9	DM6947	*ilvA219*	Gly, Ile	23 ± 0.3	100
10	DM6946	*ridA ilvA219*	Gly, Ile	16 ± 0.6	70
11	DM15035	*ABcc1*	Gly	21 ± 0.9	100
12	DM15036	*ridA ABcc1*	Gly	16 ± 0.5	76
13	DM15035	*ABcc1*	L-ara, Gly	25 ± 1.2	100
14	DM15036	*ridA ABcc1*	L-ara, Gly	14 ± 0.8	56
15	DM15035	*ABcc1*	L-ara, Ile, Gly	23 ± 2.1	100
16	DM15036	*ridA ABcc1*	L-ara, Ile, Gly	17 ± 0.5	74

The diameter of motility halos was determined on minimal glucose (lines 1-10) or glycerol (lines 11-16) motility medium (0.25% agar). When indicated, L-arabinose (0.2%), L-threonine (1 mM), L-isoleucine (1 mM), and/or L-glycine (0.67 mM) was added. Values shown as the average diameter and standard deviation (SD) of the growth zone of four biological replicates incubated for 20 h at 30°C. The data for each biological replicate was the average of two technical replicates. Percent motility is the ratio of motility of the *ridA+/ridA-* isogenic strains in each case.

Table 5.3. Effect of *ridA* on motility is not via known regulators of $FlhD_4C_2$ activity.

	Strain	Relevant genotype	Swim zone (mm)	% motility	*ridA* effect (%)
1	DM9404	wild-type	15 ± 0.6	100	100
2	DM3480	*ridA*	11 ± 0.7	73	73
3	DM15509	*micA*	13 ± 0.5	86	100
4	DM15510	*micA ridA*	9 ± 0.6	60	69
5	DM15513	*arcZ*	15 ± 0.9	100	100
6	DM15514	*arcZ ridA*	9 ± 1.0	60	60
7	DM15507	*omrBA*	11 ± 0.6	73	100
8	DM15508	*omrBA ridA*	9 ± 0.7	60	82
9	DM15511	*oxyS*	12 ± 0.6	80	100
10	DM15512	*oxyS ridA*	9 ± 0.7	60	75
11	DM15319	*ydiV*	16 ± 0.3	107	100
12	DM15320	*ydiV ridA*	10 ± 0.6	67	63
13	DM15340	*rpoS*	15 ± 1.3	100	100
14	DM15341	*rpoS ridA*	10 ± 0.9	67	67
15	DM15505	*fliZ*	9 ± 0.6	60	100
16	DM15506	*fliZ ridA*	5 ± 0.3	33	56

Motility assays were performed as described in Materials and Methods, with the strains listed. Values shown are the average diameter and standard deviation (SD) of four biological replicates incubated for 20 h at 30°C on minimal glucose (11 mM) L-glycine (0.67mM) motility plates (0.25% agar). % motility is compared to the wild-type strain (line 1), *ridA* effect is the ratio of motility for the *ridA+/ridA-* isogenic strains.

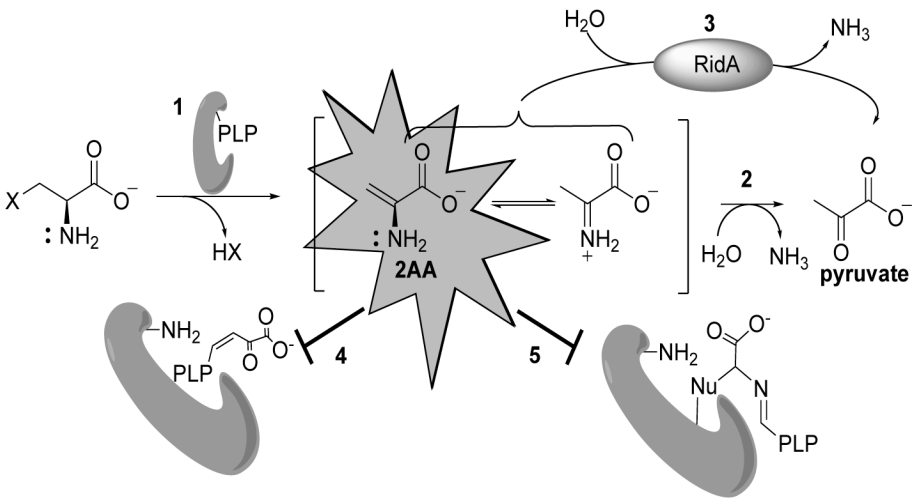


Fig. 5.1 RidA paradigm of 2AA stress in *S.* enterica. Multiple pyridoxal 5'-phosphate (PLP)-dependent enzymes can catalyze β-elimination reactions of 3-carbon alpha amino acids containing favorable leaving groups on the beta-carbon (1). This generates the reactive enamine intermediate 2-aminoacrylate (2AA), which is into solution. 2AA has three possible fates following its release from the active site: First, 2AA can tautomerize to its imine form, producing 2-iminiopropanoate, which is non-enzymatically hydrolyzed, forming the stable keto-acid pyruvate (2). Second, RidA catalyzes the deamination of 2AA to pyruvate (3). Finally, 2AA can attack the internal aldimine of various PLP enzymes, irreversibly damaging the target enzyme. 2AA attack results in the formation of an inactive pyruvate-PLP adduct (4) or covalent modification of the enzyme by the 2AA-PLP aldimine (5).

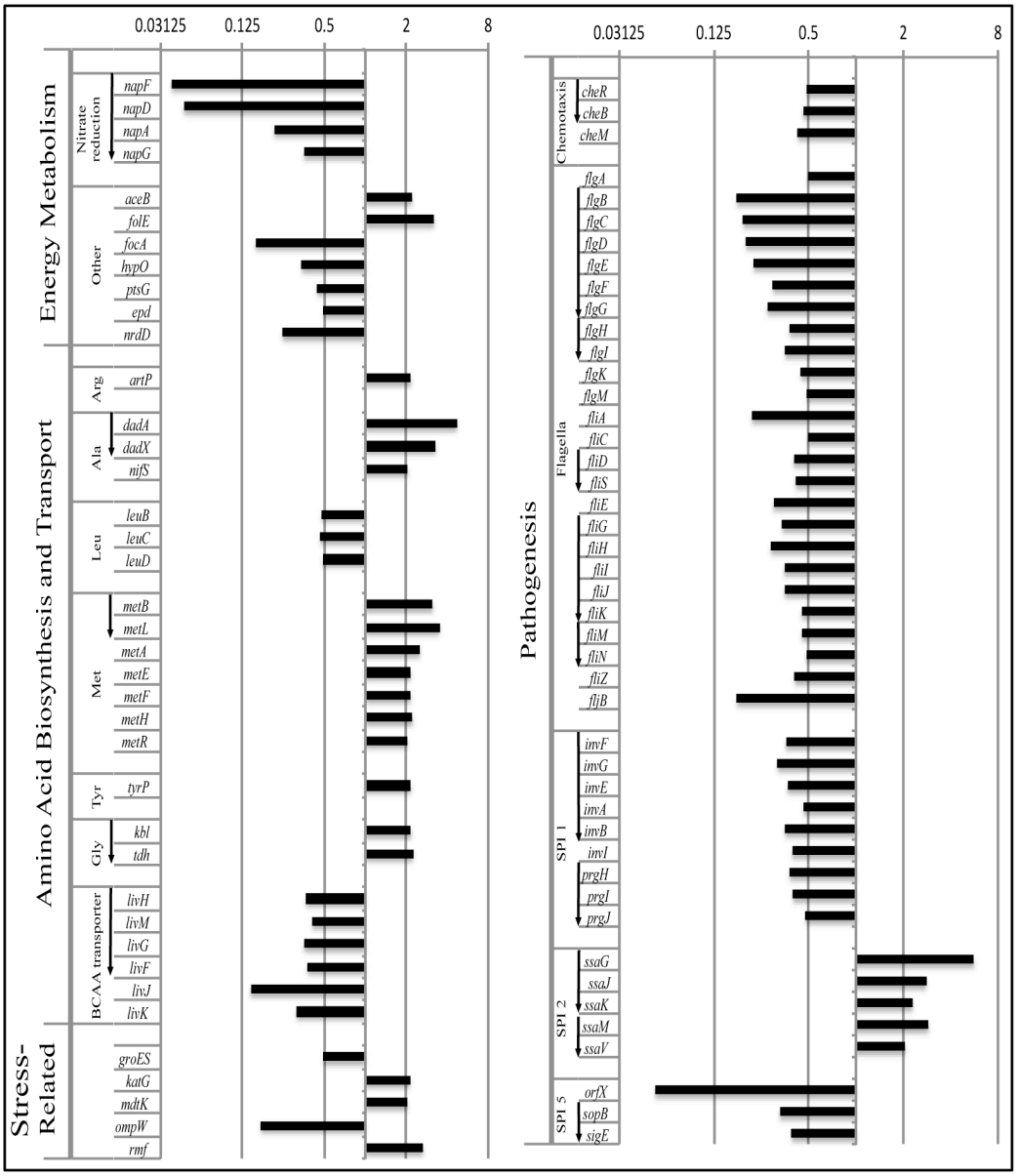


Fig. 5.2 Gene expression differs between a *ridA* mutant and wild type strain. Summary of relative expression levels of genes in *S.* Typhimurium LT2 found to differ significantly (FDR < 0.05, Fold-change > 2) between *ridA* and wild-type strains. A value of 1 indicates no detectable difference in expression, values > 1 indicate higher expression in the *ridA* mutant, and values < 1 indicate higher expression in wild-type. Arrows to the left of the gene name indicate whether that gene is contained within an operon, with the order of the operon arranged following the arrow base to the arrowhead.

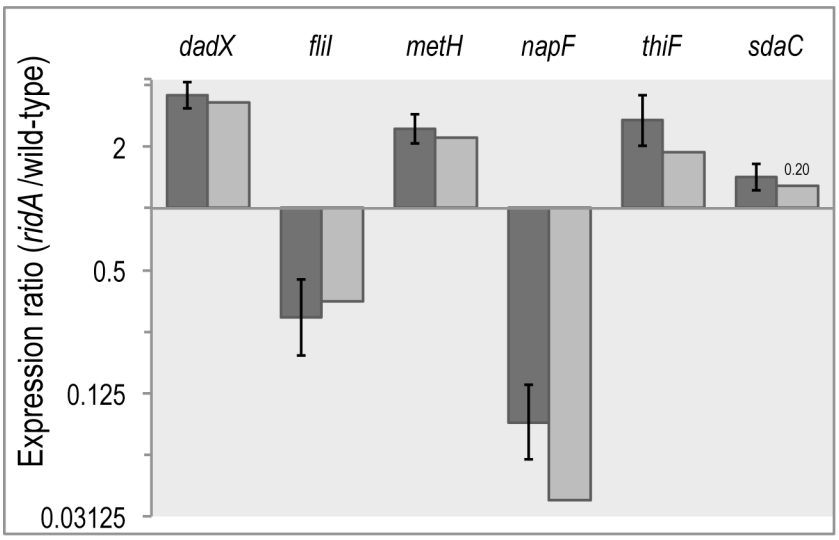


Fig. 5.3 qTR-PCR supports RNAseq differential expression results for a subset of genes. Differential expression of a subset of genes for *S. enterica* LT2 *ridA* strain (DM1480) relative to its isogenic wild-type (DM9404) as determined by qRT-PCR (dark grey bars). The RNAseq differential expression values (light grey bars) were included as a reference for comparison. Error bars on qRT-PCR data represent the 95% confidence interval. All genes presented, excluding *sdaC*, displayed differential expression (*ridA*/wild-type) with FDR < 0.05, following RNAseq analysis; the corresponding FDR value for *sdaC* is provided above the bar representative of the RNAseq data.

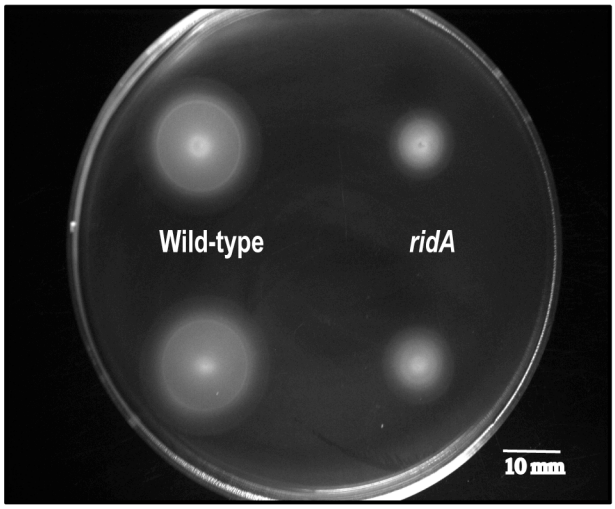


Fig. 5.4 A *ridA* mutation leads to a defect in swimming motility. Representative images of swimming motility halos for *ridA* and wild-type strains grown on minimal glucose (11 mM) motility plates (0.25% agar) Following inoculation using 1 uL overnight cultures grown in LB medium, plates were allowed to incubate 20 h at 30°C. Diameter using the outermost edge for each halo was measured, to determine zone of swimming motility. Two biological replicates are shown for each strain.

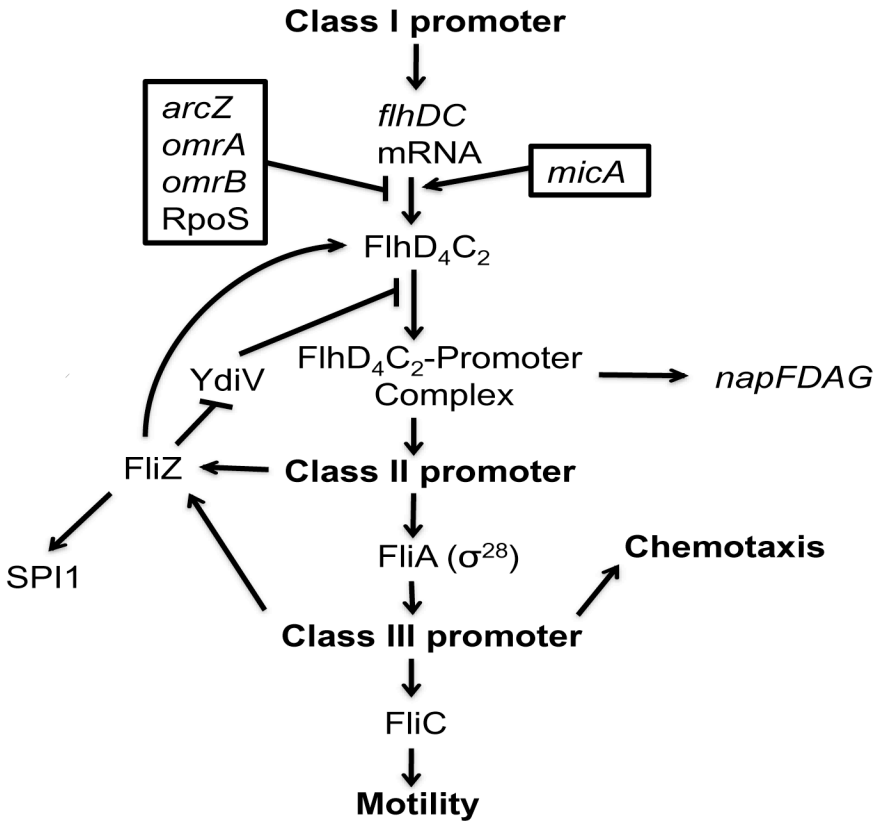


Fig. 5.5 General regulatory cascade involving $FlhD_4C_2$. Schematic of the model for regulation involving $FlhD_4C_2$ in *E. coli* is depicted. Arrows and blocked arrows indicate positive and negative regulation, respectively. The four small Hfq-dependent non-coding regulatory RNAs *arcZ, omrA, omrB, and oxyS*, negatively regulate motility by binding with the 5' untranslated region of *flhDC* and preventing translation[55]. YdiV acts as an anti-$FlhD_4C_2$ factor to disrupt promoter binding[49]. FliZ can counterbalance this affect by repressing *ydiV* expression as well as increasing $FlhD_4C_2$ protein abundance in an YdiV-independent manner[51]. FliZ also promotes increased Salmonella pathogenicity island 1 (SPI1) transcription[58].

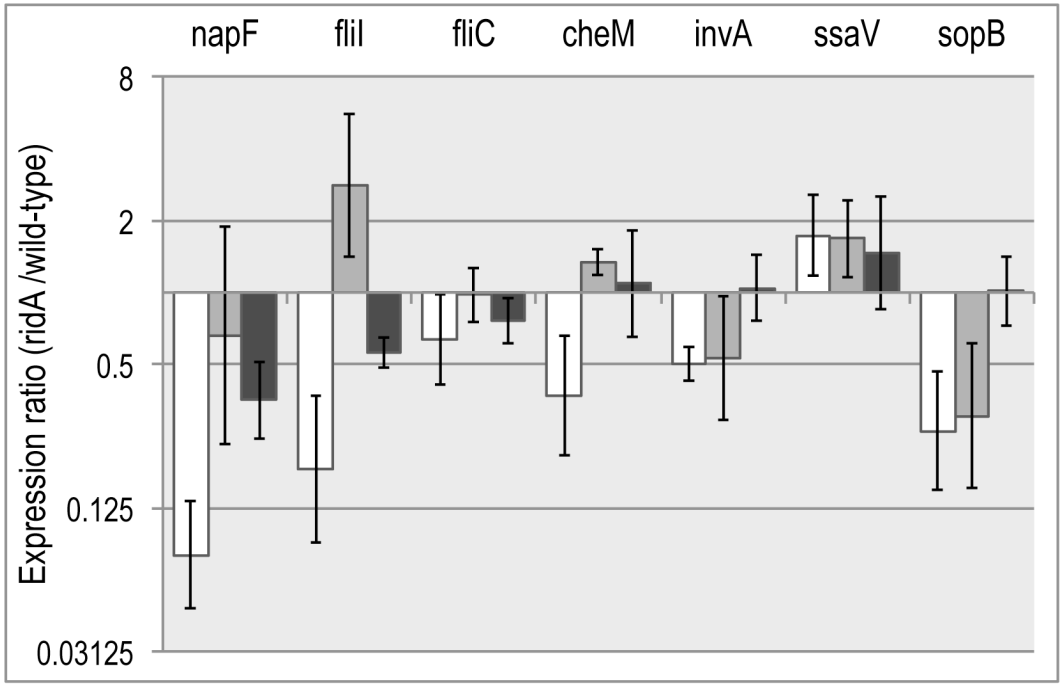


Fig. 5.6 Differential expression is impacted by status of $FlhD_4C_2$ and growth medium. Expression of a subset of genes was monitored in a *ridA* strain (DM3480) relative to the isogenic wild-type strain (DM9404) by qRT-PCR. The strains were grown in minimal glucose medium (white bars) and minimal medium with 1 mM L-glycine (dark grey bars). Gene expression was similarly measured in an *flhDC ridA* strain (DM15817) relative to the isogenic *flhDC* strain (DM15818) when they were grown in minimal glucose medium (light grey bars). Error bars represent the 95% confidence interval.

CHAPTER 6

INTEGRATED METABOLOMICS AND TRANSCRIPTOMICS UNCOVER THE GLOBAL METABOLIC RESPONSE OF ELIMINATING RIDA IN *SALMONELLA ENTERICA* [1]

[1]Borchert AJ, Walejko JM, Le Guenec A, Ernst DC, Edison AS Downs DM. 2019. To be submitted to *Metabolites* in November, 2019.

6.1 ABSTRACT

In *Salmonella enterica*, 2-aminoacrylate (2AA) is a reactive enamine intermediate generated during a number of biochemical reactions. When the 2-iminobutanoate/2-iminopropanoate deaminase (RidA; EC: 3.5.99.10) is eliminated, 2AA accumulates and inhibits the activity of multiple PLP-dependent enzymes. In this study, untargeted ^{1}H NMR metabolomics and transcriptomics data obtained from a previous study were used to uncover the global metabolic response of *S. enterica* to the accumulation of 2AA. The data showed that elimination of RidA perturbed folate and branched chain amino acid metabolism. Many of the resulting perturbations were consistent with the known effect of 2AA stress, while other results suggested additional potential enzyme targets of 2AA-dependent damage. The majority of transcriptional and metabolic changes appeared to be the consequence of downstream, indirect effects on the metabolic network, since they were not directly attributable to a PLP-dependent enzyme. In total, the results highlighted the complexity of changes stemming from multiple perturbations of the metabolic network and suggested hypotheses that will be valuable in future studies of the RidA paradigm of endogenous stress.

6.2 BACKGROUND

The metabolic network of an organism consists of a complex system of biochemical reactions that together result in the behavioral characteristics of the organism. In general, the reactions that make up the metabolic network, and their integration, are governed by chemical and physical constrains placed on the cell by its environment. Microbes in particular, face rapidly changing environments and their metabolic networks must be structured to absorb and/or respond to internal and external insults [1-3]. Technological advances continue to expand and provide deeper and more detailed global snapshots of features in a cell or population. As these large

datasets accumulate, the focus turns to analyzing these data as a means to glean insights about fundamental processes in biological systems. To that end, untargeted metabolomics and comparative transcriptomics data are increasingly being used as the basis for building metabolic models that describe the physiological state of a cell [4-6]. Examples of this approach can be found in disciplines ranging from mammalian genetic disorders and cancer biology to antibiotic modes of action and pathogen-host interaction [7-10]. While these analyses can contribute to defining the response to various cellular perturbations [11, 12], they are most significant when confirmed by rigorous genetic, biochemical, and molecular biological experimentation. In theory, the combination of global discovery approaches with focused reductionist analyses provides an optimal approach to generate insights into the fundamentals of metabolic network structure, and cellular strategies [13, 14].

An important use of global strategies is to expand the scope of a model that has been built on biochemical genetic approaches, which often define detail, but can lack a global snapshot. The RidA paradigm of enamine stress in *Salmonella enterica* defines a metabolic model with clearly defined local effects that result in predictable global consequences. Members of the broadly conserved RidA protein family are responsible for the hydrolysis of enamine/imine species, which are generated as intermediates in a variety of biochemical reactions [15-20] (Figure 1). In the absence of RidA, the enamine 2-aminoacrylate (2AA), a well-characterized inhibitor of multiple pyridoxal 5'-phosphate (PLP)-dependent enzymes, accumulates (Figure 1). Accumulated 2AA has been shown to inhibit serine hydroxymethyltransferase (GlyA, EC: 2.1.2.1), aspartate aminotransferase (AspC; EC: 2.6.1.1), alanine racemase (Alr and DadX; 5.1.1.1), and branched-chain amino acid aminotransferase (IlvE; EC:2.6.1.42) [21-24] reducing the pool of activity for each by 40-60%.

Overall, *S. enterica* has over 40 different PLP-dependent enzymes with functions in amino acid metabolism, cofactor/coenzyme biosynthesis, and cell structure component synthesis (Appendix Table A.4) [25, 26]. While a handful of specific 2AA targets have been defined, less is known about the extent of 2AA-dependent damage or how such damage, (i.e., enzyme inhibition) impacts the metabolic network. 2AA stress is expected to elicit a broad shift in the metabolome, observed as the cumulative response of damaging a subset of PLP-dependent enzymes. A previous comparative transcriptomics study highlighted transcriptional changes in a strain lacking *ridA,* and led to the identification of an uncharacterized requirement for RidA in motility [27]. A majority of the most significant transcriptional changes (fold-change >2) had no direct connection to a PLP-dependent enzyme, suggesting they were a downstream consequence 2AA damage.

This study was initiated to further our understanding of the global metabolic response of *S. enterica* to 2AA stress. Specifically, untargeted NMR metabolomic data was sought to complement the existing transcriptomics data. Our understanding of the RidA system at a physiological, genetic, and biochemical level provided a critical basis to guide interpretation of the global data sets.

6.3 MATERIALS AND METHODS

Bacterial strains, chemicals, and media. Strains used in this work are derivatives of *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2. The *ridA* null mutant (DM3480) used in this study contains a MudJ1734 transposon insertion [28] disrupting the *ridA* locus (*ridA3*::MudJ1734) [29]. Strain DM9404 is isogenic to DM3480 and has a wildtype *ridA* locus. Minimal medium was no-carbon E medium (NCE) containing 1 mM MgSO4 [30], trace metals [31], and 11 mM D-glucose. Difco nutrient broth (NB) (8 g/L) supplemented with 5 g/L

NaCl was used as rich medium. All chemicals were purchased from the Sigma-Aldrich Chemical Company (St. Louis, MO).

Metabolomics cell preparation. Ten biologically independent cultures each of wild-type and *ridA* mutant strains (DM9404 and 3480, respectively) were grown overnight in NB medium at 37 °C and used to inoculate (1% inoculum) 250 mL minimal glucose medium in 500 mL non-baffled flasks. Flasks were randomly arranged in an Innova®44 incubator and cultures were allowed to grow 16 h shaking at 180 RPM and 37 °C. Cultures were cooled on ice 5 min and then harvested by centrifugation at 7,000 x G for 10 min at 4 °C. The supernatant was decanted, pellets were resuspended in 10 mL ddH_2O and transferred to sterile 15 mL conical tubes in which they were pelleted at 7,000 x G 10 min at 4 °C. Final supernatant was decanted and pellets were frozen in liquid nitrogen and stored at -80 °C prior to cell extractions.

Acquisition and processing of NMR spectral data. *1D ^1H:* The 1D proton spectra were collected using the pulse sequence 'noesypr1d' from the Bruker library. A mixing time of 5 ms was used, and during the acquisition, 65,536 complex datapoints were collected for the FID using 128 scans, with 4 additional dummy scans for equilibration and 4 seconds between scans. The spectral width was 20 ppm. For processing, 32768 points were used for the spectrum, using an exponential window function of 0.3 Hz before the Fourier transform (FT). After FT and phase correction, a polynomial baseline correction of order 3 was used and the frequency calibrated to the DSS peak (0.00 ppm).

J-RES: The 1D ^{1}H, J-RES and TOCSY spectra were acquired on a Bruker NEO 800 MHz spectrometer equipped with a SampleJet sample changer and $^1H/^{13}C/^{15}N$ cryoprobe, at a temperature of 300 K. The 2D J-RES spectra were collected using the pulse sequence 'jresgpprqf' from the Bruker library. 32,768 points were used for the FID in the direct dimension and 64 points

for the indirect dimension, with a spectral width of 20 ppm for the direct dimension and 50 Hz for the indirect dimension. Eight scans were acquired for each t_1 increment, along with 4 dummy scans for equilibrium and 2 seconds between scans. For processing, 32,768 points were used for the direct dimension of the spectrum and 512 points for the indirect dimension. An unshifted sine window function was also used in each dimension, along with backward linear prediction of 64 points in the indirect dimension to increase the sensitivity as well as the resolution [32]. A polynomial baseline correction of order 3 in the direct dimension and order 5 in the indirect dimension was applied after the FT in both dimensions.

Compound identification/database matching. Two-dimensional ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC) and ^{1}H-^{13}C HSQC–TOCSY (HSQC–total correlation spectroscopy) experiments were used to aid in annotation of 2D J-RES metabolites and were not used for statistical analysis. A total of 37 metabolites were identified in bacterial extractions using COLMARm [33], Comprehensive Metabolite Identification Strategy Using Multiple Two-Dimensional NMR Spectra of a Complex Mixture Implemented in the COLMARm Web Server and assigned a confidence level ranging from 1 to 5, as previously described [34]. In short: (1) putatively characterized compound, (2) matched reported 1D spectra, (3) matched reported HSQC spectra, (4) matched reported HSQC and HSQC-TOCSY spectra, and (5) validated by spiking putative compound into sample. A list of these assignments can be found in the Appendix Table A.5. In addition to the database matching, using 2D NMR data we identified nicotinate mononucleotide, which was not in a spectral database. We verified this identification with a spiking experiment (Appendix Figure A.1).

Quantification of metabolomics data. Two-dimensional (2D) J-RES spectra were used for quantification and annotation of metabolites. The spectral features used for integration and

confidence values for each metabolite included in the statistical analysis are listed in Appendix Table A.5. Univariate statistics were performed on metabolites identified using 2D J-RES data after PQN-normalization. A student's t-test with an FDR correction was used to determine metabolites significantly altered in bacterial extracts from wild-type (n=9) and *ridA* mutants (n=10; FDR-corrected p-value < 0.05).

Pathway analysis. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways that were induced or repressed in *S. enterica ridA* mutants were determined using the Database for Annotation, Visualization and Integrated Discovery (DAVID) pathway enrichment analysis tool [35, 36]. Input was a list of the locus tags (STM numbers) from the significantly (FDR<0.05) differentially regulated genes separated into "up-regulated" and "repressed" subsets. Of the 413 locus tags used as input, 228 (55.2 %) could be placed into a KEGG pathway. The output for up-regulated and down-regulated genes and which KEGG pathway they mapped to is provided in the Appendix Table A.6 and Table A.7, respectively. These tables contain values representing total gene count of pathway, hit counts, overrepresentation analysis (ORA) raw P values, gene identities for hits, and ORA Benjamini-Hochberg FDR values [37]. Significantly overrepresented pathways were defined as pathways with an FDR < 0.1.

6.4 RESULTS AND DISCUSSION

Metabolite levels are altered in an *S. enterica ridA* mutant. Comprehensive ^{1}H NMR analysis was performed on the cellular contents of an isogenic pair of *S. enterica* LT2 strains that differed only at the *ridA* locus (DM9404 and DM3480) after growth to stationary phase in minimal glucose medium. These data were viewed through a biological lens, specifically considering the role PLP-enzyme-dependent enzymes have in the RidA paradigm [38, 39], to extract insights into

the effect of 2AA stress on the metabolic network that goes beyond information gleaned from past genetic, biochemical and transcriptomic analyses.

In total, 37 metabolites could be structurally identified with high confidence from the ^{1}H-NMR spectral data (Table 1). Table 1A contains data from the 23 metabolites with significantly (FDR< 0.05) altered concentrations between wild-type and *ridA* strains (FDR < 0.05). Table 1B shows the 14 metabolites detected in all samples that did not meet the statistical threshold for calling them significantly different in accumulation between strains. The data in Table 1A emphasize that the metabolic environment in a cell is impacted by the status of *ridA*. The relevant metabolites represented diverse areas of the metabolic network and did not suggest a simple model for specific site(s) of 2AA-dependent perturbation. A majority of the metabolites were related to, or the product of, amino acid biosynthetic pathways, and a significant number were components nucleotide metabolism. The former was expected based on the prevalence of PLP-dependent enzymes (targets of 2AA damage) in amino acid metabolism [22-24, 40, 41], while a direct role of PLP-dependent enzymes in nucleotide metabolism was not obvious. The 25 metabolites with differential accumulation represented a small sample, and thus limited the conclusions that could be made about the altered metabolic state of *S. enterica ridA* mutants.

Transcriptome analyses of *ridA* mutant complements metabolomics data. A transcriptomic data set obtained with the same wild-type and *ridA* mutant strains [27] was considered in combination with the metabolomics data to build a more complete picture of the metabolic changes that reflect the cellular response to 2AA accumulation. All gene expression data discussed was extracted from Borchert and Downs, 2019 [27] and is accessible at the NCBI Gene Expression Omnibus (GSE103146; http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE103146). As reported, transcription of 186 genes in a *S. enterica ridA* mutant were significantly (FDR <

0.05) elevated and transcription of 227 genes was significantly decreased, when compared to the wild-type parental strain [27]. Of the 413 genes that were differentially expressed, 113 had a greater than 2-fold difference in expression between *ridA* and WT. These data emphasized that the cellular response to 2AA stress generated transcriptional changes of modest magnitude. Over-representation analysis was performed on all differentially expressed genes, regardless of fold-change intensity, to determine pathways that showed significant enrichment for the differentially expressed genes in a *ridA* mutant (Table 2) [35, 36]. These analyses confirmed that gene networks involved in flagellar biosynthesis, chemotaxis, and epithelial cell invasion were significantly down-regulated in a *ridA* mutant, observations that previously led to the characterization of a motility defect in *ridA* mutants [27]. Genes encoding ribosomal proteins were also significantly down-regulated (Table 2). Since ribosome synthesis is tightly regulated in accordance with growth rate and amino acid availability [42-44] this change was assumed to reflect the general cell status of *ridA* mutants, i.e. perturbed amino acid metabolism and reduced growth rate [41].

Over-representation analysis showed the KEGG pathways for thiamine metabolism (stm00730), biotin metabolism (stm00780), amino acid biosynthesis (stm01230), one-carbon metabolism (stm00670), ABC transporters (stm02010), and glycine, serine and threonine metabolism (stm00260) were all significantly over-represented as up-regulated in a *ridA* mutant (Table 2). Since many of these KEGG pathways contain PLP-dependent enzymes (Table S1), they were the focus of hypotheses to define the metabolic perturbations resulting from 2AA accumulation. In cells that lack of RidA, accumulated 2AA targets PLP-dependent enzymes, and precedent with IlvE, GlyA, and Alr [21-23] suggests the activity of their population is decreased by 20-50%. The decreased activity is expected to collectively dampen flux in multiple metabolic nodes and cause metabolite imbalance(s). A subset of these imbalances would be detectable by an

untargeted metabolomics approach. Such metabolite imbalances have the potential to trigger transcriptional regulatory responses, such as those identified in the transcriptome data.

Folate metabolism is perturbed in a *ridA* mutant. Collectively, the transcriptional and metabolic shifts detected in a *ridA* mutant are consistent with a perturbation in folate metabolism. Tetrahydrofolate (THF) and its derivatives are essential cofactors used to facilitate one-carbon unit transfers in many important pathways [45]. Transcription of multiple genes involved in 5,10-mTHF biosynthesis (*glyA*, *gcvTHP*, *tdh*, *kbl*, *pabB,* and *folE*) were increased in a *ridA* mutant compared to wildtype (Table 2, Fig 2A). This expression profile indicates a response to cells that have been compromised in THF pools. Extensive biochemical genetic analyses supports this finding, showing that the 2AA that accumulates in an *S. enterica ridA* mutant inhibits GlyA and reduces both glycine and 5,10-mTHF levels [22, 41]. Interestingly, expression of *glyA* and *gcvTHP* are induced in a *ridA* mutant, a response which may serve a role in counteracting the consequence of 2AA-dependent GlyA damage. The genes encoding threonine dehydrogenase (Tdh; EC: 1.1.1.103) and 2-amino-3-ketobutyrate CoA ligase (Kbl; EC: 2.3.1.29), which are involved in the conversion of threonine to glycine, are also induced in a *ridA* mutant. Induction of these genes could lead to the rerouting of threonine for glycine production, ultimately generating additional 5,10-mTHF via GCV (Figure 2A, Figure 3). Transcription of *pabB* (1.45-fold) and *folE* (3.21-fold), is also elevated in a *ridA* mutant. Expression of these genes, which are involved in 7,8-DHF production, could increase the biosynthesis of THF for use by GlyA and the GVC complex during 5,10-mTHF production. The putative 5,10-mTHF limitation in a *ridA* mutant could explain the increased expression of multiple pathways that are dependent upon folate metabolism (methionine, purine, histidine, thiamine, biotin, and lipoate biosynthesis) (Figure 2).

Further, the transcriptional repressor, MetJ, binds S-adenosylmethionine (SAM) and represses expression of a number of genes whose expression is increased in a *ridA* mutant (*metA*, *metBL*, *metC*, *metE*, *metF, metMIQ*, *metR,* and *folE* [27, 46]). Increased expression of these *metJ/SAM* regulated genes suggests that SAM is limiting in the *ridA* background. In fact, SAM biosynthesis is reliant on both 5-mTHF and purines (ATP) making this conclusion consistent with the data and interpretations above. An expected consequence of SAM limitation is the decreased production of products of pathways that rely upon SAM as a cofactor. Consistently, expression of genes involved in the SAM-dependent pathways of biotin (*bioA*, *bioBFCD*, and *bisC*), lipoate (*lipA*), and thiamine (*thiCEFSGH* and *thiMD*) biosynthesis were induced, consistent with a limitation for these critical metabolites in a *ridA* mutant (Figure 2B).

Finally, metabolomic analysis by ^{1}H NMR showed that a *ridA* mutant had lower endogenous levels of both formate and acetate (Table 1A). During mixed acid fermentation, pyruvate-formate lyase (PflB: EC: 2.3.1.54), uses CoA and pyruvate as substrates during the production of formate and acetyl-CoA. Acetyl-CoA is further processed to acetate. Importantly, previous characterization of a *S. enterica ridA* mutant has revealed that 2AA-dependent damage of GlyA, and the resulting defect in 1-carbon units leads to a bottleneck in CoA biosynthesis [22]. Therefore, the reduced formate and acetate content in a *ridA* mutant could be explained by a limitation of CoA in a *ridA* mutant.

Branched-chain amino acid (BCAA) metabolism is altered in a *ridA* mutant. Some of the most significant changes, by magnitude, in a *ridA* mutant involved metabolites related to branched-chain amino acid biosynthesis (Figure 3). Metabolomic analysis with ^{1}H NMR showed that isoleucine and leucine content were decreased in a *ridA* mutant, while valine, 2-isopropylmalic acid, and threonine were increased (Table 1A). IlvE is a target of 2AA damage and the resulting

decrease in activity could contribute to the reduced levels of leucine and isoleucine [21]. Down-regulation of the *leu* operon and *ilvC* could also play a role in reducing leucine and isoleucine content in the *ridA* mutant background (Figure 3). Pyruvate accumulates during exponential growth of *ridA* mutants as a result of CoA limitation following 2AA-dependent damage of GlyA and 5,10-mTHF limitation [22]. Internal pyruvate increases flux toward L-valine biosynthesis [47, 48], consistent with the increased valine and 2-isopropylmalic acid levels (Table 1A).

Many metabolic perturbations consistent with transaminase damage by 2AA. The ^{1}H NMR metabolomics indicated there was a decrease in lysine, serine, and phenylalanine, and an increase in glutamate, in a *ridA* mutant. Synthesis of lysine, serine, and phenylalanine all rely upon the activity of PLP-dependent transaminases, which use glutamate as an amino group donor [49-51]. Thus, these changes could reflect that damage to one or more of these transaminases by 2AA. For example, IlvE, AspC, and aromatic amino acid aminotransferase (TyrB, EC: 2.6.1.57) facilitate the synthesis of phenylalanine from phenylpyruvate [50]. Damage of IlvE [21] and AspC [24] by 2AA may explain the phenylalanine decrease in a *ridA* background. Similarly, phosphoserine transaminase (SerC, EC: 2.6.1.52) facilitates 3-phosphooxypyruvate to O-phospho-L-serine conversion during serine biosynthesis, while SerC and N-succinyldiaminopimelate aminotransferase (ArgD, EC: 2.6.1.17) catalyze N-succinyldiaminopimelate production from N-succinyl-L-amino-6-ketopimelate during lysine biosynthesis (Fig. 3) [49, 51]. The susceptibility of SerC to 2AA damage has not been shown, but reduced SerC activity would be expected to decrease both serine and lysine levels.

6.5 CONCLUSIONS

The current study used comparative transcriptomics and untargeted ^{1}H NMR metabolomics analysis to uncover global metabolic consequences of removing RidA, and thus accumulating 2

aminoacrylate in *S. enterica*. The transcriptional changes seen for a *ridA* mutant permeated multiple KEGG pathways, including amino acid, coenzyme, and folate metabolism. Complimenting this finding, ^{1}H NMR revealed multiple metabolic changes that were involved in amino acid, mixed acid fermentation, and nucleic acid metabolism. Some of the patterns from the transcriptomics and metabolomics data were straightforward to explain by extrapolating known consequences from 2AA stress. For instance, induced expression of folate-related genes can be tied to damage of GlyA, and decreased isoleucine and leucine content with damage to IlvE. Indeed, many of the shifts in threonine, serine, lysine, and phenylalanine were consistent with 2AA-dependent damage of PLP-dependent enzymes involved in the metabolism of these metabolites.

Significantly, numerous other metabolic and transcriptional changes in a *ridA* mutant could not easily be attributed to inhibition of one or more PLP-dependent enzyme(s), the known targets of 2AA. These changes are expected to be the consequence of downstream, indirect effects on the metabolic network. Importantly, while they are difficult to tie to a specific 2AA target, these effects often prove important for the physiological state of the organism, causing visible phenotypes. The study herein took advantage of a mechanistically understood, but complex, paradigm of metabolic stress. Accumulation of 2AA causes a cellular stress that is a multipronged, yet subtle, since the activity of multiple enzymes is reduced by 40-60 %. The data herein showed that defining changes to the metabolic network created by multiple small perturbations to the system is not straightforward with the tools currently available. The RidA system can provide a valuable template for continued refinement of global approaches to understand metabolic network structure, since it is understood at a biochemical level and can be manipulated genetically to test hypotheses that arise.

6.6 REFERENCES

1. Keller, M.A. *et al.* (2015) The widespread role of non-enzymatic reactions in cellular metabolism. Curr. Opin. Biotechnol. 34, 153-161.

2. de Lorenzo, V. *et al.* (2015) Chemical reactivity drives spatiotemporal organisation of bacterial metabolism. FEMS Microbiol. Rev. 39 (1), 96-119.

3. Albert, R. *et al.* (2000) Error and attack tolerance of complex networks. Nature 406, 378-382.

4. Gertsman, I. and Barshop, B.A. (2018) Promises and pitfalls of untargeted metabolomics. J. Inherited Metab. Dis. 41 (3), 355-366.

5. Furlan, S.N. *et al.* (2015) Transcriptome analysis of GVHD reveals aurora kinase A as a targetable pathway for disease prevention. Sci. Transl. Med 7 (315), 315ra191-315ra191.

6. Koenig, D. *et al.* (2013) Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. Proc. Natl. Acad. Sci. USA 110 (28), E2655-E2662.

7. Lee-McMullen, B. *et al.* (2019) Age-dependent changes in metabolite profile and lipid saturation in dystrophic mice. NMR Biomed., e4075.

8. Clendinen, C.S. *et al.* (2019) Preoperative metabolic signatures of prostate cancer recurrence following radical prostatectomy. J Proteome Res. 18 (3), 1316-1327.

9. Vincent, I.M. *et al.* (2016) Untargeted metabolomics to ascertain antibiotic modes of action. Antimicrob. Agents Ch. 60 (4), 2281-2291.

10. Mareya, C. *et al.* (2019) Untargeted Metabolomics Reveal Defensome-Related Metabolic Reprogramming in Sorghum bicolor against Infection by Burkholderia andropogonis. Metabolites 9 (1), 8.

11. Hattori, A. *et al.* (2017) Cancer progression by reprogrammed BCAA metabolism in myeloid leukaemia. Nature 545 (7655), 500-504.

12. Baptista, R. *et al.* (2018) Untargeted metabolomics reveals a new mode of action of pretomanid (PA-824). Sci. Rep. 8 (1), 5084.

13. Downs, D. *et al.* (2018) The three-legged stool of understanding metabolism: integrating metabolomics with biochemical genetics and computational modeling. AIMS Microbiology 4 (2), 289-303.

14. Bazurto, J.V. *et al.* (2018) Untargeted metabolomics confirms and extends the understanding of the impact of aminoimidazole carboxamide ribotide (AICAR) in the metabolic network of *Salmonella enterica*. Microbial Cell 5 (2), 74-87.

15. Lambrecht, J.A. *et al.* (2012) Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. J. Biol. Chem. 287 (5), 3454-3461.

16. Niehaus, T.D. *et al.* (2014) *Arabidopsis* and maize RidA proteins preempt reactive enamine/imine damage to branched-chain amino acid biosynthesis in plastids. Plant Cell 26 (7), 3010-3022.

17. Ernst, D.C. *et al.* (2014) Endogenous synthesis of 2-aminoacrylate contributes to cysteine sensitivity in *Salmonella enterica*. J. Bacteriol. 196 (18), 3335-3342.

18. Hodge-Hanson, K.M. and Downs, D.M. (2017) Members of the Rid protein family have broad imine deaminase activity and can accelerate the *Pseudomonas aeruginosa* D-arginine dehydrogenase (DauA) reaction *in vitro*. PloS one 12 (9), e0185544.

19. Ernst, D.C. and Downs, D.M. (2018) Mmf1p couples amino acid metabolism to mitochondrial DNA maintenance in *Saccharomyces cerevisiae*. mBio 9 (1).

20. Degani, G. *et al.* (2018) Imine deaminase activity and conformational stability of UK114, the mammalian member of the Rid protein family active in amino acid metabolism. Int. J. Mol. Sci. 19 (4).

21. Lambrecht, J.A. *et al.* (2013) RidA proteins prevent metabolic damage inflicted by PLP-dependent dehydratases in all domains of life. mBio 4 (1), e00033-13-e00033-13.

22. Flynn, J.M. *et al.* (2013) Decreased coenzyme A levels in *ridA* mutant strains of *Salmonella enterica* result from inactivated serine hydroxymethyltransferase: *ridA* mutants are deficient in one carbon metabolism. Mol. Microbiol. 89 (4), 751-759.

23. Flynn, J.M. and Downs, D.M. (2013) In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5'-phosphate cofactor. J. Bacteriol. 195 (16), 3603-3609.

24. Borchert, A.J. and Downs, D.M. (2017) The response to 2-aminoacrylate differs in *Escherichia coli* and *Salmonella enterica*, despite shared metabolic components. J. Bacteriol. 199 (14), e00140-17.

25. Percudani, R. and Peracchi, A. (2003) A genomic overview of pyridoxal-phosphate-dependent enzymes. EMBO reports 4 (9), 850-854.

26. Karp, P.D. *et al.* (2019) The BioCyc collection of microbial genomes and metabolic pathways. Briefings in Bioinformatics 20 (4), 1085-1093.

27. Borchert, A.J. and Downs, D.M. (2017) Endogenously generated 2-aminoacrylate inhibits motility in *Salmonella enterica*. Sci. Rep. 7 (1).

28. Castilho, B.A. *et al.* (1984) Plasmid insertion mutagenesis and *lac* gene fusion with mini-mu bacteriophage transposons. J. Bacteriol. 158 (2), 488-495.

29. Schmitz, G. and Downs, D.M. (2004) Reduced transaminase B (IlvE) activity caused by the lack of *yjgF* is dependent on the status of threonine deaminase (IlvA) in *Salmonella enterica* Serovar Typhimurium. J. Bacteriol. 186 (3), 803-810.

30. Vogel, H.J. and Bonner, D.M. (1956) Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. 218 (1), 97-106.

31. Balch, W.E. and Wolfe, R.S. (1976) New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressureized atmosphere. Appl. Environ. Microbiol. 32 (6), 781-791.

32. Sakhaii, P. and Bermel, W. (2014) Improving the sensitivity of conventional spin echo spectra by preservation of initial signal-to-noise ratio. J. Magn. Reson. 242, 220-3.

33. Bingol, K. *et al.* (2016) Comprehensive metabolite identification strategy using multiple two-dimensional NMR spectra of a complex mixture implemented in the COLMARm web server. Anal. Chem. 88 (24), 12411-12418.

34. Walejko, J.M. *et al.* (2018) Multiomics approach reveals metabolic changes in the heart at birth. Am. J. Physiol. Endocrinol. Metab. 315 (6), E1212-E1223.

35. Huang, D.W. *et al.* (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4 (1), 44-57.

36. Huang, D.W. *et al.* (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37 (1), 1-13.

37. Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. Royal Stat. Soc. 57 (1), 289-300.

38. Downs, D.M. and Ernst, D.C. (2015) From microbiology to cancer biology: the Rid protein family prevents cellular damage caused by endogenously generated reactive nitrogen species: RidA stress response. Mol. Microbiol. 96 (2), 211-219.

39. Borchert, A.J. *et al.* (2019) Reactive enamines and imines *in vivo*: Lessons from the RidA paradigm. Trends Biochem. Sci. 44 (10), 849-860.

40. Christopherson, M.R. *et al.* (2008) YjgF is required for isoleucine biosynthesis when *Salmonella enterica* is grown on pyruvate medium. J. Bacteriol. 190 (8), 3057-3062.

41. Ernst, D.C. and Downs, D.M. (2016) 2-aminoacrylate stress induces a context-dependent glycine requirement in *ridA* strains of *Salmonella enterica*. J. Bacteriol. 198 (3), 536-543.

42. Miura, A. *et al.* (1981) Growth-rate-dependent regulation of ribosome synthesis in *Escherichia coli* - Expression of the *Iacz* and *galK* genes fused to ribosomal promoters. Cell 25 (3), 773-782.

43. Aseev, L.V. *et al.* (2016) Regulation of ribosomal protein operons *rplM-rpsI*, *rpmB-rpmG*, and *rplU-rpmA* at the transcriptional and translational levels. J. Bacteriol. 198 (18), 2494-2502.

44. Lemke, J.J. *et al.* (2011) Direct regulation of *Escherichia coli* ribosomal protein promoters by the transcription factors ppGpp and DksA. Proc. Natl. Acad. Sci. USA 108 (14), 5712-5717.

45. Lucock, M. (2000) Folic acid: Nutritional biochemistry, molecular biology, and role in disease processes. Mol. Genet. Metab. 71 (1-2), 121-138.

46. White, R.J. and Phillips, D.R. (1989) Bidirectional transcription footprinting of DNA-binding ligands. Biochemistry 28 (15), 6259-6269.

47. Webb, M. (1968) Pyruvate accumulation in growth-inhibited cultures of *Aerobacter aerogenes*. Biochem. J. 106 (2), 375-380.

48. Webb, M. (1958) Aminopterin inhibition in *Aerobacter aerogenes*: alanine and valine accumulation during the inhibition and their utilization on recovery. Biochem. J. 70 (3), 472-489.

49. Lal, P.B. *et al.* (2014) The redundant aminotransferases in lysine and arginine synthesis and the extent of aminotransferase redundancy in *Escherichia col*i: Aminotransferase redundancy. Mol. Microbiol. 94 (4), 843-856.

50. Yang, J. and Pittard, J. (2008) Biosynthesis of the aromatic amino acids. EcoSal Plus 3 (1).

51. Ravnikar, P.D. and Somerville, R.L. (1987) Genetic characterization of a highly efficient alternate pathway of serine biosynthesis in *Escherichia coli*. J. Bacteriol. 169, 7.

Table 6.1 A. Significantly altered (FDR < 0.05) metabolites in *S. enterica ridA* background

Compound	Wild-type (Mean ± Std)	*ridA* (Mean ± Std)	Fold Change	P Value	FDR value
Valine	1.23 ± 0.13	2.71 ± 0.13	2.20	8.47E-14	3.56E-12
Methionine	0.08 ± 0.01	0.17 ± 0.01	2.20	4.43E-12	9.31E-11
GSSG	0.54 ± 0.20	2.02 ± 0.18	3.71	1.99E-11	2.78E-10
CMP	0.05 ± 0.02	0.13 ± 0.02	2.94	1.38E-08	8.29E-08
2-isopropylmalic acid	0.32 ± 0.04	1.30 ± 0.45	4.03	7.54E-06	3.52E-05
Glucose	0.04 ± 0.01	0.06 ± 0.01	1.40	1.41E-05	5.37E-05
Maltose	1.41 ± 0.11	1.60 ± 0.10	1.13	2.13E-03	5.17E-03
UTP	0.04 ± 0.01	0.06 ± 0.01	1.43	8.79E-03	0.02
Glutamate	0.18 ± 0.04	0.25 ± 0.06	1.41	9.82E-03	0.02
Threonine	0.05 ± 0.01	0.06 ± 0.01	1.22	0.02	0.03
Pantothenic acid	0.13 ± 0.02	0.16 ± 0.02	1.17	0.03	0.04
Leucine	1.90 ± 0.25	0.80 ± 0.06	0.42	6.61E-10	6.94E-09
Lysine	0.59 ± 0.13	0.07 ± 0.02	0.12	1.53E-09	1.29E-08
Phenylalanine	0.13 ± 0.03	0.02 ± 0.01	0.12	1.38E-08	8.29E-08
dCMP	0.68 ± 0.08	0.44 ± 0.03	0.65	4.88E-07	2.56E-06
Succinic acid	2.49 ± 0.35	1.67 ± 0.20	0.67	1.39E-05	5.37E-05
AMP	0.70 ± 0.14	0.42 ± 0.05	0.60	2.92E-05	1.02E-04
NADP	0.05 ± 0.01	0.04 ± 0.01	0.74	9.71E-05	3.14E-04
Formate	0.16 ± 0.02	0.12 ± 0.02	0.76	2.79E-04	8.38E-04
N-acetylputrescine	0.16 ± 0.02	0.12 ± 0.02	0.77	1.11E-03	3.10E-03
Isoleucine	0.37 ± 0.12	0.22 ± 0.02	0.61	1.89E-03	4.95E-03
Acetate	2.50 ± 0.26	2.12 ± 0.18	0.85	2.22E-03	5.17E-03
Serine	0.06 ± 0.01	0.05 ± 0.01	0.87	0.01	0.03
Table 6.1 B. Metabolites below statistical threshold for differential accumulation					
N-acetyl alanine	0.02 ± 0.01	0.02 ± 0.01	1.24	0.07	0.10
Alanine	0.58 ± 0.12	0.68 ± 0.11	1.17	0.09	0.12
CDP	0.01 ± 0.01	0.02 ± 0.01	1.36	0.12	0.15
Uracil	0.06 ± 0.02	0.07 ± 0.03	1.25	0.22	0.27
Ethanolamine	0.04 ± 0.01	0.04 ± 0.02	1.15	0.45	0.53
NMN	0.34 ± 0.08	0.36 ± 0.07	1.05	0.61	0.68
Tyrosine	0.04 ± 0.02	0.04 ± 0.01	1.06	0.66	0.69
2-oxoglutaric acid	0.08 ± 0.12	0.09 ± 0.16	1.05	0.95	0.95
Malic acid	0.04 ± 0.01	0.03 ± 0.01	0.88	0.05	0.07
Putrescine	0.47 ± 0.21	0.34 ± 0.15	0.72	0.15	0.19
UMP	0.19 ± 0.04	0.17 ± 0.02	0.93	0.36	0.43
CTP	0.21 ± 0.04	0.20 ± 0.03	0.95	0.57	0.65
NAD	0.09 ± 0.04	0.08 ± 0.03	0.92	0.64	0.69
2-aminobutyric acid	1.17 ± 0.45	1.14 ± 0.09	0.97	0.84	0.86

[1]Abbreviations: GSSG: glutathione disulfide, CMP: cytidine monophosphate, UTP: uridine triphosphate, dCMP: deoxycytidine monophosphate, AMP: adenosine monophosphate, NADP: nicotinamide adenine dinucleotide phosphate, CDP: cytidine diphosphate, NMN: nicotinate mononucleotide, UMP: uridine monophosphate, CTP: cytidine triphosphate, NAD: nicotinamide adenine dinucleotide.

[2]Fold-change provided as *ridA*/WT

Table 6.2. KEGG pathways differentially regulated between wild-type and *ridA S. enterica*

Pathway Description	Relative Abundance[1]	Differential genes/Total genes	P value	FDR value
Thiamine metabolism (stm00730)	*ridA*	8/12	8.60E-07	4.80E-05
Biotin metabolism (stm00780)	*ridA*	6/14	6.60E-04	1.20E-02
Biosynthesis of amino acids (stm01230)	*ridA*	18/130	5.50E-04	1.50E-03
One-carbon pool by folate (stm00670)	*ridA*	5/13	4.40E-03	4.90E-02
ABC transporters (stm02010)	*ridA*	18/173	1.30E-02	8.40E-02
Glycine, serine, and threonine metabolism (stm00260)	*ridA*	7/35	1.10E-02	8.70E-02
Flagellar assembly (stm02040)	WT	31/37	3.00E-28	1.40E-26
Ribosome (stm03010)	WT	28/78	4.20E-12	1.00E-10
Bacterial chemotaxis (stm02030)	WT	11/23	3.60E-06	5.70E-05
Bacterial invasion of epithelial cells (stm05100)	WT	5/9	3.60E-03	4.20E-02

[1]Relative abundance denotes the *S. enterica* strain (wild-type or *ridA*) where expression of the pathways in question is higher.

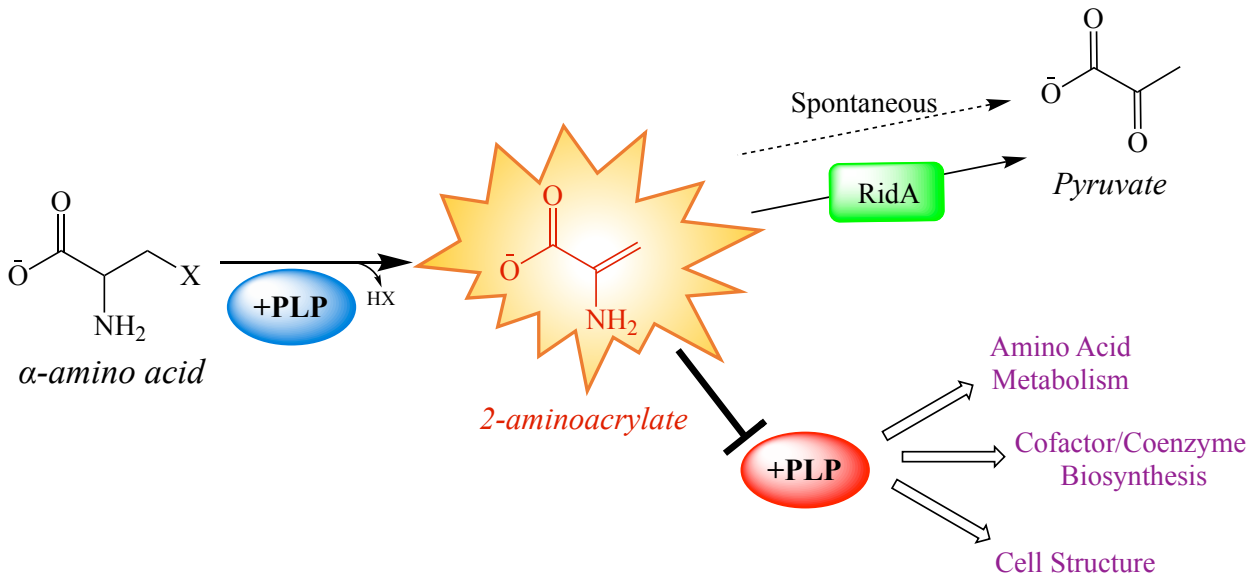


Figure 6.1. ***2-aminoacrylate stress paradigm in Salmonella enterica.*** Multiple pyridoxal 5'-phosphate(PLP)-dependent α,β-eliminases catalyze the β-elimination of 3-carbon α-amino acids, generating and releasing a reactive enamine product, 2-aminoacrylate (2AA). 2AA can be non-enzymatically hydrolyzed to pyruvate, a process that is catalyzed by RidA *in vivo*. In the absence of RidA, 2AA can accumulate endogenously, where it is able to attack the active site of various PLP-dependent enzymatic targets, rendering them inactive. Generally, PLP-dependent enzymes play biochemical roles related to amino acid metabolism, cofactor and coenzyme production, and cell structure synthesis (Appendix Table A.4). Consequently, many of the pathways making up these broad physiological processes are perturbed following 2AA accumulation [21-23, 41].

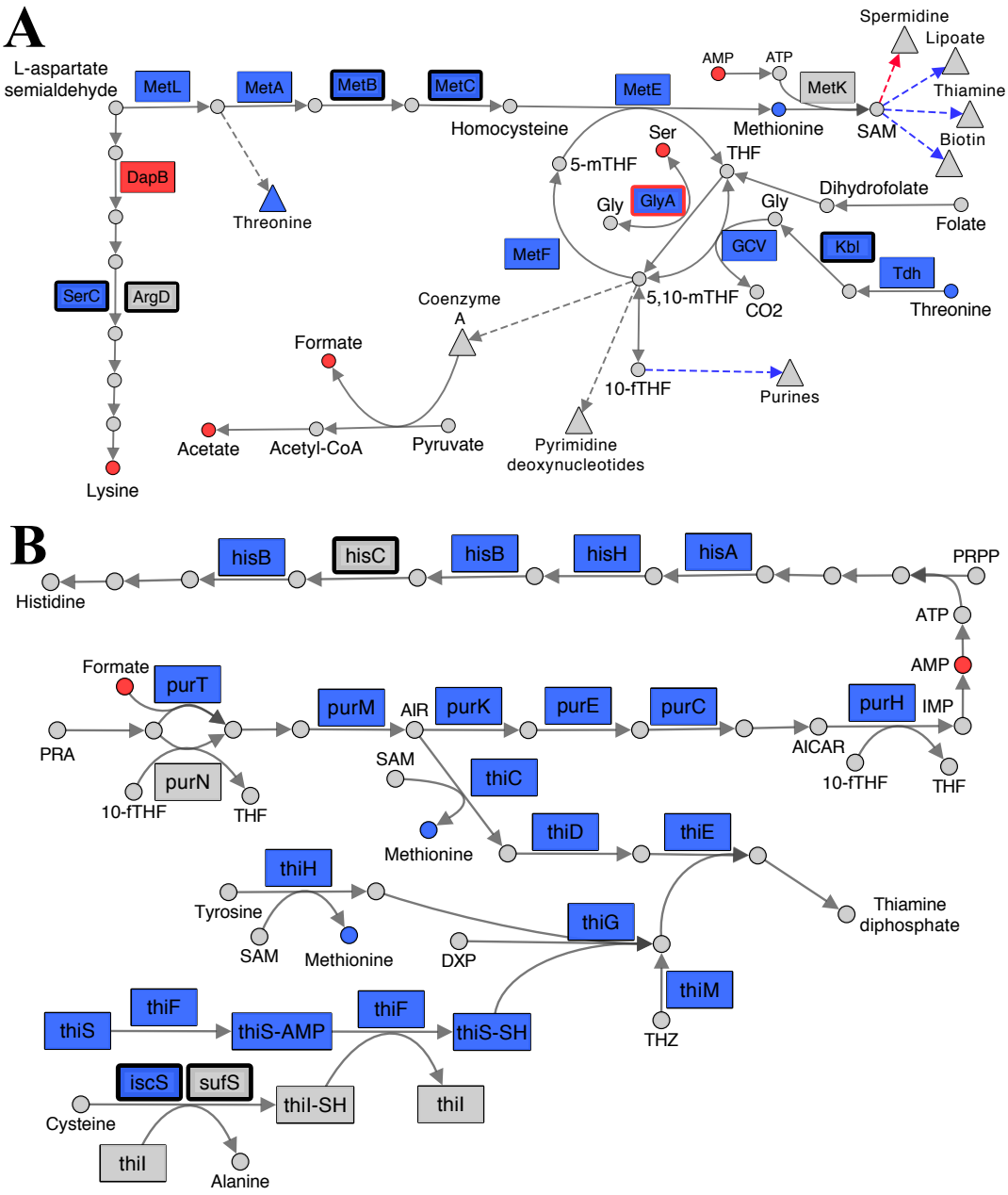


Figure 6.2. ***Folate and S-adenosylmethionine pathways are reorganized in a S. enterica ridA mutant.*** Metabolites (circles) and gene products (rectangles) involved in (A) lysine, methionine, and folate biosynthesis and (B) purine, histidine, and thiamine biosynthesis are depicted, with each arrow denoting a biochemical step in the pathway. Blue shading denotes significantly (FDR >0.05) increased abundance of the respective metabolite/transcript in *ridA* cells, while red shading indicates significantly decreased abundance in *ridA* cells. Grey gene products showed no significant difference between wild-type and *ridA* cells, while grey metabolites were either not identified or significantly altered. A bold border around a gene product denotes enzymes that use PLP as a cofactor. GlyA is a known target for 2AA damage, and is shown using a red border. Dashed arrows and triangles indicated general pathways which utilize the metabolite at the base of the arrow, while blue and red colored arrows denote that genes involved in biosynthesis of that compound up- or down-regulated, respectively.

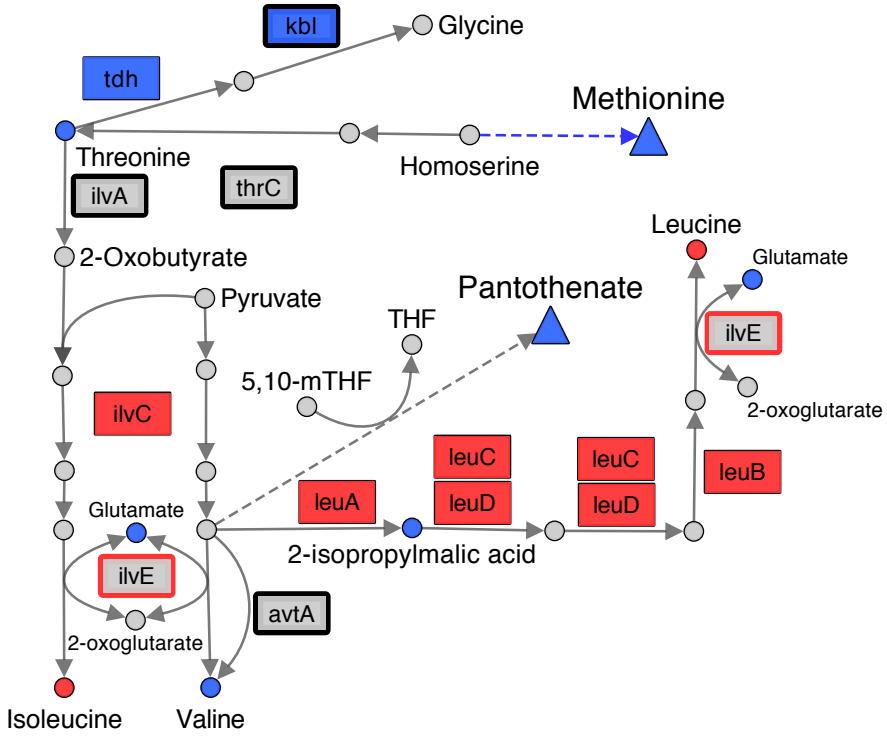


Figure 6.3. ***Branched-chain amino acid biosynthesis-related gene expression and metabolite abundance are altered in a S. enterica ridA mutant.*** Metabolites and gene products involved in branched-chain amino acid biosynthesis are shown using circles and rectangles, respectively. Blue shading denotes significantly (FDR >0.05) increased abundance of the respective metabolite/transcript in *ridA* cells, while red shading indicates significantly decreased abundance in *ridA* cells. Grey gene products showed no significant difference between wild-type and *ridA* cells, while grey metabolites were either not identified or significantly altered. A bold border around a gene product denotes enzymes that use PLP as a cofactor. IlvE is a known target for 2AA damage, and is shown using a red border. Dashed arrows and triangles indicated general pathways that utilize the metabolite at the base of the arrow.

CHAPTER 7

DAMAGE OF SERINE HYDROXYMETHYLTRANSFERASE BY 2-AMINOACRYLATE ELICITS A WIDESPREAD METABOLIC SHIFT IN *SALMONELLA ENTERICA* [1]

[1]Borchert AJ, Walejko JM, Le Guenec A, Ernst DC, Edison AS Downs DM. 2019. To be submitted to *mSystems* December, 2019.

7.1 ABSTRACT

In *Salmonella enterica*, the reactive intermediate deaminase (RidA; EC: 3.5.99.10) is responsible for deaminating reactive enamine species that form as intermediates in a number of amino acid metabolic pathways. When *S. enterica ridA* mutants are grown in minimal medium containing glucose as the sole carbon source, 2AA accumulates and damages serine hydroxymethyltransferase (GlyA). This damage leads to a growth-limiting depletion of 5,10-methelenetetrahydrofolate (5,10-mTHF), an import carbon carrier in the cell. The downstream metabolic consequence from GlyA damage by 2AA remains relatively unexplored. This study sought to use untargeted ^{1}H NMR metabolomics to identify the global downstream metabolic response of *S. enterica* undergoing a 2AA-dependent 5,10-mTHF limitation. Since supplementation of the growth medium circumvents the need for GlyA to generate 5,10-mTHF, glycine supplementation was used to discern between the effects stemming from GlyA damage and other consequences of RidA elimination and/or 2AA accumulation. The data supported the conclusion that expected perturbations to mixed acid fermentation and branched chain amino acid metabolism were consequences of 2AA-dependent GlyA damage. A number of other metabolic shifts in phenylalanine and putrescine were observed for a *ridA* mutant that were glycine status-dependent, though the mechanism linking their abundance to glycine/5,10-mTHF remains unclear. In total, the results helped demonstrate the value in using nutrient supplementations during ^{1}H NMR metabolomics experiments to effectively disentangle complex metabolic outcomes stemming from a general metabolic stress.

7.2 BACKGROUND

The metabolic state of the cell at any given time is complex and reflects the cumulative result from a number of inputs upon the network (transcription, translation, enzyme activity,

pathway integration, etc.). To this end, ascribing a metabolic output to a specific cellular process without the context that reductionist biochemical-genetic experiments provide can be difficult [1, 2]. Nonetheless, metabolomics approaches have been successful in accelerating the elucidation of complex metabolic and physiological states of an organism, where biochemical and genetic approaches can be slower and require vast resources [1, 3, 4]. Examples of studies that employ proton nuclear magnetic resonance (^{1}H NMR) spectroscopy metabolomics to dissect the complex physiological outcomes following genetic and environmental stress abound [5-7]. Additionally, ^{1}H NMR-based metabolomics shows many advantages over mass spectroscopy-associated approaches in simple sample preparation, wide chemical coverage, confident chemical assignments, and straightforward quantification of metabolites [4, 8-10]. Therefore, this work sought to use ^{1}H NMR metabolomics to broaden our understanding of the complex RidA/2-aminoacrylate paradigm in *S. enterica*.

Pyridoxal 5'-phosphate (PLP)-dependent α,β-eliminases generate reactive enamine species as important reaction intermediates from amino acid substrates. A subset of these α,β-eliminases can release these enamine intermediates into the cellular milieu. The **r**eactive **i**ntermediate **d**eaminase, RidA, catalyzes the deamination of free enamine species. 2-aminoacrylate is a reactive enamine species generated from L-serine by the biosynthetic serine/threonine dehydratase (IlvA; EC: 4.3.1.19) [11]. Importantly, IlvA is subject to allosteric control by L-isoleucine [12, 13]. During growth in minimal glucose medium, the absence of RidA in *S. enterica* allows the 2AA produced by IlvA to accumulate and damage multiple other PLP-dependent enzymes [11, 14-16]. This 2AA stress is manifest as a defined growth defect. Importantly, serine hydroxymethyltransferase (GlyA; EC: 2.1.2.1) acts as the most physiologically significant 2AA target, and the subsequent reduction in 5,10-methylenetetrahydrofolate (5,10-

mTHF) elicits the growth defect in these conditions [17]. The inclusion of exogenous glycine can circumvent this 5,10-mTHF limitation via the activity of the glycine cleavage complex (GCV). A summary of the RidA paradigm for *S. enterica* is provided in Figure 7.1.

However, while GlyA damage by 2AA is known to perturb glycine and 5,10-methylenetetrahydrofolate (5,10-mTHF) production, it remains unclear what global metabolic consequences stem from these perturbations. In this study, we used untargeted $_1$H NMR metabolomics and careful nutrient supplementation to expand our understanding of the global metabolic consequences associated with the accumulation of 2AA, specifically those stemming from 2AA-dependent damage of GlyA. Taking advantage of the simple sample preparation, high sample throughput, wide chemical coverage, and confident chemical assignments associated with $_1$H NMR, we measured both the endogenous and excreted metabolomes of *S. enterica* wild-type and *ridA* mutant strains following growth in various media conditions to further elucidate the RidA metabolic paradigm. More generally, this work demonstrates the utility in using well-characterized genetic perturbations and careful nutrient supplementation alongside metabolomics approaches as an effective means to dissect complex metabolic states of an organism.

7.3 MATERIALS AND METHODS

Bacterial strains, chemicals, and media. Strains used in this work are derivatives of *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2. Construction of the *ridA* null mutant (DM3480, *ridA3*::MudJ1734) is described elsewhere [18, 19]. The wild-type used in this work (DM9404) is an LT2 derivative isogenic to the *ridA* mutant. Rich medium was Difco nutrient broth (NB) (8 g/L) supplemented with 5 g/L NaCl. No-carbon E medium (NCE) containing 1 mM MgSO4 [20], trace metals [21], and 11 mM D-glucose was the designated minimal medium.

Minimal medium was supplemented with 1 mM L-isoleucine or 1 mM glycine, as indicated. All chemicals were purchased from the Sigma-Aldrich Chemical Company (St. Louis, MO).

Cell pellet and spent media acquisition. Ten biologically independent wild-type and *ridA* mutant cultures were grown overnight in NB medium shaking at 37 °C and used to inoculate (1% inoculum) 250 mL non-baffled flasks holding 125 mL one each of the three media types (minimal medium, minimal medium with isoleucine, and minimal medium with glycine), for a total of 60 flasks. Flasks were randomly arranged in an Innova®44 incubator and cultures allowed to grow 16 h at 37 °C, shaking at 180 RPM. Cultures were chilled 5 min on ice and then harvested by centrifugation at 7,000 x G for 10 min at 4 °C. The supernatant was decanted, with 10 mL transferred to sterile 15 mL conical tubes and flash-frozen using liquid nitrogen for downstream analysis of excreted metabolites. Cell pellets were transferred to sterile 15 mL conical tubes using 10 mL ddH_2O and then pelleted at 7,000 x G for 10 min at 4 °C. The supernatant was decanted and pellets were flash-frozen using liquid nitrogen and stored at -80 °C.

Preparation of growth medium samples. Spent media from each bacterial culture was lyophilized (VirTis Benchtop K) for 48 h. Once dry, each lyophilized sample was reconstituted in 150 μL of 100 mM sodium phosphate buffer, pH 7.0, containing 1/3 mM DSS (d6 4,4-dimethyl-4-silapentane-1-sulfonic acid) as an internal standard. Each sample was centrifuged at 20,000 X G for 30 min and 50 μL of supernatant was transferred by a Bruker SamplePro liquid handler into 1.7 mm SampleJet NMR tubes (Bruker Biospin).

Metabolite extraction from bacterial pellets. Each frozen bacterial pellet was thawed on ice and 1 mL of ice cold 80/20 methanol/water together with approximately 200 mL of 0.7mm silica beads (BioSpec products) added. Homogenization was carried out using a FastPrep 96 (MPBIO). The samples and extraction blanks went through three cycles of homogenization at

1800 rpm for 300 s each. At the end of each cycle samples and controls were centrifuged at 20000 x G for 30 min. Each supernatant was transferred to a new tube and 1 mL of ice-cold methanol/water added to the original tubes before each new cycle. The combined supernatants from each cycle were pooled and concentrated overnight using a CentriVap Benchtop Vacuum Concentrator (Labconco) down to 0.1 mL. The samples were then diluted with 0.5 mL of methanol/water and transferred into 0.6 mL centrifuge tube and concentrated to dryness. The extracts were reconstituted in 150 uL of deuterated 100 mM sodium phosphate buffer containing 1/3 mM of the internal standard DSS (d6 4,4-dimethyl-4-silapentane-1-sulfonic acid) at pH 7.0 and vortex mixed for 5 min. Each sample was centrifuged at 20000 x G for 30 min and transferred by a Bruker SamplePro liquid handler into 1.7 mm SampleJet NMR tubes.

Acquisition and processing of NMR data. One-dimensional ^{1}H NMR spectra for each sample were acquired using an optimized PROJECT (periodic refocusing of *J* evolution by coherence transfer) pulse sequence [22] on an Avance III HD 600 MHz Bruker NMR spectrometer equipped with a TCI cryoprobe and a Bruker SampleJet autosampler cooled to 5.6 °C. During acquisition, 32768 complex datapoints were collected for the FID, using 64 scans with 16 additional dummy scans. The spectral width was 20 ppm. A Fourier transform (FT), a polynomial baseline correction of order 3, a 2 Hz line broadening and phase correction were applied to each spectrum.

Two-dimensional ^{1}H-^{1}H total correlation spectroscopy (TOCSY), ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC) and ^{1}H-^{13}C HSQC–total correlation spectroscopy (HSQC–TOCSY) experiments were collected on pooled samples, composed from each study sample, for metabolite identification. During acquisition, all three experiments were collected for 32 scans and an additional 16 dummy scans, with 512 and 1024 datapoints recorded on the direct and indirect

dimensions respectively and, a spectral width of 200 ppm for ^{13}C and 12 ppm for ^{1}H. A 90 ms mixing time was used for both HSQC-TOCSY and TOCSY experiments.

All spectral processing was carried out using NMRpipe [23]. Spectra referencing, baseline correction, and statistical analysis were carried out using in-house Matlab (Mathworks, R2019a) scripts which are publicly available. (https://github.com/artedison/Edison_Lab_Shared_Metabolomics_UGA).

Compound identification/database matching. All three two-dimensional experiments were used for spectral matching to identify a total of 9 exogenous and 16 endogenous metabolites against the BBiorefcode library using COLMARm [24]. A confidence level ranging from 1 to 5, was assigned to each metabolite (Appendix Table A.8) as described elsewhere [25]. Briefly this scale is defined as: (1) putatively characterized compound, (2) matched reported 1D spectra, (3) matched reported HSQC spectra, (4) matched reported HSQC and HSQC-TOCSY spectra, and (5) validated by spiking putative compound into sample.

Statistical analysis. The data were normalized using probabilistic quotient normalization (PQN) and range-scaled before multivariate statistical analysis [26, 27]. The principal component analysis (PCA) scores were calculated using the NIPALS algorithm [28]. The partial least squares discriminant analysis (PLS-DA) using the SIMPLS algorithm was used to identify features that differed between the wild-type and *ridA* mutant for both endogenous and exogenous datasets [29]. Univariate Student's *t*-test statistics was performed on a subset of these features that were annotated to metabolites and had one or more isolated features in the one-dimensional NMR spectra. All raw and processed data are available on the Metabolomics Workbench (www.metabolomicsworkbench.org), along with detailed experimental NMR and statistical analysis methods.

7.4 RESULTS AND DISCUSSION

RidA elimination elicits global metabolic shifts in *S. enterica*. In a *ridA* mutant, 2AA accumulates and damages a number of PLP-dependent enzymes, causing a mild growth defect in minimal glucose medium (Fig. 7.2A,B). To gain a better understanding of the metabolic shifts which accompany 2AA stress in these conditions, the endogenous and exogenous metabolomes of a *ridA* mutant (DM3480) were compared to those obtained for the isogenic parental strain (DM9404, wild-type) using untargeted ^{1}H NMR metabolomics. Consistent with a significant metabolic shift accompanying the growth defect, principle component analysis (PCA) showed clear visual separation of both the exogenous and endogenous metabolomes from the *ridA* mutant and wild-type strains grown in minimal glucose medium (Fig. 7.2C-F). Notably, despite its advantages in reproducibility, sample preparation, and quantitative analysis, ^{1}H NMR's relatively low sensitivity and high level of peak overlap biases its output to a small number of abundant metabolites [4, 30]. Nonetheless, a clearly defined molecular signature was established describing the metabolic deviation of a *ridA* mutant, indicating that this approach could identify a number of biomarkers associated with RidA elimination and which represent a useful resource to explore metabolic pathways effected during 2AA accumulation.

Importantly, all RidA orthologs described to date share enamine/imine deaminase activity [11, 16, 31, 32]. However, other studies have suggested other functions for RidA, including a role as a translational inhibitor/ribonuclease [33-36] or a molecular chaperone [37-39]. Therefore, it was critical to determine whether the all of the metabolic attributes of a *ridA* mutant were 2AA-dependent or related to some other RidA-associated activity. During *S. enterica* growth in minimal glucose medium, IlvA acts as the predominant generator of 2AA [11]. As a consequence, supplementation of the growth medium with L-isoleucine, which allosterically inhibits IlvA

activity, eliminates production of 2AA and restores wild-type growth to the *ridA* mutant [11] (Fig 7.2A). Therefore, L-isoleucine supplementation was used to discern between 2AA-dependent and 2AA-independent metabolic effects. Addition of 1 mM exogenous L-isoleucine to the growth medium abolished the differences between the *ridA* and wild-type strains endogenous (Fig 7.2C) and exogenous (Fig 7.2E) metabolomes, as gathered from PCA analyses. This finding was consistent with the dominant consequence of RidA elimination on the metabolic network being mediated through 2AA accumulation and not a 2AA-independent mechanism.

The free 2AA that accumulates in an *S. enterica ridA* mutant is able to covalently modify and damage the active site of GlyA, whose activity is reduced to only 20 % of wild-type activity in minimal glucose medium [14, 17]. GlyA catalyzes the transfer of the hydroxymethyl group from L-serine substrate to THF, forming glycine and 5,10-mTHF (Fig. 7.1). Importantly, 2AA-dependent damage of GlyA results in a growth-limiting reduction to 5,10-mTHF production, prompting GlyA's designation as the most physiologically sensitive target in *S. enterica* [14, 17]. Supplementation of the medium with glycine restores growth to a *ridA* mutant by bypassing the need for GlyA-dependent 5,10-mTHF synthesis (Fig 7.1; Fig 7.2B). However, it was possible that the detectable metabolic shifts seen for a *ridA* mutant strain did not reflect the most physiologically pertinent metabolic perturbations (i.e. those resulting from a decrease in 5,10-mTHF production). Therefore, glycine supplementation was used to discern between the metabolic effects resultant from 2AA-dependent damage of GlyA and those stemming from some other 2AA-dependent perturbations. PCA analysis showed that, upon addition of 1 mM glycine to the growth medium, the *ridA* mutant strain-associated endogenous metabolome could no longer be distinguished from wild-type's (Fig 7.2D). Therefore, the bulk of the metabolic perturbations observed for a *ridA* mutant strain's endogenous metabolites appear linked to GlyA damage. Similarly, when the

excreted metabolome of a *ridA* mutant grown in minimal glucose medium containing glycine was compared to the corresponding metabolome from wild-type by PCA analysis, a clear decrease in the separation of these metabolic signatures was observed, however their separation was not completely eliminated.

Altogether, these data supported the conclusions that (i) ^{1}H NMR metabolomics could generate a unique molecular signature describing an *S. enterica ridA* mutant strain, (ii) the deviation in molecular signatures between *ridA* and wild-type strains was dependent upon the generation of 2AA, and (iii), in large part, described the physiologically relevant metabolic consequence of 2AA-dependent GlyA damage.

2AA stress affects amino acid metabolism and mixed acid fermentation. In total, 16 endogenous metabolites and 9 exogenous metabolites could be identified from the NMR spectra (Appendix Table A.8). To determine the specific metabolic pathways perturbed by 2AA stress, partial least squares differential analysis (PLSda) scores plots were used to further discern the metabolic differences between *ridA* mutant and wild-type strain grown in minimal glucose medium (Fig. 7.3A,C). Variable importance in projections (VIP) scores were determined for all NMR features in the endogenous and exogenous datasets that contributed most to the PLSda separation in the first component. From these data, VIP plots were created from identified metabolites contributing to separation of endogenous (Fig. 7.3B) and exogenous (Fig. 7.3D) metabolomes. Importantly, the VIP score reported in these plots represent the average of VIP scores taken from all the features comprising a given metabolite. The VIP analysis revealed that threonine, valine, N-acetyl putrescine, glutamine, phenylalanine, alanine, glutamate, and pyruvate, which were elevated in a *ridA* mutant background, and ethanolamine, formate, putrescine, and coenzyme A (CoA), which were diminished in the *ridA* mutant background, drove PLSda

separation for endogenous metabolites, with VIP scores > 1 (Fig. 7.3B). Similarly, VIP analysis revealed that lactate, valine, putrescine, 2-isopropylmalate, and acetyl phosphate, which were elevated in a *ridA* mutant background, and formate, uracil, and 2-aminobutyrate, which were diminished in the *ridA* mutant background, drove PLSda separation for exogenous metabolites (Fig. 7.3D). Generally, these data suggested that several metabolites relating to amino acid metabolism and mixed acid fermentation were altered in a *ridA* mutant. Integration of representative peaks corresponding to all the identifiable metabolites and comparison by Student's unpaired two-samples *t*-test revealed that 12/16 endogenous metabolites and 8/10 exogenous metabolites were significantly altered in a *ridA* mutant strain ($P < 0.05$, Appendix Table A.9, Fig. 7.4).

Integration of the same representative peaks from the samples grown in minimal glucose medium supplemented with isoleucine revealed that concentrations of the significantly altered metabolites were restored to wild-type levels. (Appendix Table A.9, Fig. 7.4). The only exception was in the case of exogenous uracil abundance, which was 3.1-fold lower for a *ridA* mutant compared to wild-type following growth in minimal glucose medium, but was elevated 1.3-fold compared to wild-type when grown in minimal glucose medium containing isoleucine. In this case, the peak used for integration was poorly defined and likely influenced accurate integration (Data not shown). Nonetheless, isoleucine reversed the metabolic alteration observed for a *ridA* mutant grown in minimal glucose medium. Altogether, these results supported the conclusion that the measured changes to identifiable metabolite abundance observed in a *ridA* mutant strain were all the consequence of IlvA-associated 2AA stress.

Untargeted metabolomics captures the known CoA limitation of a *ridA* mutant. During growth on minimal glucose medium, *S. enterica* derives most of its one-carbon units from

serine via generation of 5,10-mTHF by PLP-dependent GlyA [40]. GlyA activity is decreased more than five-fold in a *ridA* mutant, when compared to wild-type, as a result of its damage by 2AA [14]. A major consequence of this damage is reduced flux through the 5,10-mTHF-dependent 3-methyl-2-oxobutanoate hydroxymethyltransferase (PanB; EC: 2.1.2.11), which is important for the biosynthesis of pantothenate, the precursor to CoA [14]. Accordingly, previous work identified a significant (3-fold) reduction to CoA levels in a *ridA* mutant [14], a finding also captured by this untargeted metabolomic experiment. Since the degradation of glycine by the inducible GCV complex generates additional 5,10-mTHF [41], exogenous glycine was shown to circumvent the 5,10-mTHF requirement of a *ridA* mutant and significantly increase, but not fully restore to wild-type, the CoA levels in this background [14, 17]. The addition of exogenous glycine to the growth medium also reduced the discrepancy between the *ridA* mutant and wild-type glycine levels assessed from the untargeted metabolomics data (from 1.6-fold to 1.3-fold reduced in a *ridA* mutant grown in minimal medium vs minimal medium containing glycine, respectively) (Appendix Table A.9, Fig. 7.4).

The finding that glycine supplementation did not fully restore CoA to wild-type levels may indicate that the inclusion of 1 mM exogenous glycine was insufficient to completely recover wild-type 5,10-mTHF levels or that CoA biosynthesis is constricted in a 5,10-mTHF-independent manner. Importantly, this level of glycine supplementation was able to restore growth to a *ridA* mutant (Fig. 7.2B), indicating that sufficient amounts of 5,10-mTHF were generated for growth. Additionally, a previously characterized consequence of CoA limitation during 2AA stress was the accumulation and secretion of pyruvate into the growth medium during late exponential growth phase [14]. This work found that glycine supplementation eliminated the pyruvate accumulation

phenotype but only restored CoA levels to 80 % of wild-type, suggesting normal flux through central carbon metabolism was possible in the presence of a modest reduction to CoA [14].

Identification of downstream metabolic effects of 2AA-dependent damage to GlyA. Since glycine supplementation overcame most of the metabolic shifts observed for a *ridA* mutant, it was gathered that a majority of the metabolic alterations in this background were a consequence of 2AA-dependent damage to GlyA. Specifically, endogenous valine, threonine, glutamine, N-acetyl putrescine, phenylalanine, glutamic acid, ethanolamine, acetic acid, putrescine and formic acid levels were indistinguishable from wild-type following glycine supplementation of the growth medium (Fig. 7.4A). Exogenous levels of 2-isopropylmalic acid, putrescine, acetyl-phosphate, and formic acid also returned to wild-type following glycine addition (Fig. 7.4B). Importantly, the discrepancy between exogenous levels of both valine and uracil was reduced following the addition of exogenous glycine (5-fold higher to 1.4-fold higher and 3.1-fold lower to 1.5-fold lower in the *ridA* background for valine and uracil, respectively), suggesting that the state of glycine/5,10-mTHF was in part responsible for their relative concentration shift in a *ridA* mutant (Appendix Table A.9, Fig. 7.4). Also, while the difference between lactic acid excretion for the *ridA* strain vs. wild-type could not be determined since wild-type lactic acid levels were below the limit of detection, the absolute quantity of lactic acid observed for a *ridA* mutant decreased six-fold upon the addition of exogenous glycine (Appendix Table A.9). This again suggested that lactic acid accumulation was a consequence, at least in part, of the glycine/5,10-mTHF limitation in a *ridA* mutant, though this could not be determined definitively since the effect of glycine on wild-type lactic acid formation is unknown.

Most striking, the glycine status-dependent decrease in endogenous formic acid and acetic acid, as well as increase in exogenous lactic acid suggested a model whereby the CoA limitation

described above provoked a shift in flux through mixed acid fermentation (Fig. 7.5). During mixed acid fermentation, pyruvate-formate lyase (PflB: EC: 2.3.1.54) uses CoA and pyruvate as substrates for the production of formate and acetyl-CoA, and then acetyl-CoA is further processed to acetate [42]. A bottleneck in CoA biosynthesis is expected to reduce PflB-dependent formation of formate and downstream production of acetate. The accumulation of pyruvate during late exponential phase growth in a *ridA* mutant is consistent with this model [14]. An increase in endogenous pyruvate was also observed for a *ridA* mutant by ^{1}H NMR, however, with a p-value of only 0.1, this finding was not statistically significant (Appendix Table A.9). The observation that lactic acid, which is derived from pyruvate in a CoA-independent manner, accumulates in the growth medium of a *ridA* mutant suggests that overflow pyruvate may be diverted toward lactate synthesis (Fig. 7.5). Additionally, since two molecules of pyruvate are used during valine synthesis, work in *Klebsiella aerogenes* shows that valine production can increase as a consequence of pyruvate accumulation [43, 44]. Therefore, the glycine-dependent increase to valine content may also indicate that the overflow pyruvate in a *ridA* mutant is diverted toward valine synthesis (Fig. 7.4, Fig. 7.5).

Excretion of 2-isopropylmalic acid is also increased in a glycine-dependent manner for a *ridA* mutant (Fig. 7.4B). 2-isopropylmalate synthase (LeuA, EC: 2.3.3.13) uses 2-keto-isovalerate, which is a valine intermediate, and the acyl group from acetyl-CoA to synthesize 2-isopropylmalate. Since 2-isopropylmalate production is dependent upon acetyl-CoA, and acetyl-CoA is likely decreased in a *ridA* mutant, increased 2-isopropylmalate excretion was surprising. Importantly, other intermediates in this pathway were not detected, so it is unclear whether 2-isopropylmalate is the only leucine intermediate that accumulates in a *ridA* background. Since 2-isopropylmalate's precursor is a valine intermediate, the most straightforward explanation is that

increased 2-isopropylmalate excretion is also a consequence of pyruvate overflow toward valine biosynthesis (Fig. 7.5).

7.5 CONCLUSIONS

This study used untargeted ^{1}H NMR metabolomics to inform the global metabolic consequence RidA elimination in *S. enterica*. Significant metabolic shifts centered around mixed acid fermentation, and branched-chain amino acid metabolism. The addition of exogenous L-isoleucine, which eliminates 2AA production by IlvA, abolished separation of the wild-type and *ridA* mutant metabolomes by PCA and eliminated all concentration shifts for the metabolites detected. This supported the assumption that the majority of metabolic effects stemming from RidA elimination were a consequence of 2AA accumulation in this background and not a result of some RidA-dependent 2AA-independent effect. Further, glycine addition decreased PCA separation between wild-type and *ridA* mutant metabolomes and restored wild-type levels for 10/12 endogenous and 4/8 exogenous metabolites that were confidently identified and differed between the two populations. This finding supports the conclusion that the bulk of global metabolic shifts that occur during 2AA stress of *S. enterica* in minimal medium are a consequence of GlyA damage by 2AA.

A model describing the shifts in mixed-acid fermentation and BCAA products was constructed based upon the previously characterized CoA limitation a *ridA* mutant experiences during 2AA stress (Fig. 7.5). Further validation of this model may incorporate the supplementation of growth medium with pantothenate, which restores CoA levels back to wild-type in a manner independent from 5,10-mTHF. Additionally, alterations in the endogenous content of threonine, glutamine, glutamate, putrescine, N-acetyl putrescine, and ethanolamine were also dependent upon supplementation of glycine in the growth medium. This observation suggests that the changes in

these metabolites may also be a consequence downstream from GlyA damage in a *ridA* mutant, though the direct link to glycine/5,10-mTHF remains unknown. Pantothenate addition could help determine whether alterations in these metabolites are dependent upon the CoA state of the cell. Glycine supplementation in wild-type and *ridA* mutant cells where the GCV complex is disrupted may also prove useful in determining whether the ability for glycine supplementation to suppress the phenotype is 5,10-mTHF-related or mediated through glycine itself. The small RNA (sRNA) GcvB, whose expression is controlled by glycine, is known to regulate translation of a number of amino acid uptake and biosynthesis enzymes [45]. It is possible that GlyA damage by 2AA alters expression of GcvB, leading to a number of downstream metabolic effects. Therefore, glycine supplementation in a *gcvB* mutant vs. a *ridA gcvB* mutant can help inform whether the effect of glycine in a *ridA* mutant is a regulatory one [17].

In the context of the RidA paradigm, the exogenous metabolome shifts for mixed-acid fermentation products, valine, and putrescine mimicked many of the endogenous metabolic shifts in a *ridA* mutant. Since fewer metabolites are present in the spent media samples, the spectra associated with the exogenous metabolome displayed markedly better peak separation. Media samples also did not require homogenization or methanol extraction, making their processing more expedient and straightforward. Therefore, analysis of exogenous samples may offer a high-throughput and simplified way to continue characterization of the RidA paradigm. Such simplified and expedited analysis would be particularly useful during time-course experiments or studies containing more genetic backgrounds and/or media conditions, where hundreds of samples may be required.

Overall, the combination of ^{1}H NMR metabolomics and careful nutrient supplementation was successful in expanding the complex RidA/2-aminoacrylate paradigm in *S. enterica* and in

making first steps toward delineating downstream consequences of GlyA damage from metabolic effects independent of glycine/5,10-mTHF perturbation. The RidA system provides an interesting case study in refining the experimental approach of metabolomics studies to include genetic manipulations and nutrient supplementations as a means to probe the underlying factors contributing to the metabolic shifts observed following a perturbation. This approach may prove useful during high-throughput characterization of other complex metabolic systems, such as with oxidative stress or the microbial response to an antibiotic, where the outcome from a given stress is expected to be multifactored.

7.6 REFERENCES

1. Downs, D. *et al.* (2018) The three-legged stool of understanding metabolism: integrating metabolomics with biochemical genetics and computational modeling. AIMS Microbiology 4 (2), 289-303.

2. Johnson, C.H. *et al.* (2016) Metabolomics: beyond biomarkers and towards mechanisms. Nat. Rev. Mol. Cell Biol. 17 (7), 451-9.

3. Benjamin, D.I. *et al.* (2012) Global profiling strategies for mapping dysregulated metabolic pathways in cancer. Cell Metab. 16 (5), 565-77.

4. Markley, J.L. *et al.* (2017) The future of NMR-based metabolomics. Curr. Opin. Biotechnol. 43, 34-40.

5. Walejko, J.M. *et al.* (2019) Chronic maternal cortisol excess during late gestation leads to metabolic alterations in the newborn heart. Am. J. Physiol. Endocrinol. Metab. 316 (3), E546-E556.

6. Lee-McMullen, B. *et al.* (2019) Age-dependent changes in metabolite profile and lipid saturation in dystrophic mice. NMR Biomed., e4075.

7. Hattori, A. *et al.* (2017) Cancer progression by reprogrammed BCAA metabolism in myeloid leukaemia. Nature 545 (7655), 500-504.

8. Aretz, I. and Meierhofer, D. (2016) Advantages and pitfalls of mass spectrometry based metabolome profiling in systems biology. Int. J. Mol. Sci. 17 (5).

9. Nagana Gowda, G.A. *et al.* (2015) Expanding the limits of human blood metabolite quantitation using NMR spectroscopy. Anal. Chem. 87 (1), 706-15.

10. Nagana Gowda, G.A. and Raftery, D. (2015) Can NMR solve some significant challenges in metabolomics? J. Magn. Reson. 260, 144-60.

11. Lambrecht, J.A. *et al.* (2013) RidA proteins prevent metabolic damage inflicted by PLP-dependent dehydratases in all domains of life. mBio 4 (1), e00033-13-e00033-13.

12. LaRossa, R.A. and Van Dyk, T.K. (1987) Metabolic mayhem caused by 2-ketoacid imbalances. Bioessays 7, 125-130.

13. Umbarger, H.E. and Brown, B. (1957) Threonine deamination in *Escherichia coli* ii.: Evidence for two l-threonine deaminases. J. Bacteriol. 73 (1), 105.

14. Flynn, J.M. *et al.* (2013) Decreased coenzyme A levels in *ridA* mutant strains of *Salmonella enterica* result from inactivated serine hydroxymethyltransferase: *ridA* mutants are deficient in one carbon metabolism. Mol. Microbiol. 89 (4), 751-759.

15. Flynn, J.M. and Downs, D.M. (2013) In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5'-phosphate cofactor. J. Bacteriol. 195 (16), 3603-3609.

16. Borchert, A.J. and Downs, D.M. (2017) The response to 2-aminoacrylate differs in *Escherichia coli* and *Salmonella enterica*, despite shared metabolic components. J. Bacteriol. 199 (14), e00140-17.

17. Ernst, D.C. and Downs, D.M. (2016) 2-aminoacrylate stress induces a context-dependent glycine requirement in *ridA* strains of *Salmonella enterica*. J. Bacteriol. 198 (3), 536-543.

18. Schmitz, G. and Downs, D.M. (2004) Reduced transaminase B (IlvE) activity caused by the lack of *yjgF* is dependent on the status of threonine deaminase (IlvA) in *Salmonella enterica* Serovar Typhimurium. J. Bacteriol. 186 (3), 803-810.

19. Castilho, B.A. *et al.* (1984) Plasmid insertion mutagenesis and *lac* gene fusion with mini-mu bacteriophage transposons. J. Bacteriol. 158 (2), 488-495.

20. Vogel, H.J. and Bonner, D.M. (1956) Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. 218 (1), 97-106.

21. Balch, W.E. and Wolfe, R.S. (1976) New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressureized atmosphere. Appl. Environ. Microbiol. 32 (6), 781-791.

22. Le Guennec, A. *et al.* (2017) Alternatives to nuclear Overhauser enhancement spectroscopy presat and Carr-Purcell-Meiboom-Gill presat for NMR-based metabolomics. Anal. Chem. 89 (17), 8582-8588.

23. Delaglio, F. *et al.* (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6 (3), 277-293.

24. Bingol, K. *et al.* (2016) Comprehensive metabolite identification strategy using multiple two-dimensional NMR spectra of a complex mixture implemented in the COLMARm web server. Anal. Chem. 88 (24), 12411-12418.

25. Walejko, J.M. *et al.* (2018) Global metabolomics of the placenta reveals distinct metabolic profiles between maternal and fetal placental tissues following delivery in non-labored women. Metabolites 8 (1).

26. Dieterle, F. *et al.* (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in $_1$H NMR metabonomics. Anal. Chem. 78 (13), 4281-4290.

27. Smilde, A.K. *et al.* (2005) Fusion of mass spectrometry-based metabolomics data. Anal. Chem. 77 (20), 6729-6736.

28. Wold, H. (1975) Soft modelling by latent variables: The non-linear iterative partial least squares (NIPALS) approach. Journal of Applied Probability 12 (S1), 117-142.

29. de Jong, S. (1993) SIMPLS: An alternative approach to partial least squares regression. Chemometrics Intellig. Lab. Syst. 18 (3), 251-263.

30. Emwas, A.H. *et al.* (2019) NMR spectroscopy for metabolomics research. Metabolites 9 (7).

31. Niehaus, T.D. *et al.* (2014) *Arabidopsis* and maize RidA proteins preempt reactive enamine/imine damage to branched-chain amino acid biosynthesis in plastids. Plant Cell 26 (7), 3010-3022.

32. ElRamlawy, K.G. *et al.* (2016) Der f 34, a novel major house dust mite allergen belonging to a highly conserved Rid/YjgF/YER057c/UK114 family of imine deaminases. J. Biol. Chem. 291 (41), 21607-21615.

33. Oka, T. *et al.* (1995) Isolation and characterization of a novel perchloric acid-soluble protein inhibiting cell-free protein synthesis. J. Biol. Chem. 270 (50), 30060-7.

34. Morishita, R. *et al.* (1999) Ribonuclease activity of rat liver perchloric acid-soluble protein, a potent inhibitor of protein synthesis. J. Biol. Chem. 274 (29), 20688-92.

35. Schmiedeknecht, G. *et al.* (1996) Isolation and characterization of a 14.5-kDa trichloroacetic-acid-soluble translational inhibitor protein from human monocytes that is upregulated upon cellular differentiation. Eur. J. Biochem. 242 (2), 339-51.

36. Walejko, J.M. *et al.* (2018) Gut microbiota and serum metabolite differences in African Americans and White Americans with high blood pressure. Int. J. Cardiol. 271, 336-339.

37. Samuel, S.J. *et al.* (1997) Hrp12, a novel heat-responsive, tissue-specific, phosphorylated protein isolated from mouse liver. Hepatology 25 (5), 1213-22.

38. Farkas, A. *et al.* (2004) DUK114, the *Drosophila* orthologue of bovine brain calpain activator protein, is a molecular chaperone. Biochem. J. 383 (Pt 1), 165-70.

39. Müller, A. *et al.* (2014) Activation of RidA chaperone function by N-chlorination. Nat. Commun. 5, 5804.

40. Green, J.M. *et al.* (1996) Folate biosynthesis, reduction, and polyglutamylation. In *Escherichia coli* and *Salmonella* Cellular and Molecular Biology (Neidhart, F.C. ed), pp. 665-673, ASM Press.

41. Stauffer, L.T. and Stauffer, G.V. (1999) Role for the leucine-responsive regulatory protein (Lrp) as a structural protein in regulating the *Escherichia coli gcvTHP* operon. Microbiology 145 (Pt 3), 569-76.

42. Trotter, E.W. *et al.* (2011) Reprogramming of *Escherichia coli* K-12 metabolism during the initial phase of transition from an anaerobic to a micro-aerobic environment. PLoS ONE 6 (9), e25501.

43. Webb, M. (1958) Aminopterin inhibition in *Aerobacter aerogenes*: alanine and valine accumulation during the inhibition and their utilization on recovery. Biochem. J. 70 (3), 472-489.

44. Webb, M. (1968) Pyruvate accumulation in growth-inhibited cultures of *Aerobacter aerogenes*. Biochem. J. 106 (2), 375-380.

45. Sharma, C.M. *et al.* (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. Mol. Microbiol. 81 (5), 1144-65.

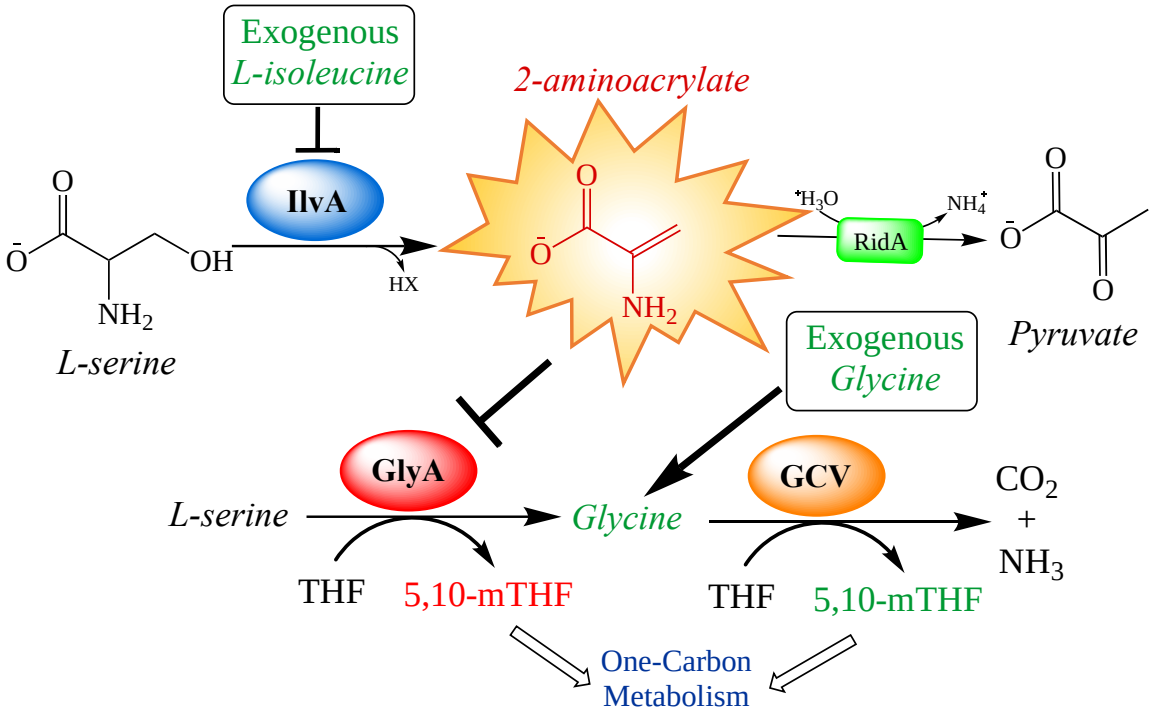


Figure 7.1: The RidA paradigm of 2-aminoacrylate stress in *S. enterica*. Biosynthetic serine/threonine dehydratase (IlvA), which is sensitive to allosteric inhibition by L-isoleucine, catalyzes the β-elimination of L-serine to generate the reactive enamine 2-aminoacrylate (2AA). 2AA is hydrolyzed to pyruvate by the reactive intermediate deaminase A (RidA). In RidA's absence, 2AA accumulates and can damage a number of PLP-dependent enzymes. The most physiologically sensitive target of 2AA damage in *S. enterica* grown in minimal glucose medium is serine hydroxymethyltransferase (GlyA). GlyA is responsible for the reversible transfer of the hydroxymethyl from serine to tetrahydrofolate (THF), generating glycine and 5,10-methylenetetrahydrofolate (5,10-mTHF). The glycine cleavage complex (GCV) can further catabolize glycine, generating additional 5,10-mTHF.

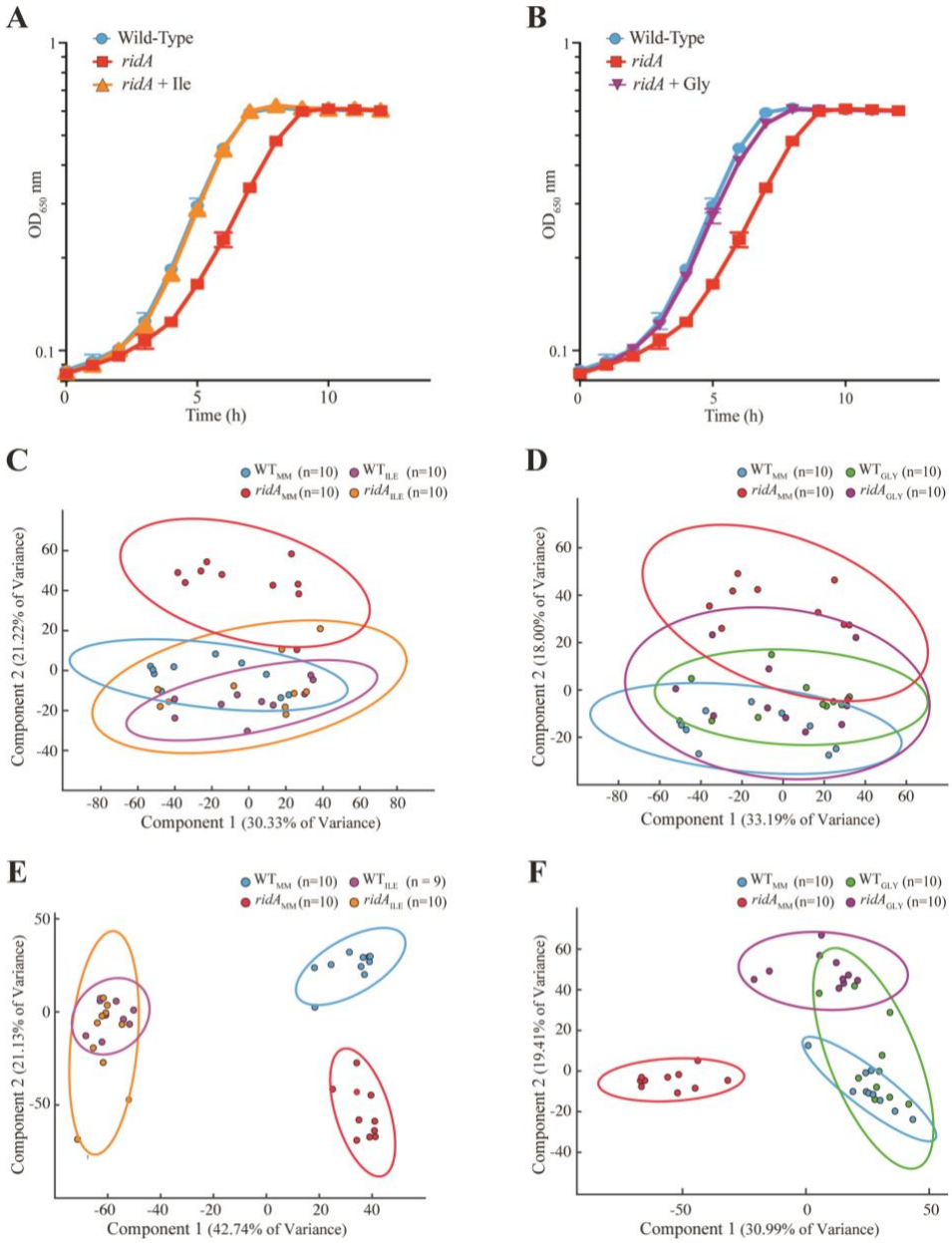


Figure 7.2. Supplementation of the growth medium with isoleucine or glycine restores the growth behavior and metabolic profile of a *ridA* strain back to wild-type levels. *S. enterica* wild-type and *ridA* mutant strains were grown at 37 °C in minimal glucose (11 mM) medium. Supplementation of the medium with (A) 1 mM L-isoleucine or (B) 1 mM glycine restored wild-type growth to a *ridA* mutant. Data are the means from three biological replicates, where error bars represent the 95% confidence interval. Principle component analysis (PCA) scores plots show separation of endogenous and exogenous metabolic profiles for *S. enterica* wild-type and *ridA* mutant strains following 16 h growth in minimal glucose medium at 37 °C. Supplementation of medium with isoleucine eliminated separation of wild-type and *ridA* mutant groups for both endogenous and exogenous metabolites (A and B). Glycine supplementation eliminated separation of wild-type and *ridA* mutant groups for endogenous metabolites (C) and reduced separation of these groups for exogenous metabolites (D). Colored ellipses represent the 95% confidence interval for each group.

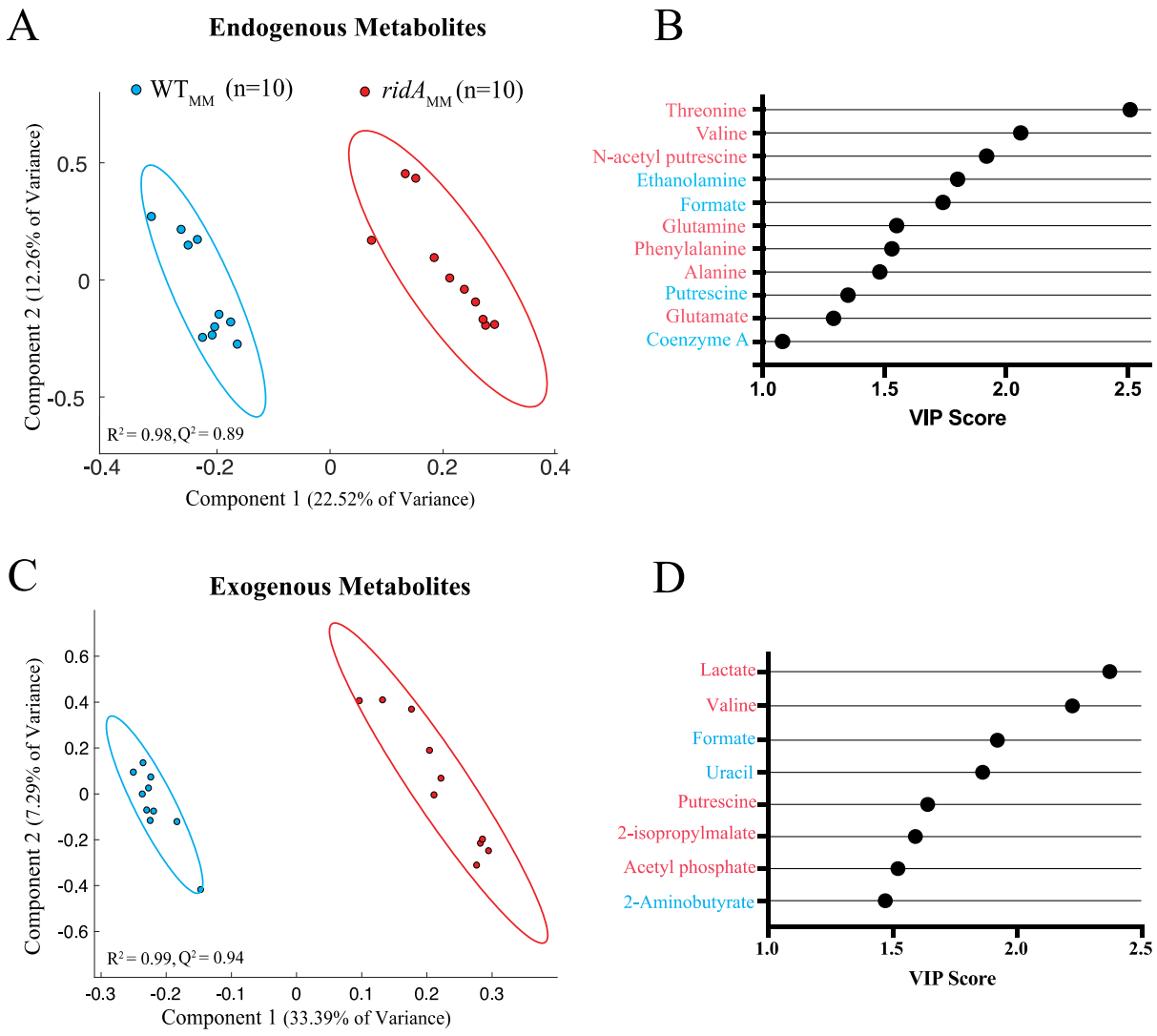


Figure 7.3. Partial least squares discriminant analysis (PLSda) highlights separation of *S. enterica ridA* and wild-type metabolomic profiles following growth in minimal glucose medium. PLSda scores plots show clear separation by PLS component 1 of wild-type (n=10, blue) and *ridA* mutant (n=10, red) endogenous (A) and exogenous (B) metabolite samples, following growth in minimal glucose medium. Variable importance of projection (VIP) scores were plotted for the metabolites that contributed significantly (VIP > 1) to separation by PLSda Component 1 for endogenous (C) and exogenous (D) samples. A higher VIP score indicates a greater contribution of that metabolite to separation by PLSda component 1. VIP scores were determined as the average from all peak VIP scores belonging to the given metabolite. Metabolites colored blue were elevated in wild-type samples, while those colored red were elevated in *ridA* mutant samples.

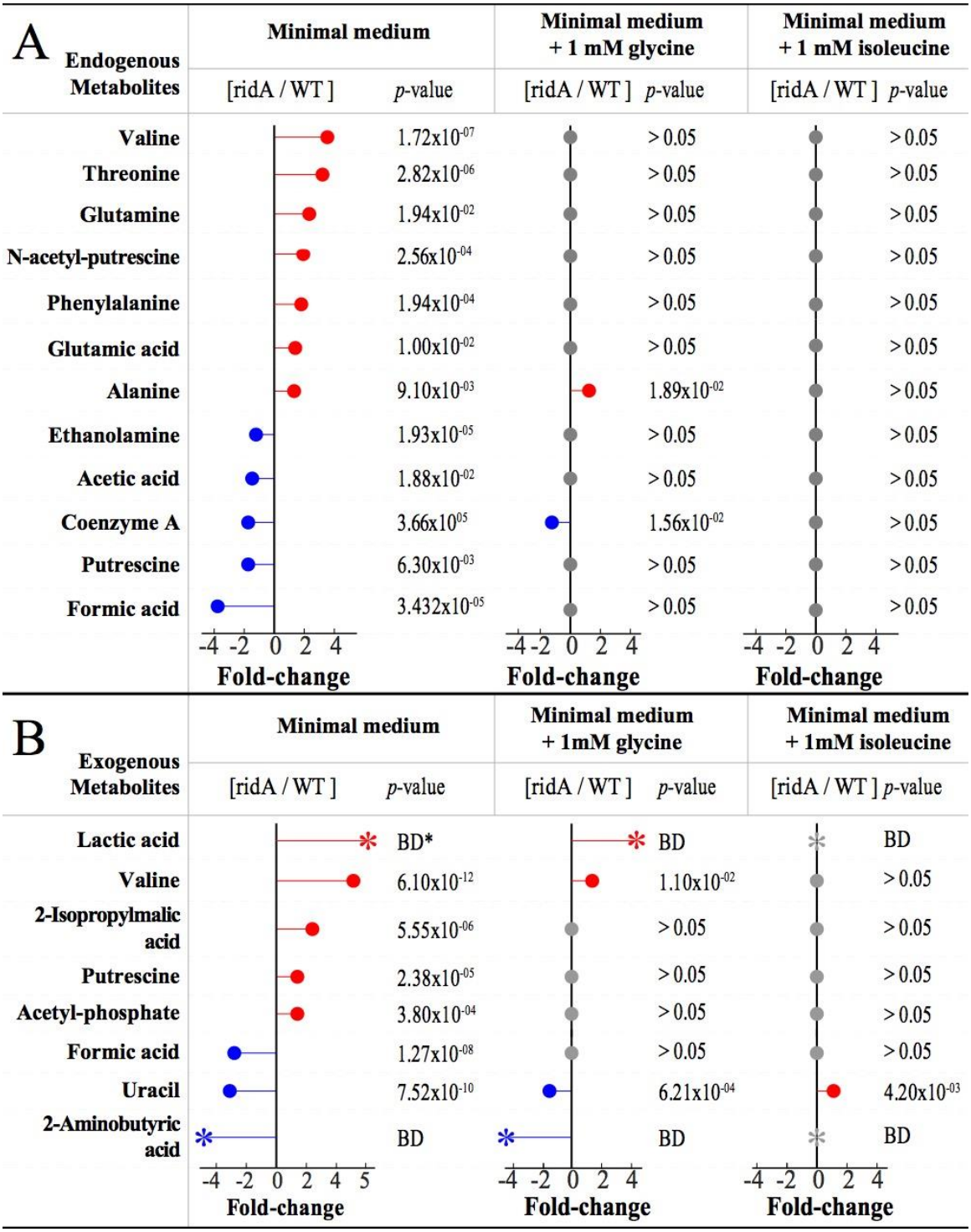


Figure 7.4. Significantly altered metabolites in a *ridA* mutant grown under different conditions. Fold-changes were calculated for both intracellular metabolites (A) and excreted metabolites (B). Red colored circles indicate higher abundance in *ridA* mutants while blue colored circles indicate lower metabolite levels in respect to wild type. No significant differences between metabolite levels are indicated by grey colored circles at a significance level of 0.05. Asterisks, with corresponding 'BD' *p*-value designation, indicate that a fold-change was not determined since the feature in WT (red), *ridA* (blue), or both (grey) groups was below the detection limit.

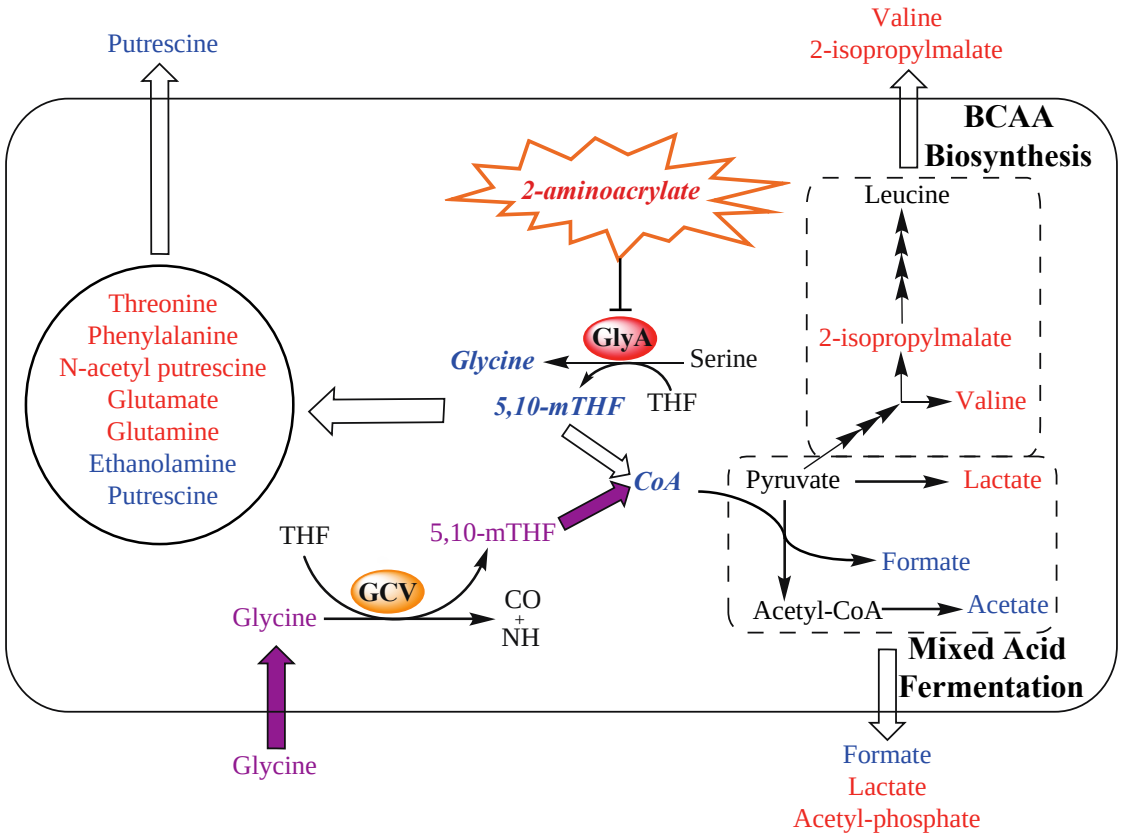


Figure 7.5. Metabolic outcomes stemming from 2-aminoacrylate-dependent damage of GlyA. 2AA-dependent damage of GlyA in a *ridA* mutant of *S. enterica* causes a glycine and 5,10-mTHF limitation. The 5,10-mTHF limitation leads to a downstream decrease in CoA, which in turn elicits a decrease in flux through CoA-dependent mixed acid fermentation. Accumulated pyruvate is rerouted toward lactate and branched chain amino acid (BCAA) biosynthesis. Addition of exogenous glycine increases the internal glycine content, which is utilized by the GCV complex to produce additional units of 5,10-mTHF, restoring wild-type levels of mixed acid fermentation and BCAA products. The levels of threonine, phenylalanine, N-acetyl putrescine, glutamate, glutamine, ethanolamine, and putrescine are believed to be influenced by GlyA damaged since glycine supplementation restores their concentration to wild-type levels, though the mechanism is unknown. Metabolites colored blue are elevated in wild-type and those colored red are elevated in the *ridA* mutant. Metabolites which are bolded and italicized were not detected by ^{1}H NMR. Compounds listed outside the rounded rectangle represent metabolites observed by ^{1}H NMR in the growth medium.

CHAPTER 8

CONCLUSIONS

8.1 DISSECTING METABOLIC NETWORK STRUCTURE USING 2-AMINOACRYLATE

The work in chapter 3 demonstrated that despite sharing all relevant metabolic components and regulatory machinery with *S. enterica*, an *E. coli ridA* mutant did not display the same growth defect observed for an *S. enterica ridA* mutant grown in minimal glucose medium. This discrepancy was traced back to an inability to accumulate 2AA in a *ridA* mutant of *E. coli* and introduction of additional units of the 2AA generator IlvA by means of an artificial construct was able to perturb growth. Up to a quarter of orthologous genes between *E. coli* and *S. enterica* have different expressional profiles [1], so it was possible that expressional differences in *ilvA* explained the 2AA generation discrepancy. However, crude activity assays of IlvE, which is encoded on an operon with *ilvA* [2], revealed that activity in wild-type *E. coli* was more than twice the level seen for wild-type *S. enterica*.

Chapter 4 demonstrated that elevated threonine content increases flux toward isoleucine biosynthesis, which allosterically inhibits IlvA, and therefore threonine plays an important indirect role in modulating 2AA generation. Previous work demonstrated that 2AC accumulates in wild-type *E. coli*, despite the presence of intact RidA [3], whereas 2AC is only able to accumulate in a *ridA* mutant of *S. enterica* [4]. Increased flux through threonine in *E. coli* may explain both the 2AC and 2AA phenotypic differences between *E. coli* and *S. enterica*. That is, increased flux through threonine may cause a buildup in 2AC, which can be repurposed for PRA synthesis, but

is also likely to increase the isoleucine pool in *E. coli* relative to *S. enterica*. Elevated isoleucine increases allosteric inhibition of IlvA-dependent 2AA production, preventing 2AA stress in these conditions.

Surprisingly, when 2AA stress is induced in *E. coli*, the failure for glycine to restore growth suggests that GlyA is not the primary growth-limiting target for 2AA damage. *E. coli* GlyA was still damaged when expressed in *S. enterica*, indicating that GlyA immunity to 2AA damage could not explain the discrepancy and again pointed to metabolic network structure differences as a likely explanation. Aside from cleavage of the hydroxymethyl group from serine by GlyA, *S. enterica* and *E. coli* both contain two alternate routes of glycine production from threonine [5-7]. One attractive possibility is that these routes for glycine production via threonine, possibly owing to the elevated threonine content in *E. coli* vs. *S. enterica*, provide sufficient glycine for an *E. coli* strain to circumvent the requirement for GlyA during 2AA stress. Interestingly, the transcriptomics work from chapter 5 showed that expression of the threonine dehydrogenase (Tdh; EC:1.1.1.103) / 2-amino-3-ketobutyrate CoA ligase (Kbl; EC: 2.3.1.29) route of glycine production from threonine is induced in a *ridA* mutant of *S. enterica*, though it is unclear what effect this has in contributing to glycine biosynthesis.

8.2 THE FAR-REACHING PHYSIOLOGICAL CONSEQUENCES OF 2-AMINOACRYLATE STRESS

The most striking finding from chapter 5 was that a large number of significant transcriptional changes between wild-type and a *ridA* mutant of *S. enterica* grown in minimal glucose medium were in genes relating to chemotaxis, flagellar assembly, and virulence (Salmonella pathogenicity island (SPI) 1, SPI 2, and SPI 5). The downregulation of flagellar biosynthesis machinery genes was predictive of a motility defect in a *ridA* mutant, and further

characterization of the motility defect revealed that it was contingent upon 2AA accumulation, but independent from GlyA damage (glycine supplementation could not restore motility). In general, the expression profile of chemotaxis, flagellar machinery, SPI 1, SPI 2, SPI 5, nitrate reduction (*nap* operon), and methionine biosynthesis (*met* genes) in a *ridA* mutant mimicked the expressional response of Salmonella, when growing inside of a macrophage [8]. This suggested that *S. enterica* experiencing metabolic crisis during 2AA stress may experience physiological signals similar to those experienced during growth inside a macrophage. This signal remains unknown, though recent work found that SPI 2, whose expression is induced in a *ridA* mutant, plays an important role in intracellular amino acid and carbon metabolism as well as in controlling an antioxidant defense [9]. It is possible that perturbations to amino acid and/or carbon metabolism leads to induction of SPI 2 gene expression, which may control downstream expression of tightly coordinated flagellar synthesis, SPI 1, chemotaxis, and *nap* genes [10-12].

Chapters 6 and 7 employed ^{1}H NMR metabolomics as a means to assess the global metabolic effect upon eliminating RidA (chapter 6). Since IlvA acts as the primary generator for 2AA under the growth conditions utilized, and since IlvA is allosterically inhibited by isoleucine, isoleucine supplementation was used to confirm that the metabolic effects seen were the result of 2AA stress and not some other effect of RidA elimination (chapter 7). Expanding on this approach, glycine addition supported the claim that the bulk of 2AA stress-mediated metabolic perturbations in a *ridA* mutant were linked to GlyA damage and glycine/5,10-mTHF limitation. The metabolic perturbations in a *ridA* mutant were widespread across amino acid metabolism, mixed acid fermentation, and putrescine metabolism. Viewing these data alongside the transcriptomics data from chapter 5 revealed additional points of perturbation in folate, purine, thiamine, and histidine metabolism. Altogether, this work helped demonstrate that since untargeted ^{1}H NMR can only

provide limited datasets consisting of tens of metabolites, that its true potential in effectively disentangling complex metabolic outcomes is realized when it is coupled with other -omics analyses and genetics-based approaches.

8.3 FUTURE DIRECTIONS

In the face of all the interesting work that has come from characterization of the RidA paradigm for 2AA stress, a number of important questions still remain. Many PLP-dependent enzymes exist in the cell, and an interesting dichotomy exists where both the generators and targets for 2AA-dependent damage are PLP-dependent enzymes. Little is known regarding the structural features which dictate damage susceptibility, though generators seem to belong to one fold-type of PLP-dependent enzymes, while targets belong to the other fold-types [13]. Continued use of -omics approaches may also prove useful in the identification of new targets for 2AA damage. Historically, 2AA targets have been identified genetically through their damage resulting in an observable phenotype. Transcriptomics and metabolomics offer a means to probe the perturbation of a metabolic network without the need for an observable growth defect.

To that end, continued dissection of the "nodes" of metabolism effected by 2AA stress in *S. enterica* is likely to prove fruitful. Glycine addition does not recover the expressional discrepancy relating to flagellar gene expression, nor does it recover the motility defect of a *ridA* mutant. However, ~400 genes were observed as differentially regulated in a *ridA* mutant grown in minimal glucose medium. Nutrient additions have the potential to inform which genes are regulated in a 2AA stress-dependent and a GlyA damage-dependent manner, much like the approach described in chapter 7. Other studies have suggested other functions for RidA, including a role as a translational inhibitor/ribonuclease [14-17] or a molecular chaperone [18-20], isoleucine addition could help reveal whether any of the transcriptional observations made in a *ridA* mutant

are linked to a 2AA-independent function of RidA or even related to an enamine/imine generated by an enzyme other than IlvA. Finally, damage to GlyA causes a depletion to both glycine and 5,10-mTHF levels. Glycine content influences the activation of the sRNA GcvB. Additionally, while 5,10-mTHF limitation is known to perturb CoA production, its limitation likely perturbs other areas of metabolism such as purine metabolism. Successive nutrient additions (pantothenate addition to restore CoA independent of 5,10-mTHF abundance, for example) and the careful use of mutant backgrounds (*gcvB* mutant, for example) would prove useful in dissecting whether the glycine status-dependent effect of 2AA is mediated through *gcvB* expression, CoA limitation, or some other effect of 5,10-mTHF limitation.

8.4 REFERENCES

1. Meysman, P. *et al.* (2013) Expression divergence between *Escherichia coli* and *Salmonella enterica* serovar Typhimurium reflects their lifestyles. Mol. Biol. Evol. 30 (6), 1302-1314.

2. Berg, C.M. and Shaw, K.J. (1981) Organization and regulation of the *ilvGEDA* operon in *Salmonella* Typhimurium LT2. J. Bacteriol. 145 (2), 984-9.

3. Bazurto, J.V. *et al.* (2016) An unexpected route to an essential cofactor: *Escherichia coli* relies on threonine for thiamine biosynthesis. mBio 7 (1), e01840-15.

4. Lambrecht, J.A. *et al.* (2010) Members of the YjgF/YER057c/UK114 family of proteins inhibit phosphoribosylamine synthesis *in vitro*. J. Biol. Chem. 285 (45), 34401-34407.

5. Fraser, J. and Newman, E.B. (1975) Derivation of glycine from threonine in Escherichia coli K-12 mutants. J. Bacteriol. 122 (3), 810-817.

6. Liu, J.-Q. *et al.* (1998) Gene cloning, biochemical characterization and physiological role of a thermostable low-specificity L-threonine aldolase from Escherichia coli. Eur. J. Biochem. 255 (1), 220-226.

7. Chan, T.T. and Newman, E.B. (1981) Threonine as a carbon source for Escherichia coli. J. Bacteriol. 145 (3), 1150-1153.

8. Srikumar, S. *et al.* (2015) RNA-seq brings new insights to the intra-macrophage transcriptome of *Salmonella* Typhimurium. PLoS Pathog. 11 (11), e1005262.

9. Noster, J. *et al.* (2019) Proteomics of intracellular Salmonella enterica reveals roles of Salmonella pathogenicity island 2 in metabolism and antioxidant defense. PLoS Pathog. 15 (4), e1007741.

10. Uchiya, K.-i. *et al.* (2009) Involvement of SPI-2-encoded SpiC in flagellum synthesis in Salmonella enterica serovar Typhimurium. BMC Microbiol. 9 (1), 179.

11. Chubiz, J.E.C. *et al.* (2010) FliZ regulates expression of the *Salmonella* pathogenicity island 1 invasion locus by controlling HilD protein activity in *Salmonella enterica* serovar Typhimurium. J. Bacteriol. 192 (23), 6261-6270.

12. Pruss, B.M. *et al.* (2001) FlhD/FlhC-regulated promoters analyzed by gene array and *lacZ* gene fusions. FEMS Microbiol. Lett. 197 (1), 91-97.

13. Borchert, A.J. *et al.* (2019) Reactive enamines and imines *in vivo*: Lessons from the RidA paradigm. Trends Biochem. Sci. 44 (10), 849-860.

14. Oka, T. *et al.* (1995) Isolation and characterization of a novel perchloric acid-soluble protein inhibiting cell-free protein synthesis. J. Biol. Chem. 270 (50), 30060-7.

15. Morishita, R. *et al.* (1999) Ribonuclease activity of rat liver perchloric acid-soluble protein, a potent inhibitor of protein synthesis. J. Biol. Chem. 274 (29), 20688-92.

16. Schmiedeknecht, G. *et al.* (1996) Isolation and characterization of a 14.5-kDa trichloroacetic-acid-soluble translational inhibitor protein from human monocytes that is upregulated upon cellular differentiation. Eur. J. Biochem. 242 (2), 339-51.

17. Walejko, J.M. *et al.* (2018) Gut microbiota and serum metabolite differences in African Americans and White Americans with high blood pressure. Int. J. Cardiol. 271, 336-339.

18. Samuel, S.J. *et al.* (1997) Hrp12, a novel heat-responsive, tissue-specific, phosphorylated protein isolated from mouse liver. Hepatology 25 (5), 1213-22.

19. Farkas, A. *et al.* (2004) DUK114, the *Drosophila* orthologue of bovine brain calpain activator protein, is a molecular chaperone. Biochem. J. 383 (Pt 1), 165-70.

20. Müller, A. *et al.* (2014) Activation of RidA chaperone function by N-chlorination. Nat. Commun. 5, 5804.

APPENDIX

SUPPLEMENTAL TABLES AND FIGURES

Table A.1. *S. enterica* LT2 *ridA* /wild-type transcriptome comparison

Locus tag	Fold Change	FDR value	Gene	Locus tag	Fold Change	FDR value	Gene
STM1264	1.76	2.32E-04	aadA	STM1173	-2.02	1.96E-02	flgA
STM4184	1.55	5.12E-03	aceA	STM1174	-5.70	6.47E-04	flgB
STM4183	2.21	1.40E-08	aceB	STM1175	-5.19	8.72E-05	flgC
STM0463	1.50	8.09E-03	amtB	STM1176	-4.98	1.02E-07	flgD
STM1446	1.41	4.25E-02	anmK	STM1177	-4.46	9.68E-09	flgE
STM1584	-1.47	1.75E-02	ansP	STM1178	-3.37	7.93E-07	flgF
STM2992	1.58	5.74E-04	argA	STM1179	-3.65	1.58E-08	flgG
STM4121	1.57	7.97E-04	argC	STM1180	-2.63	9.77E-07	flgH
STM4469	2.00	2.99E-08	argI	STM1181	-2.82	9.56E-08	flgI
STM0760	1.60	9.98E-04	aroG	STM1182	-2.00	8.61E-04	flgJ
STM0150	-1.66	4.46E-04	aroP	STM1183	-2.23	3.27E-08	flgK
STM0890	1.90	6.07E-07	artI	STM1184	-1.98	4.31E-05	flgL
STM0887	1.84	2.05E-03	artJ	STM1172	-2.05	9.75E-05	flgM
STM0888	1.92	7.93E-07	artM	STM1913	-1.75	2.48E-04	flhA
STM0891	2.17	1.46E-09	artP	STM1914	-1.85	1.74E-03	flhB
STM0889	1.94	2.63E-07	artQ	STM1924.S	1.58	8.19E-04	flhC
STM3876	-1.54	2.67E-02	asnC	STM1925	1.64	2.31E-04	flhD
STM1304	1.57	3.90E-02	astA	STM1912	-1.93	9.00E-03	flhE
STM1306	1.45	4.95E-02	astB	STM1956	-4.57	2.05E-07	fliA
STM3864	-1.40	3.60E-02	atpC	STM1959	-2.02	2.67E-07	fliC
STM3870	-1.44	2.78E-02	atpE	STM1960	-2.49	4.23E-06	fliD
STM3869	-1.42	3.02E-02	atpF	STM1968	-3.35	1.86E-06	fliE
STM3866	-1.39	3.86E-02	atpG	STM1969	-1.65	6.20E-03	fliF
STM0028	1.49	2.42E-02	bcfH	STM1970	-2.93	1.31E-11	fliG
STM0793	1.51	5.98E-03	bioA	STM1971	-3.47	2.73E-12	fliH
STM0794	1.70	4.49E-05	bioB	STM1972	-2.86	4.49E-05	fliI
STM0796	1.71	2.83E-02	bioC	STM1973	-2.84	3.62E-04	fliJ
STM0797	1.56	3.44E-03	bioD	STM1974	-2.21	2.19E-06	fliK
STM0795	1.78	9.00E-06	bioF	STM1975	-1.75	3.32E-03	fliL
STM3644	1.99	5.36E-08	bisC	STM1976	-2.20	3.53E-08	fliM
STM0068	-1.60	1.35E-02	caiF	STM1977	-2.07	2.15E-04	fliN
STM1921	-1.68	7.32E-04	cheA	STM1978	-1.96	4.88E-04	fliO
STM1917	-2.17	6.15E-06	cheB	STM1979	-1.77	1.04E-03	fliP
STM1919	-2.34	2.09E-07	cheM	STM1961	-2.41	9.65E-06	fliS
STM1918	-2.04	2.20E-04	cheR	STM1962	-1.81	2.05E-03	fliT
STM1916	-1.78	1.58E-03	cheY	STM1955	-2.45	1.54E-05	fliZ
STM1915	-1.48	3.12E-02	cheZ	STM2771	-5.70	1.02E-07	fljB
STM1142	1.55	3.18E-02	csgD	STM0974	-6.40	2.11E-02	focA
STM4398	1.43	3.12E-02	cycA	STM2193	3.21	1.81E-21	folE
STM2933	-1.39	3.40E-02	cysC	STM4084	-1.46	2.33E-02	fpr
STM2946	-1.40	2.43E-02	cysH	STM3446	-1.44	1.67E-02	fusA
STM1803	4.71	4.30E-35	dadA	STM2791	1.75	2.28E-03	gabD
STM1802	3.29	2.93E-20	dadX	STM2191	1.91	5.58E-04	galS
STM0064	-1.50	9.36E-03	dapB	STM1290	-1.61	1.87E-02	gapA
STM3967	1.48	1.31E-02	dlhH	STM3055	1.49	4.82E-03	gcvT
STM3629	1.39	4.43E-02	dppB	STM3551	1.50	1.15E-02	ggt
STM1885	-1.57	3.57E-02	edd	STM0462	1.57	3.07E-03	glnK
STM3070	-2.03	7.40E-03	epd	STM3525	1.54	1.90E-02	glpE
STM2391	-1.76	5.08E-05	fadL	STM4086	-1.45	2.25E-02	glpK

Table A.1 *continued*. *S. enterica* LT2 *ridA* /wild-type transcriptome comparison

Locus tag	Fold Change	FDR value	Gene	Locus tag	Fold Change	FDR value	Gene
STM4085	-1.45	2.31E-02	glpX	STM4182	2.52	7.46E-10	metA
STM3330	-1.38	4.56E-02	gltB	STM4100	3.13	5.36E-17	metB
STM3331	-1.51	4.87E-03	gltD	STM3161	1.48	6.82E-03	metC
STM0664	1.45	2.27E-02	gltJ	STM3965	2.16	2.74E-10	metE
STM4283	1.39	3.65E-02	gltP	STM4105	2.16	4.76E-07	metF
STM2555	1.37	4.44E-02	glyA	STM4188.S	2.23	1.41E-10	metH
STM3542	-1.49	4.33E-02	gntK	STM0246	1.38	2.25E-02	metI
STM3512	-1.88	3.51E-05	gntT	STM4099	1.38	3.62E-02	metJ
STM4330	-1.64	7.32E-04	groEL	STM4101	3.51	7.68E-26	metL
STM4329	-2.03	1.95E-04	groES	STM0247	1.66	9.98E-04	metN
STM2876	-1.72	9.84E-03	hilA	STM3964	2.03	1.94E-08	metR
STM2076	1.52	3.56E-03	hisA	STM1840	-1.47	1.29E-02	mgrB
STM2074	1.38	4.48E-02	hisB	STM3763	-1.77	3.26E-05	mgtB
STM2075	1.47	1.02E-02	hisH	STM1922	-1.63	1.04E-02	motB
STM2539	1.47	8.55E-03	hscA	STM0684	-1.48	2.33E-02	nagB
STM0659	-1.63	4.72E-02	hscC	STM0685	-1.54	8.67E-03	nagE
STM0487.S	-1.72	9.23E-03	htpG	STM2259	-4.68	4.96E-02	napA
STM3150	-2.93	7.07E-04	hypO	STM2260	-21.63	7.08E-03	napD
STM3909	-1.63	1.77E-02	ilvC	STM2261	-26.56	3.03E-04	napF
STM2896	-2.15	4.62E-06	invA	STM2258	-2.83	2.87E-02	napG
STM2895	-2.83	1.88E-05	invB	STM2257	-1.44	5.06E-02	napH
STM2894	-1.57	3.24E-02	invC	STM1761	-1.49	3.90E-02	narI
STM2897	-2.73	2.63E-07	invE	STM3776	1.57	5.98E-03	nepI
STM2899	-2.79	2.96E-07	invF	STM2543	2.05	7.64E-05	nifS
STM2898	-3.15	3.10E-12	invG	STM4452	-4.05	2.52E-02	nrdD
STM2893	-2.54	1.25E-03	invI	STM3494.S	1.66	8.01E-03	nudE
STM2541	1.84	2.85E-06	iscA	STM2325	-1.43	3.08E-02	nuoE
STM4106	2.12	2.90E-09	katG	STM3352	1.60	4.17E-02	oadA
STM3709	2.17	7.00E-09	kbl	STM3218	1.47	4.75E-02	oat
STM0086	1.76	7.17E-05	kefC	STM1572	-1.73	1.22E-03	ompD
STM0113	-1.55	2.98E-03	leuA	STM1732	-5.82	1.74E-03	ompW
STM0112	-2.10	1.65E-08	leuB	STM1092	-18.90	3.53E-03	orfX
STM0111	-2.14	2.14E-09	leuC	STM1824	1.45	2.23E-02	pabB
STM0110	-2.07	1.09E-08	leuD	STM1246	-1.61	4.37E-02	pagC
STM0115	1.67	4.70E-02	leuO	STM1244	-1.95	4.31E-05	pagD
STM0633	1.42	3.34E-02	lipA	STM1004	1.45	3.65E-02	pncB
STM3560	-2.66	5.72E-15	livF	STM1479	-1.75	4.88E-04	pntA
STM3561	-2.83	1.71E-16	livG	STM1480	-1.68	1.47E-03	pntB
STM3563	-2.72	5.72E-15	livH	STM0877	-1.47	2.27E-02	potF
STM3567	-6.99	7.46E-55	livJ	STM3910	-1.71	4.97E-05	ppiC
STM3564	-3.20	9.11E-15	livK	STM2874	-2.63	6.31E-11	prgH
STM3562	-2.46	2.84E-12	livM	STM2873	-2.51	2.63E-07	prgI
STM4227	1.62	2.31E-03	malG	STM2872	-2.09	1.59E-04	prgJ
STM3514	1.52	9.92E-03	malP	STM2871	-1.47	2.28E-02	prgK
STM1830	-1.78	6.88E-03	manX	STM4392	-1.41	2.41E-02	priB
STM1831	-1.55	4.29E-02	manY	STM0322	1.40	3.65E-02	proA
STM1832	-1.47	1.76E-02	manZ	STM0321	1.56	2.05E-03	proB
STM1149	1.52	2.38E-02	mdoC	STM4290	1.42	4.44E-02	proP
STM1425	2.04	2.63E-07	mdtK	STM2809	1.98	1.85E-06	proV

Table A.1 *continued*. *S. enterica* LT2 *ridA* /wild-type transcriptome comparison

Locus tag	Fold Change	FDR value	Gene	Locus tag	Fold Change	FDR value	Gene
STM2810	1.95	3.66E-06	proW	STM0510	1.57	1.75E-03	sfbA
STM2811	1.64	1.74E-03	proX	STM0511	1.84	6.25E-06	sfbB
STM1203	-2.26	7.44E-11	ptsG	STM0512	1.85	2.56E-06	sfbC
STM2487	1.44	1.29E-02	purC	STM1090	-2.56	2.20E-05	sigE
STM0534	1.73	1.87E-02	purE	STM2885	-1.83	3.60E-04	sipB
STM0533	1.54	1.67E-02	purK	STM2884	-1.98	2.33E-06	sipC
STM2499.S	1.69	2.19E-02	purM	STM2883	-1.55	4.61E-02	sipD
STM1883	1.62	3.59E-04	purT	STM2387	1.40	3.52E-02	sixA
STM4459	-1.81	4.21E-05	pyrI	STM2066	1.48	1.82E-02	sopA
STM2271	1.38	4.94E-02	rcsC	STM1091	-3.04	7.45E-12	sopB
STM3195	1.45	2.16E-02	ribB	STM2945	-1.81	3.77E-02	sopD
STM2675	-1.39	3.80E-02	rimM	STM2892	-1.86	1.35E-02	spaN
STM1066	2.66	4.80E-13	rmf	STM2891	-1.96	2.09E-03	spaO
STM4150	-1.46	1.29E-02	rplA	STM0165	-1.40	3.38E-02	speD
STM3437	-1.59	7.92E-04	rplB	STM0166	-1.39	4.50E-02	speE
STM3440	-1.50	6.10E-03	rplC	STM1593	-1.59	1.14E-02	srfA
STM3439	-1.41	2.28E-02	rplD	STM1594	-1.80	5.49E-05	srfB
STM3425	-1.52	4.01E-03	rplF	STM1406	5.65	4.19E-02	ssaG
STM4151	-1.73	5.24E-05	rplJ	STM1409	2.79	2.31E-04	ssaJ
STM4149	-1.38	4.59E-02	rplK	STM1411	2.30	4.04E-03	ssaK
STM4152	-1.60	8.71E-04	rplL	STM1412	1.74	1.54E-02	ssaL
STM3345	-1.42	3.08E-02	rplM	STM1413	2.90	3.24E-02	ssaM
STM3421	-1.43	1.95E-02	rplO	STM1419	1.78	3.94E-02	ssaR
STM3433	-1.44	1.54E-02	rplP	STM1414	2.07	1.08E-04	ssaV
STM3414	-1.48	1.02E-02	rplQ	STM1398	1.58	3.08E-02	sseB
STM3424	-1.43	1.87E-02	rplR	STM1400	1.75	2.42E-03	sseC
STM2673	-1.53	3.35E-03	rplS	STM0108	1.41	1.76E-02	tbpA
STM3435	-1.53	3.44E-03	rplV	STM0690	1.78	2.96E-03	tcuB
STM3438	-1.56	1.76E-03	rplW	STM3708	2.24	4.03E-10	tdh
STM3432	-1.45	1.66E-02	rpmC	STM4164	1.84	2.17E-06	thiC
STM3422	-1.42	2.24E-02	rpmD	STM2146	1.42	1.75E-02	thiD
STM3415	-1.54	3.32E-03	rpoA	STM4163	1.71	5.24E-05	thiE
STM2924	1.75	2.19E-03	rpoS	STM4162	1.88	2.02E-06	thiF
STM3434	-1.44	1.77E-02	rpsC	STM4160	1.66	1.76E-04	thiG
STM3416	-1.43	2.42E-02	rpsD	STM4159	1.58	1.01E-03	thiH
STM3423	-1.43	1.77E-02	rpsE	STM2147	1.55	2.18E-03	thiM
STM3426	-1.54	2.90E-03	rpsH	STM0745	1.37	5.03E-02	tolQ
STM3441	-1.51	4.73E-03	rpsJ	STM2674	-1.64	3.73E-04	trmD
STM3418	-1.49	9.84E-03	rpsM	STM4583	-1.44	4.35E-02	trpR
STM3427.S	-1.41	3.34E-02	rpsN	STM3402	1.42	3.66E-02	tsaC
STM3431	-1.56	2.98E-03	rpsQ	STM0217	-1.41	2.75E-02	tsf
STM4393	-1.53	3.56E-03	rpsR	STM4533	-1.44	4.44E-02	tsr
STM3436	-1.50	5.41E-03	rpsS	STM3445	-1.42	2.77E-02	tuf
STM4165	1.74	1.68E-04	rsd	STM1937	2.16	1.05E-08	tyrP
STM1471	1.40	3.65E-02	rstB	STM0085	2.02	8.37E-06	yabF
STM4063	-1.77	1.12E-04	sbp	STM0171	1.51	5.22E-03	yadF
STM3420	-1.40	3.45E-02	secY	STM0204.S	1.74	7.64E-05	yadR
STM3062	1.87	9.07E-07	serA	STM0601	1.40	3.47E-02	ybdD
STM0977	1.36	4.34E-02	serC	STM0602	1.70	8.86E-05	ybdH

Table A.1 *continued*. *S. enterica* LT2 *ridA* /wild-type transcriptome comparison

Locus tag	Fold Change	FDR value	Gene	Locus tag	Fold Change	FDR value	Gene
STM0603	2.25	9.52E-12	ybdL	STM4552	-2.78	6.99E-03	-
STM0604	1.91	1.15E-06	ybdM	STM4510	1.57	9.55E-03	-
STM0713	1.39	3.68E-02	ybgK	STM4509.S	1.44	4.70E-02	-
STM0758	-1.37	4.87E-02	ybgR	STM4457	-1.64	8.02E-04	-
STM0938	1.43	2.77E-02	ybjE	STM4390	-1.56	3.65E-02	-
STM0968	1.88	2.08E-02	ycaD	STM4315	-3.95	1.38E-03	-
STM0997	1.38	3.77E-02	ycbL	STM4272	1.50	2.75E-02	-
STM1217.S	1.38	4.71E-02	ycfU	STM4258	-2.13	1.71E-02	-
STM1625	2.23	1.74E-02	ydcI	STM4257	-3.42	2.25E-02	-
STM1522	1.55	3.56E-03	ydeA	STM4161	1.56	2.07E-03	-
STM1517	1.45	2.32E-02	ydeD	STM4102	3.63	2.83E-23	-
STM1476	1.73	3.81E-02	ydgC	STM4070	1.73	1.86E-02	-
STM1477	1.62	3.57E-03	ydgI	STM4069	1.73	2.19E-03	-
STM1344	1.67	5.93E-03	ydiV	STM3980	-1.78	5.03E-02	-
STM1965	-1.83	1.68E-04	yedE	STM3966	2.27	6.78E-10	-
STM1993	1.49	2.37E-02	yedJ	STM3956	1.72	1.51E-03	-
STM4499	-1.55	7.37E-03	yeeN	STM3650	-4.07	2.45E-02	-
STM2158	-1.48	1.96E-02	yehT	STM3566	-6.74	6.30E-23	-
STM2279	-1.52	1.08E-02	yfaE	STM3516	-1.45	1.99E-02	-
STM2308	-1.62	2.19E-03	yfbB	STM3411	-1.58	9.85E-03	-
STM2347	1.39	3.34E-02	yfcE	STM3160	-4.65	1.62E-02	-
STM2348	1.73	5.24E-05	yfcF	STM3138	-2.14	5.49E-05	-
STM2413	-1.56	2.20E-02	yfeC	STM3085	1.81	1.66E-02	-
STM2544	2.82	3.03E-14	yfhP	STM3050	-1.42	4.90E-02	-
STM2794	1.53	3.57E-02	ygaE	STM2986.Sc	-1.98	1.09E-02	-
STM2790	1.81	2.05E-03	ygaF	STM2628	-1.95	2.33E-06	-
STM3060	-1.44	2.63E-02	ygfE	STM2610	-1.94	7.80E-03	-
STM3225	1.39	2.87E-02	ygjU	STM2542	1.75	1.68E-04	-
STM3274	-9.71	3.35E-03	yhbU	STM2537	1.40	3.48E-02	-
STM3473	1.44	3.20E-02	yhfC	STM2505	-1.72	5.55E-04	-
STM3511	1.67	2.22E-02	yhgI	STM2361	-1.48	7.78E-03	-
STM3565	-3.01	1.02E-14	yhhK	STM2281	-1.59	5.03E-02	-
STM3611	-1.96	2.58E-03	yhjH	STM2192	2.64	1.45E-14	-
STM3930	2.33	6.31E-11	yifK	STM2179	1.88	2.11E-02	-
STM4181	-2.00	2.25E-05	yjaB	STM1966	-1.93	1.44E-02	-
STM4226	1.90	2.53E-06	yjbA	STM1872	2.11	5.20E-04	-
STM4264	1.46	3.08E-02	yjcC	STM1868A	-1.90	2.11E-04	-
STM4458	-21.96	1.22E-34	yjgF	STM1857	1.68	4.25E-02	-
STM4532	1.69	8.14E-04	yjiY	STM1795	2.15	1.58E-08	-
STM4566	-4.98	4.59E-02	yjjI	STM1731	-1.72	2.30E-02	-
STM4599	-1.43	3.95E-02	yjjY	STM1722a	2.04	1.41E-04	-
STM0471	1.48	1.66E-02	ylaC	STM1657	1.57	1.08E-02	-
STM1652	-1.68	4.25E-02	ynaF	STM1633	-1.46	3.71E-02	-
STM3164	-1.45	2.45E-02	yqhD	STM1555	1.43	3.48E-02	-
STM3666	-1.51	4.40E-03	ysaA	STM1528	1.71	8.41E-03	-
STM4402	-1.64	3.52E-02	ytfH	STM1527	2.25	1.45E-04	-
STM1059	1.79	8.86E-05	zapC	STM1317	1.49	4.04E-02	-
STM1891	-1.36	4.59E-02	znuA	STM1266	1.61	2.02E-02	-
STM4241	-1.39	4.96E-02	zur	STM1263	-1.52	1.87E-02	-

Table A.1 *continued*. *S. enterica* LT2 *ridA* /wild-type transcriptome comparison

Locus tag	Fold Change	FDR value	Gene	Locus tag	Fold Change	FDR value	Gene
STM1156	1.74	3.44E-03	-	STM05910	3.06	5.89E-03	-
STM1126	1.61	2.90E-03	-	STM05680	-5.00	4.08E-02	-
STM1012	-1.95	2.26E-06	-	STM05520	-1.75	9.85E-03	-
STM0898	-1.68	3.56E-03	-	STM05320	-2.98	2.49E-03	-
STM0840	1.55	3.46E-02	-	STM05010	1.66	2.20E-02	-
STM0771	-1.49	3.09E-02	-	STM0402	-1.74	5.98E-03	-
STM0714	1.50	1.08E-02	-	STM0382	2.01	1.39E-06	-
STM0710	1.52	3.01E-03	-	STM0009	-1.95	4.04E-06	-
STM0692	1.58	8.28E-03	-	STM3895	-1.49	4.82E-02	-
STM0691	1.78	2.52E-04	-	STM2966	1.44	1.08E-02	-
STM0660	-1.42	4.25E-02	-				

Table A.2. Bacterial strains and plasmids from Chapter 5

Strain	Genotype[a]	Source
DM9404	Wild-type (isogenic to DM3480)	Lab Collection
DM3480	*ridA3* ::MudJ[b]	Lab Collection
DM6946	*ilvA219 ridA3* ::MudJ	Lab Collection
DM6947	*ilvA219*	Lab Collection
DM15035	ABcc1[c]	Lab Collection
DM15036	*ridA3* ::MudJ ABcc1	Lab Collection
DM15418	WT / pCV1	This study
DM15419	*ridA3* ::MudJ / pCV1	This study
DM15420	WT / pDM1439	This study
DM15421	*ridA3* ::MudJ/ pDM1439	This study
DM15319	*ydiV* ::*cat*[d]	This study
DM15320	*ridA3* ::MudJ *ydiV::cat*	This study
DM15340	*rpoS1117::amp*	This study
DM15341	*ridA3* ::MudJ *rpoS1117::amp*	This study
DM15505	*fliZ::cat*	This study
DM15506	*ridA3* ::MudJ *fliZ::cat*	This study
DM15507	*omrBA::cat*	This study
DM15508	*ridA3* ::MudJ *omrBA::cat*	This study
DM15509	*micA::cat*	This study
DM15510	*ridA3* ::MudJ *micA::cat*	This study
DM15511	*oxyS::cat*	This study
DM15512	*ridA3* ::MudJ *oxyS::cat*	This study
DM15513	*arcZ::cat*	This study
DM15514	*ridA3* ::MudJ *arcZ::cat*	This study
DM15817	*flhDC::cat*	This study
DM15818	*ridA3* ::MudJ *flhDC::cat*	This study
Plasmid	**Description**	**Source**
pCV1	BspQI modified pBAD24	Lab collection
pDM1439	pCV1-*ridA*[e]	Lab collection

[a]All strains were derived from *Salmonella enterica* serovar Typhimurium LT2

[b]MudJ refers to the MudJ1734 transposon [Castilho, Olfson, and Casadaban,1984].

[c]ABcc1 is a chromosomal cassette (araBAD::PBAD-ilvA219 cat) that has the ilvA219 allele under the inducible control of the PBAD promoter and encodes chloramphenicol resistance [Borchert, 2017].

[d]*cat* alleles are insertion deletions of the relevant loci as described in Chapter 5.3

[e]Plasmid constructed using pCV1 [Vandrisse, 2016].

Table A.3. Primers used in Chapter 5.

Primer Name	Primer Sequence 5' -> 3'
Strain construction	
LT2 Wanner ydiV 5'	ACTGGATGGCGAATAGCGCCCTAACCATGG GACTGGCGTAGTGTAGGCTGGAGCTGCTTC
LT2 Wanner ydiV 3'	GCAATTCATCGCGCCAACGGGCAGTAAAAG ACAGGGTCATCATATGAATATCCTCCTTAG
LT2 Wanner fliZ 5'	CGAAAAGTGCCGCACAACGTATAGACTACC AGGAGTTCTCGTGTAGGCTGGAGCTGCTTC
LT2 Wanner fliZ 3'	CACGTTTCACCAACACGACTCTGCTACATCT TATGCTTTTCATATGAATATCCTCCTTAG
LT2 Wanner omrBA 5'	GTGACCTGCACGGGAGAAATGCGTCTGTAA AGTGACAATAGTGTAGGCTGGAGCTGCTTC
LT2 Wanner omrBA 3'	TAACAGGAGCGATAGCAAAATAGGTTCGAAA AAAAACCTGCATATGAATATCCTCCTTAG
LT2 Wanner micA 5'	ATAAACTGAACTCTTTGTTCCGGGGCGAGTC TGAGTATATGTGTAGGCTGGAGCTGCTTC
LT2 Wanner micA 3'	CGAGCCGTTTGCCGCGTGGCTTGCAAAACA CGCCTGACCCCATATGAATATCCTCCTTAG
LT2 Wanner oxyS 5'	CGATTATCCCTATCAACCTTTCTGATTAATAA TACATCACGTGTAGGCTGGAGCTGCTTC
LT2 Wanner oxyS 3'	CCGAGCATCAGCGACGCCGGGTTTTTTTAC GCAAAAAAAACATATGAATATCCTCCTTAG
LT2 Wanner arcZ 5'	TTCATGTAACAAATCATTTAGGATTTGCTATC TTAACTGCGTGTAGGCTGGAGCTGCTTC
LT2 Wanner arcZ 3'	GGCAAACGCGGAAAAAAAATGACCCCGGTT TGACCGGGGTCATATGAATATCCTCCTTAG
LT2 Wanner flhDC 5'	GTGCGGCTACGTCGCACAAAAATAAAGTTGG TTATTCTGGGTGTAGGCTGGAGCTGCTTC
LT2 Wanner flhDC 3'	TCTGTTCATCCAGCAGTTGTGGAATAATATC GGCAGCATCCATATGAATATCCTCCTTAG
Cloning	
LT2 ridA cloned into pCV1 5'	NNGCTCTTCNTTCATGAGCAAAACTATCGC
LT2 ridA cloned into pCV1 3'	NNGCTCTTCNTTATTAGCGACGAACAGC

Table A.3 *continued*. Primers used in Chapter 5.

Primer Name	Primer Sequence 5' -> 3'
Sequence verification primers	
LT2 check ydiV 5'	GGAACCGGTTCACCGGTAAA
LT2 check ydiV 3'	AACGGTGTGGGTAAAAGCGC
LT2 check fliZ 5'	TTGGCGCGGTACTGGAA
LT2 check fliZ 3'	TCATAGTAACCCCGAATATTGCC
LT2 check omrBA 5'	CGCTTACATTACGCCAGT
LT2 check omrBA 3'	TTCAACTCCCTTTGGCC
LT2 check micA 5'	TTAGTCACCTCCGATAAT
LT2 check micA 3'	GTTTATTGCTGAACGTGA
LT2 check oxyS 5'	TATTCATCCTCCGTCGC
LT2 check oxyS 3'	GCCTTATAAGCATAGCGC
LT2 check arcZ 5'	GGTGAACATCCTCCTGC
LT2 check arcZ 3'	CAGCAAATCCAGTCGC
LT2 check flhDC 5'	CTAAAGGTAAAATAAAAGCG
LT2 check flhDC 3'	CATCCTTCCGCTGTTGAC
pBAD sequencing 5'	TACCTGACGCTTTTTATCGC
pBAD sequencing 3'	GAAAATCTTCTCTCATCCGC
RT-qPCR primers	
LT2 dadX qRT-PCR 5'	CGGGTGAAAGGGTAGGCTAT
LT2 dadX qRT-PCR 3'	GTGGATAGCCATCGGCATAA
LT2 fliI qRT-PCR 5'	GACCAACAAGATCCCATTGC
LT2 fliI qRT-PCR 3'	CGATATCAATGGCCGGATAG
LT2 metH qRT-PCR 5'	ACATTGCCCGTGTACCAATC
LT2 metH qRT-PCR 3'	GACGATGCCTTTACCCTGAA
LT2 napF qRT-PCR 5'	GCCAGGACAGTTGTGAACCT
LT2 napF qRT-PCR 3'	TACAGGCCTGCGAGTCAAG
LT2 thiF qRT-PCR 5'	GCCAGCTAATGGTCCTCACA
LT2 thiF qRT-PCR 3'	GTACGGCAGTTGCGTTCAG
LT2 sdaC qRT-PCR 5'	AAGGCTTTAACGGGATGGTG
LT2 sdaC qRT-PCR 3'	GTGGTGACCAGCATGAACAG
LT2 fliC qRT-PCR 5'	GCAAGTAAAGCCGAAGGTCA
LT2 fliC qRT-PCR 3'	GTCAACCTGTGCCAAAGCAG
LT2 cheM qRT-PCR 5'	GTAACGCAGGTTCGTGAAGG
LT2 cheM qRT-PCR 3'	TGTTCGGTACGGGAAGAGAG
LT2 invA qRT-PCR 5'	GTCCTCCGCCCTGTCTACTT
LT2 invA qRT-PCR 3'	CGGCACTAATCGCAATCAAC
LT2 ssaV qRT-PCR 5'	TCGGCACCTTAATTGACTGG
LT2 ssaV qRT-PCR 3'	CGGATTAAGACGACGCAGAA
LT2 sopB qRT-PCR 5'	CGGGCATCACTATACCAACAC
LT2 sopB qRT-PCR 3'	CCCTCATAAGCACTGGGAAA

Table A.4. *S. enterica* PLP-dependent and their corresponding pathway

Metabolic Pathway	PLP-Dependent Enzyme [EC Number]			
Amino Acid Metabolism				
Alanine	**Alr** [5.1.1.1]	**DadX** [5.1.1.1]	**AlaA** [2.6.1.2]	**AlaC** [2.6.1.2]
	AvtA [2.6.1.66]	**IscS** [2.8.1.7]	**SufS** [2.8.1.7]	**CsdA** [4.4.1.16]
Arginine	**AdiA [**4.1.1.19]	**ArgD** [2.6.1.81]	**SpeA** [4.1.1.19]	
Aspartate	**AspC** [2.6.1.1]	**TyrB** [2.6.1.57]		
BCAA	**IlvE** [2.6.1.42]	**AvtA** [2.6.1.66]		
Cysteine	**CsdA** [4.4.1.16]	**CysK** [2.5.1.47]	**CysM** [2.5.1.47]	**CdsH** [2.5.1.47]
	DcyD [4.4.1.15]	**MetB** [2.5.1.48]	**IscS** [2.8.1.7]	**SufS** [2.8.1.7]
DAP	**YgeX** [4.3.1.15]			
Glycine	**GlyA** [2.1.2.1]	**GcvP** [1.4.4.2]	**Kbl** [2.3.1.29]	**LtaA** [4.1.2.48]
Lysine	**SerC** [2.6.1.17]	**ArgD** [2.6.1.17]	**CadA** [4.1.1.18]	**EpmB** [5.4.3.2]
	LysA [4.1.1.20]	**STM2360** [-]		
Methionine	**MetB** [2.5.1.48]	**MetC** [4.4.1.13]	**YbdL** [2.6.1.88]	
Ornithine	**ArgD** [2.6.1.11]	**PatA** [2.6.1.13]	**SpeC** [4.1.1.17]	**SpeF** [4.1.1.17]
Phenylalanine	**AspC** [2.6.1.57]	**IlvE** [2.6.1.57]	**TyrB** [2.6.1.57]	
Proline	**PatA** [2.6.1.13]			
Selenocysteine	**SelA** [2.9.1.1]			
Serine	**SerC** [2.6.1.52]	**IlvA** [4.3.1.17]	**TdcB** [4.3.1.17]	**DsdA** [4.3.1.18]
	GlyA [2.1.2.1]			
Threonine	**IlvA** [4.3.1.19]	**TdcB** [4.3.1.17	**Kbl** [2.3.1.29	**LtaA** [4.1.2.48]
	ThrC [4.2.3.1]			
Tryptophan	**TrpB** [4.2.1.20]			
Tyrosine	**TyrB** [2.6.1.5]	**AspC** [2.6.1.5]		
Putrescine	**SpeA** [4.1.1.19]	**SpeC** [4.1.1.17]	**SpeF** [4.1.1.17]	**PatA** [2.6.1.82]
	GabT [2.6.1.19]			
Cofactor/Coenzyme Biosynthesis				
Biotin	**BioA** [2.6.1.62]	**BioF** [2.3.1.47]		
PLP	**SerC** [2.6.1.52]			
Tetrapyrrole	**HemL** [5.4.3.8]			
[2Fe-2S]	**IscS** [2.8.1.7	**SufS** [2.8.1.7]		
Folate	**GlyA** [2.1.2.1]	**GcvP** [1.4.4.2	**PabC** [4.1.3.38]	
Cell Structure				
Glycogen	**GlgP** [2.4.1.1	**MalP** [2.4.1.1]		
LPS	**ArnB** [2.6.1.87]	**WecE** [2.6.1.59]		
D-alanine	**Alr** [5.1.1.1]	**DadX** [5.1.1.1]		

Table A.5. Spectral feature used for quantification and metabolite confidence score

Compound	Binned_chemical_shift	Confidence_level
2-isopropylmalic acid	0.844	4
Pantothenic acid	0.8847	3
Isoleucine	0.928	4
Leucine	0.952	4
2-Aminobutyric acid	0.97	4
Valine	0.98	4
Threonine	1.1378	4
N-acetyl alanine	1.3258	4
Alanine	1.47	4
N-Acetylputrescine	1.576	4
Putrescine	1.765	4
Acetate	1.906	3
Glutamate	2.327	4
Succinic acid	2.392	3
Oxoglutaric acid	2.4288	4
Methionine	2.63	4
Malic acid	2.659	4
Lysine	3.015	4
Ethanolamine	3.1297	4
GSSG	3.3	4
Serine	4.40335	4
CTP	4.418	4
Glucose	5.213	3
Maltose	5.407	4
Uracil	5.79	4
Uridine 5'-triphosphate (UTP)	5.968	4
CMP	5.996	3
dCMP	6.13	4
Tyrosine	6.887	4
Phenylalanine	7.4158	4
Cytidine diphosphate (CDP)	7.968	4
UMP	8.108	4
Nicotinate Mononucleotide	8.265	5
Formate	8.447	3
Adenosine monophosphate (AMP)	8.602	3
Nicotinamide Adenine Dinucleotide Phosphate (NADP)	9.095	4
Nicotinamide Adenine Dinucleotide (NAD)	9.1374	4

Table A.6. KEGG pathways represented by up-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm00730:Thiamine metabolism	8	STM4159, STM4162, STM4163, STM4164, STM4160, STM2146, STM2543, STM2147	11.79	8.52E-04
stm01230:Biosynthesis of amino acids	18	STM3965, STM0322, STM0321, STM2074, STM4188.S, STM4121, STM4469, STM0760, STM3161, STM2992, STM4101, STM2075, STM2076, STM4100, STM4182, STM3062, STM0977, STM2555	2.45	0.541526
stm00780:Biotin metabolism	6	STM0797, STM3644, STM0796, STM0794, STM0795, STM0793	7.58	0.64803
stm01100:Metabolic pathways	56	STM3965, STM2074, STM3218, STM3967, STM0533, STM1004, STM0534, STM2075, STM2076, STM0633, STM4162, STM1304, STM4163, STM4164, STM1306, STM4160, STM3055, STM0977, STM2543, STM2487, STM0322, STM0321, STM4121, STM3514, STM0760, STM3161, STM2791, STM1802, STM2499.S, STM3551, STM3195, STM3062, STM2555, STM1883, STM3352, STM4159, STM4188.S, STM4469, STM1795, STM2193, STM0794, STM0795, STM0793, STM0797, STM4182, STM0796, STM3644, STM4183, STM4184, STM4105, STM2146, STM4106, STM2147, STM4101, STM2992, STM4100	1.38	0.952251
stm00670:One carbon pool by folate	5	STM4188.S, STM4105, STM3055, STM2555, STM1883	6.80	4.299035
stm01110:Biosynthesis of secondary metabolites	28	STM3965, STM2074, STM4188.S, STM3967, STM0533, STM4469, STM0534, STM2075, STM2076, STM4182, STM4183, STM4184, STM3055, STM2487, STM0322, STM4121, STM3514, STM0760, STM4106, STM3161, STM2992, STM4101, STM4100, STM2499.S, STM3195, STM2555, STM1883, STM4070	1.64	5.386926
stm00260:Glycine, serine and threonine metabolism	7	STM3709, STM3062, STM3708, STM0977, STM3055, STM2555, STM4101	3.54	10.61463

Table A.6 *cont*. KEGG pathways represented by up-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm02010:ABC transporters	18	STM0247, STM0246, STM0891, STM0890, STM0664, STM0108, STM4227, STM2809, STM0512, STM3629, STM0889, STM0888, STM1217.S, STM0510, STM0511, STM2811, STM0887, STM2810	1.84	11.68299
stm00270:Cysteine and methionine metabolism	6	STM3965, STM4188.S, STM4182, STM3161, STM4101, STM4100	3.66	18.09814
stm00330:Arginine and proline metabolism	5	STM0322, STM3218, STM0321, STM1304, STM1306	4.02	26.61548
stm00450:Selenocompound metabolism	4	STM3965, STM4188.S, STM3161, STM4100	4.72	37.97152
stm01130:Biosynthesis of antibiotics	18	STM0322, STM0321, STM3967, STM4121, STM0533, STM4469, STM0534, STM0760, STM2992, STM4101, STM4100, STM2499.S, STM3062, STM0977, STM2555, STM3055, STM1883, STM2487	1.58	39.73111
stm00220:Arginine biosynthesis	4	STM4121, STM4469, STM1795, STM2992	3.93	53.72737
stm00332:Carbapenem biosynthesis	2	STM0322, STM0321	17.69	67.91703
stm00340:Histidine metabolism	3	STM2074, STM2075, STM2076	4.42	77.93238
stm04122:Sulfur relay system	3	STM4161, STM4162, STM2543	3.12	93.78712
stm00630:Glyoxylate and dicarboxylate metabolism	4	STM4183, STM4184, STM3055, STM2555	2.14	96.08187
stm00360:Phenylalanine metabolism	2	STM1803, STM4106	5.90	96.71159
stm00460:Cyanoamino acid metabolism	2	STM3551, STM2555	5.05	98.141
stm00230:Purine metabolism	6	STM0533, STM2499.S, STM0534, STM3494.S, STM1883, STM2487	1.36	99.70112
stm01200:Carbon metabolism	8	STM4183, STM4184, STM3062, STM4105, STM1795, STM0977, STM3055, STM2555	1.24	99.76055

Table A.6 *cont*. KEGG pathways represented by up-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm03070:Bacterial secretion system	4	STM1406, STM1414, STM1409, STM1419	1.54	99.84508
stm00250:Alanine, aspartate and glutamate metabolism	3	STM1795, STM4510, STM2791	1.77	99.90867
stm00680:Methane metabolism	3	STM3062, STM0977, STM2555	1.77	99.90867
stm00790:Folate biosynthesis	2	STM2193, STM1824	2.21	99.98921
stm00920:Sulfur metabolism	3	STM4182, STM3525, STM4100	1.40	99.99564
stm00910:Nitrogen metabolism	2	STM0171, STM1795	1.61	99.99966
stm05132:Salmonella infection	2	STM1398, STM1400	1.36	99.99997
stm01210:2-Oxocarboxylic acid metabolism	2	STM4121, STM2992	1.22	99.99999
stm02040:Flagellar assembly	2	STM1924.S, STM1925	0.96	100
stm01120:Microbial metabolism in diverse environments	10	STM3967, STM3525, STM4183, STM4184, STM3062, STM4105, STM0977, STM2555, STM4101, STM2791	0.73	100
stm00620:Pyruvate metabolism	2	STM4183, STM3352	0.63	100
stm02020:Two-component system	5	STM1471, STM1924.S, STM2271, STM1925, STM0664	0.53	100
stm00500:Starch and sucrose metabolism	1	STM3514	0.57	100
stm01054:Nonribosomal peptide structures	1	STM4510	17.69	100
stm00471:D-Glutamine and D-glutamate metabolism	1	STM1795	3.54	100
stm00300:Lysine biosynthesis	1	STM4101	1.18	100

Table A.6 *cont*. KEGG pathways represented by up-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm00785:Lipoic acid metabolism	1	STM0633	5.90	100
stm00310:Lysine degradation	1	STM2791	1.97	100
stm00623:Toluene degradation	1	STM3967	17.69	100
stm00650:Butanoate metabolism	1	STM2791	0.63	100
stm00400:Phenylalanine, tyrosine and tryptophan biosynthesis	1	STM0760	0.77	100
stm00760:Nicotinate and nicotinamide metabolism	1	STM1004	0.98	100
stm00380:Tryptophan metabolism	1	STM4106	2.53	100
stm00473:D-Alanine metabolism	1	STM1802	4.42	100
stm00261:Monobactam biosynthesis	1	STM4101	1.97	100
stm00750:Vitamin B6 metabolism	1	STM0977	1.77	100
stm00350:Tyrosine metabolism	1	STM2791	0.88	100
stm00480:Glutathione metabolism	1	STM3551	0.93	100
stm00361:Chlorocyclohexane and chlorobenzene degradation	1	STM3967	8.85	100
stm00564:Glycerophospholipid metabolism	1	STM4070	0.61	100
stm00364:Fluorobenzoate degradation	1	STM3967	17.69	100
stm00740:Riboflavin metabolism	1	STM3195	2.53	100

Table A.6 *cont*. KEGG pathways represented by up-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm01502:Vancomycin resistance	1	STM1802	2.21	100
stm00430:Taurine and hypotaurine metabolism	1	STM3551	4.42	100
stm04931:Insulin resistance	1	STM3514	5.90	100

Table A.7. KEGG pathways represented by down-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm02040:Flagellar assembly	31	STM1172, STM1179, STM1976, STM1973, STM1177, STM1974, STM1178, STM1979, STM1175, STM1176, STM1977, STM1173, STM1978, STM1174, STM1972, STM1971, STM1970, STM1913, STM1914, STM1959, STM2771, STM1180, STM1181, STM1183, STM1960, STM1961, STM1184, STM1969, STM1968, STM1962, STM1922	10.40	2.87E-25
stm03010:Ribosome	28	STM3345, STM3421, STM3422, STM3423, STM3424, STM4151, STM4150, STM4152, STM3440, STM3441, STM3427.S, STM3426, STM3425, STM4149, STM2673, STM3435, STM3434, STM3433, STM3432, STM3431, STM3438, STM3439, STM3436, STM3437, STM3416, STM3414, STM4393, STM3418	4.46	3.98E-09
stm02030:Bacterial chemotaxis	11	STM1919, STM4533, STM1915, STM1916, STM1917, STM1970, STM1918, STM1976, STM1922, STM1921, STM1977	5.94	0.0034175
stm05100:Bacterial invasion of epithelial cells	5	STM2885, STM2884, STM2883, STM1091, STM2945	6.89	3.3749837
stm05132:Salmonella infection	6	STM2885, STM2884, STM2883, STM1091, STM1959, STM2771	2.86	39.240525
stm03070:Bacterial secretion system	8	STM2894, STM2896, STM3420, STM2871, STM2873, STM2891, STM2897, STM2898	2.16	49.324575
stm00290:Valine, leucine and isoleucine biosynthesis	5	STM0110, STM0111, STM0112, STM0113, STM3909	3.10	49.653171
stm01210:2-Oxocarboxylic acid metabolism	5	STM0110, STM0111, STM0112, STM0113, STM3909	2.14	87.696258
stm00910:Nitrogen metabolism	4	STM1761, STM2259, STM3331, STM3330	2.26	93.979749

Table A.7 *cont*. KEGG pathways represented by down-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm00660:C5-Branched dibasic acid metabolism	3	STM0110, STM0111, STM0112	2.33	98.790477
stm00190:Oxidative phosphorylation	5	STM2325, STM3864, STM3870, STM3869, STM3866	1.48	99.58207
stm00520:Amino sugar and nucleotide sugar metabolism	6	STM0685, STM0684, STM1203, STM1831, STM1830, STM1832	1.31	99.82235
stm02060:Phosphotransferase system (PTS)	5	STM0685, STM1203, STM1831, STM1830, STM1832	1.07	99.998844
stm00250:Alanine, aspartate and glutamate metabolism	3	STM3331, STM3330, STM4459	1.24	99.999177
stm00051:Fructose and mannose metabolism	4	STM4085, STM1831, STM1830, STM1832	1.03	99.999844
stm00760:Nicotinate and nicotinamide metabolism	2	STM1479, STM1480	1.38	99.999944
stm02020:Two-component system	12	STM1919, STM1761, STM1956, STM4533, STM1246, STM1916, STM1917, STM1918, STM1244, STM1959, STM1921, STM2771	0.89	99.999981
stm00920:Sulfur metabolism	3	STM2933, STM4063, STM2946	0.98	99.999993
stm00010:Glycolysis / Gluconeogenesis	3	STM4085, STM1290, STM1203	0.95	99.999996
stm01230:Biosynthesis of amino acids	9	STM3331, STM3330, STM0110, STM0111, STM0112, STM0064, STM0113, STM1290, STM3909	0.86	99.999997
stm00330:Arginine and proline metabolism	2	STM0165, STM0166	1.13	99.999998
stm00030:Pentose phosphate pathway	3	STM4085, STM1885, STM3542	0.89	99.999999
stm00270:Cysteine and methionine metabolism	2	STM0165, STM0166	0.86	100
stm00240:Pyrimidine metabolism	3	STM4452, STM3415, STM4459	0.70	100

Table A.7 *cont*. KEGG pathways represented by down-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm02010:ABC transporters	10	STM3564, STM1891, STM3567, STM3561, STM0877, STM3560, STM3563, STM3562, STM0771, STM4063	0.72	100
stm00230:Purine metabolism	3	STM2933, STM4452, STM3415	0.48	100
stm01120:Microbial metabolism in diverse environments	12	STM4085, STM1761, STM2259, STM3331, STM3330, STM1885, STM2933, STM0064, STM3150, STM1290, STM3542, STM2946	0.62	100
stm01200:Carbon metabolism	4	STM4085, STM1885, STM1290, STM3542	0.44	100
stm01130:Biosynthesis of antibiotics	7	STM4085, STM3331, STM3330, STM0064, STM1290, STM3542, STM3909	0.43	100
stm01110:Biosynthesis of secondary metabolites	12	STM4085, STM3331, STM3330, STM2308, STM0110, STM0111, STM0112, STM0064, STM0113, STM1290, STM3542, STM3909	0.49	100
stm01100:Metabolic pathways	33	STM3870, STM3869, STM1290, STM3866, STM4085, STM3864, STM4086, STM1479, STM3909, STM3542, STM1831, STM1830, STM1832, STM2946, STM3331, STM3330, STM1885, STM2933, STM4452, STM0064, STM3070, STM0165, STM1480, STM0166, STM2325, STM2308, STM0110, STM0111, STM0112, STM0113, STM0684, STM3415, STM4459	0.57	100
stm00130:Ubiquinone and other terpenoid-quinone biosynthesis	1	STM2308	0.59	100
stm00620:Pyruvate metabolism	1	STM0113	0.22	100
stm00410:beta-Alanine metabolism	1	STM0166	1.13	100
stm00480:Glutathione metabolism	1	STM0166	0.65	100
stm00680:Methane metabolism	1	STM4085	0.41	100

Table A.7 *cont*. KEGG pathways represented by down-regulated genes in *S. enterica ridA* mutants

Term	#	Genes	Fold Enrichment	FDR
stm00261:Monobactam biosynthesis	1	STM0064	1.38	100
stm00970:Aminoacyl-tRNA biosynthesis	1	STM3895	0.11	100
stm00770:Pantothenate and CoA biosynthesis	1	STM3909	0.52	100
stm03060:Protein export	1	STM3420	0.65	100
stm03018:RNA degradation	1	STM4330	0.78	100
stm00300:Lysine biosynthesis	1	STM0064	0.83	100
stm00633:Nitrotoluene degradation	1	STM3150	1.38	100
stm00750:Vitamin B6 metabolism	1	STM3070	1.24	100
stm03020:RNA polymerase	1	STM3415	3.10	100
stm03440:Homologous recombination	1	STM4392	0.46	100
stm00640:Propanoate metabolism	1	STM3164	0.34	100
stm00561:Glycerolipid metabolism	1	STM4086	1.03	100

Table A.8. Metabolites Identified by ^{1}H-NMR in Pellet and Medium Samples and Confidence Levels

Exogenous Metabolites			
Metabolite	Assignment	^{1}H Chemical shift peaks (ppm) [Multiplicity]	Confidence Level
Lactic acid	CH_3	1.259, 1.266 [d]	4
	CH	4.062 [q]	4
Valine	αCH	3.636, 3.641 [d]	4
	βCH	2.253-2.302 [m]	4
	γCH_3	0.997, 1.006 [d]	4
	γCH_3	1.039, 1.048 [d]	4
Formate	CH	8.488 [s]	4
Uracil	CH	5.830, 5.839 [d]	4
	CHNH	7.617, 7.625 [d]	4
2-isopropylmalic acid	CH_2	2.547, 2.567, 2.681, 2.700 [d of d]	4
	CH_3	0.838, 0.845 [d]	4
	CH_3	0.899, 0.907 [d]	4
Putrescine	CH_2	1.793 [m]	4
	CH_2NH_2	3.050 [t]	4
2-aminobutyric acid	CH_3	0.974, 0.984 [t]	4
	CH_2	1.90, 1.91, 1.92 [m]	4
	CH	3.75, 3.75, 3.76 [t]	4
Acetate	CH_3	1.96 [s]	4
acetyl-phosphate	CH_3	2.095 [s]	4

Table A.8 *cont.* Metabolites Identified by ^{1}H-NMR in Pellet and Medium Samples and Confidence Levels

Endogenous Metabolites			
Metabolite	Assignment	^{1}H Chemical shift peaks (ppm) [Multiplicity]	Confidence Level
Acetate	CH_3	1.912 [s]	4
Alanine	CH_3	1.469, 1.484 [d]	4
Coenzyme A	CH_3	0.726 [s]	4
	CH_3	0.853 [s]	4
	CH_2	2.445, 2.456, 2.470 [t]	4
	CH_2	3.308 [s]	4
	CH	6.182, 6.191 [d]	4
Ethanolamine	CH_2NH_2	3.128, 3.136, 3.145 [t]	4
Formate	CH	8.453 [s]	4
Glutamate	αCH_2	2.017 [m]	4
	βCH_2	2.334-2.346 [m]	4
	CH	3.774 [m]	4
Glutamine		2.170 [s]	4
N-acetylputrescine	CH_2	1.565-1.602 [m]	4
	CH_2	1.667-1.695 [m]	4
	CH_3	1.983 [s]	4
	CH_2NH_2	2.992, 3.006, 3.019 [t]	4
	CH_2	3.192, 3.203, 3.214 [t]	4
Nicotinate		8.596 [s]	4
		8.603 [s]	4
		9.289 [s]	4
Phenylalanine	CH	7.311, 7.323 [d]	4
	CH	7.349-7.374 [m]	4
	CH	7.403-7.428 [m]	4
Putrescine	CH_2	1.770 [m]	4
	CH_2NH_2	3.042 [t]	4
Pyruvate	CH_3	2.359 [s]	4
Succinate	CH_2	2.398 [s]	4
Threonine	CH_3	1.319, 1.330 [d]	4
	CH	4.238-4.245 [m]	4
Uracil	CH	5.784, 5.798 [d]	4
	CHNH	7.525, 7.537 [d]	4
Valine	γCH_3	0.976 [d]	4
	γCH_3	1.028 [d]	4
	αCH	3.653, 3.66 [d]	4
	βCH	2.251-2.291 [m]	4

Table A.9 Integration values of exogenous metabolites identified in Chapter 7

EXOGENOUS METABOLITES

Metabolite	1H Peak (ppm) Integrated [Multiplicity]	Growth Medium	WT Mean (SD)[A]	*ridA* Mean (SD)[A]	FC [*ridA/WT*][B] (p-value)
Lactic acid	1.259, 1.266 [d]	Gly	N/A	0.01 (± 0.01)	**Undetermined**
		Min	N/A	0.06 (± 0.01)	**Undetermined**
		Ile	N/A	N/A	**Undetermined**
Valine	1.039, 1.048 [d]	Gly	1.32 (± 0.48)	1.93 (± 0.45)	**1.47 (0.0086)**
		Min	1.23 (± 0.33)	6.11 (± 0.93)	**4.97 (6.17×10^{-12})**
		Ile	0.93 (± 0.36)	0.82 (± 0.31)	**Not Significant**
Uracil	7.617, 7.625 [d]	Gly	0.14 (± 0.02)	0.10 (± 0.03)	**-1.47 (6.29×10^{-4})**
		Min	0.14 (± 0.03)	0.05 (± 0.01)	**-3.12 (9.90×10^{-10})**
		Ile	0.07 (± 0.02)	0.09 (± 0.01)	**Not Significant**
2-aminobutyric acid	1.90, 1.91, 1.92 [m]	Gly	0.91 (± 0.44)	N/A	**Undetermined**
		Min	0.28 (± 0.19)	N/A	**Undetermined**
		Ile	N/A	N/A	**Undetermined**
2-isopropylmalic acid	0.838, 0.845 [d]	Gly	0.31 (± 0.09)	0.45 (± 0.23)	**Not Significant**
		Min	0.22 (± 0.08)	0.51 (± 0.12)	**2.35 (6.01×10^{-6})**
		Ile	0.05 (± 0.03)	0.05 (± 0.03)	**Not Significant**
Putrescine	3.050 [t]	Gly	0.50 (± 0.08)	0.53 (± 0.16)	**Not Significant**
		Min	0.54 (± 0.06)	0.79 (± 0.16)	**1.44 (2.63×10^{-4})**
		Ile	0.36 (± 0.09)	0.38 (± 0.04)	**Not Significant**
Acetyl-phosphate	2.095 [s]	Gly	0.04 (± 0.01)	0.04 (± 0.01)	**Not Significant**
		Min	0.03 (± 0.01)	0.05 (± 0.01)	**1.33 (6.82×10^{-5})**
		Ile	0.05 (± 0.05)	0.03 (± 0.01)	**Not Significant**
Formic acid	8.488 [s]	Gly	34.91 (± 5.38)	34.43 (± 7.50)	**Not Significant**
		Min	38.12 (± 5.35)	13.85 (± 5.79)	**-2.75 (1.36 x 10-8)**
		Ile	30.54 (± 11.64)	33.27 (± 3.20)	**Not Significant**
Acetic acid	1.96 [s]	Gly	30.83 (± 16.66)	39.56 (± 13.11)	**Not Significant**
		Min	41.66 (± 14.65)	41.18 (± 9.12)	**Not Significant**
		Ile	27.65 (± 16.03)	30.59 (± 15.59)	**Not Significant**

[A]N/A- Peak not detectable above baseline signal

[B]Not Significant - P-value < 0.05;

Undetermined - Fold-change could not be calculated since one or both conditions did not show peak instensity above the baseline;

Blue shading indicated significantly more metabolite [test condition/strain vs. control condition/strain] and red shading indicated the opposite

Table A.9 *cont*. Integration values of exogenous metabolites identified in Chapter 7

ENDOGENOUS METABOLITES

Metabolite	1H Peak (ppm) Integrated [Multiplicity]	Growth Medium	WT Mean (SD)[A]	*ridA* Mean (SD)[A]	FC [*ridA/WT*][B] (p-value)
Acetate	1.912 [s]	Gly	0.91 (± 0.25)	1.02 (± 0.29)	**Not Significant**
		Min	1.03 (± 0.33)	0.69 (± 0.24)	**-1.47 (0.0188)**
		Ile	0.08 (± 0.04)	0.08 (± 0.04)	**Not Significant**
Alanine	1.469, 1.484 [d]	Gly	0.32 (± 0.11)	0.42 (± 0.06)	**1.32 (0.0189)**
		Min	0.21 (± 0.08)	0.30 (± 0.05)	**1.42 (0.0090)**
		Ile	0.32 (± 0.11)	0.28 (± 0.07)	**Not Significant**
Coenzyme A	0.726 [s]	Gly	0.25 (± 0.06)	0.19 (± 0.04)	**-1.31 (0.0156)**
		Min	0.31 (± 0.06)	0.19 (± 0.04)	**-1.61 (3.67×10^{-5})**
		Ile	0.20 (± 0.07)	0.25 (± 0.12)	**Not Significant**
Ethanolamine	3.128, 3.136, 3.145 [t]	Gly	0.14 (± 0.02)	0.14 (± 0.01)	**Not Significant**
		Min	0.14 (± 0.01)	0.11 (± 0.01)	**-1.31 (1.92×10^{-5})**
		Ile	0.14 (± 0.02)	0.15 (± 0.01)	**Not Significant**
Formate	8.453 [s]	Gly	0.18 (± 0.10)	0.16 (± 0.07)	**Not Significant**
		Min	0.22 (± 0.09)	0.06 (± 0.03)	**-3.80 (3.43×10^{-5})**
		Ile	0.17 (± 0.13)	0.25 (± 0.08)	**Not Significant**
Glutamate	2.334-2.346 [m]	Gly	0.72 (± 0.14)	0.74 (± 0.23)	**Not Significant**
		Min	0.51 (± 0.17)	0.75 (± 0.20)	**1.47 (0.0100)**
		Ile	0.52 (± 0.14)	0.48 (± 0.15)	**Not Significant**
Glutamine	2.170 [s]	Gly	0.10 (± 0.07)	0.12 (± 0.09)	**Not Significant**
		Min	0.08 (± 0.03)	0.19 (± 0.13)	**2.40 (0.0194)**
		Ile	0.06 (± 0.02)	0.07 (± 0.03)	**Not Significant**
N-acetylputrescine	3.192, 3.203, 3.214 [t]	Gly	0.09 (± 0.02)	0.08 (± 0.04)	**Not Significant**
		Min	0.07 (± 0.02)	0.14 (± 0.04)	**1.92 (2.56×10^{-4})**
		Ile	0.09 (± 0.05)	0.07 (± 0.03)	**Not Significant**

[A]N/A- Peak not detectable above baseline signal

[B]Not Significant - P-value < 0.05;
Undetermined - Fold-change could not be calculated since one or both conditions did not show peak
Blue shading indicated significantly more metabolite [test condition/strain vs. control

Table A.9 *cont*. Integration values of exogenous metabolites identified in Chapter 7

ENDOGENOUS METABOLITES

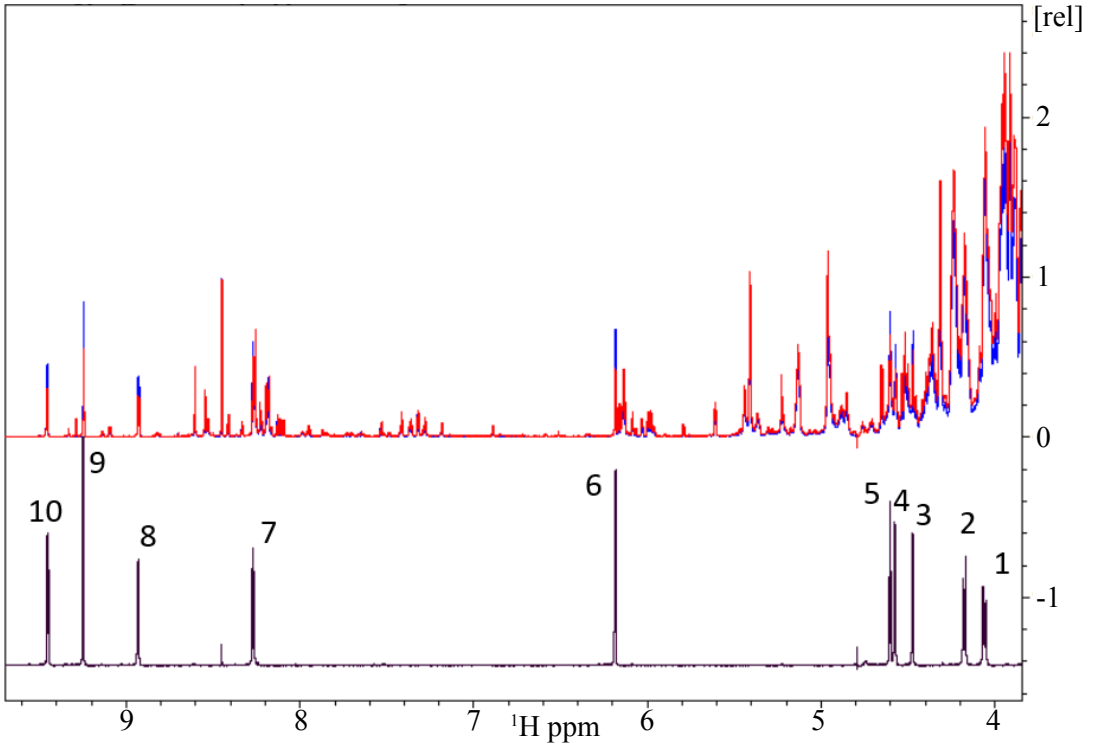
Metabolite	1H Peak (ppm) Integrated [Multiplicity]	Growth Medium	WT Mean (SD)[A]	*ridA* Mean (SD)[A]	FC [*ridA/WT*][B] (p-value)
Nicotinate	8.596 [s]	Gly	0.03 (± 0.01)	0.03 (± 0.01)	**Not Significant**
		Min	0.03 (± 0.01)	0.03 (± 0.01)	**Not Significant**
		Ile	0.03 (± 0.01)	0.03 (± 0.01)	**Not Significant**
Phenylalanine	7.311, 7.323 [d]	Gly	0.03 (± 0.01)	0.03 (± 0.02)	**Not Significant**
		Min	0.02 (± 0.01)	0.04 (± 0.01)	**1.83 (1.19 x 10^{-4})**
		Ile	0.04 (± 0.02)	0.03 (± 0.01)	**Not Significant**
Putrescine	1.770 [m]	Gly	0.77 (± 0.19)	0.78 (± 0.29)	**Not Significant**
		Min	0.92 (± 0.27)	0.59 (± 0.20)	**-1.55 (0.0063)**
		Ile	0.75 (± 0.29)	0.90 (± 0.30)	**Not Significant**
Pyruvate	2.359 [s]	Gly	0.05 (± 0.01)	0.04 (± 0.02)	**Not Significant**
		Min	0.03 (± 0.01)	0.04 (± 0.01)	**Not Significant**
		Ile	0.03 (± 0.01)	0.03 (± 0.01)	**Not Significant**
Succinate	2.398 [s]	Gly	0.08 (± 0.06)	0.09 (± 0.05)	**Not Significant**
		Min	0.08 (± 0.06)	0.12 (± 0.11)	**Not Significant**
		Ile	0.06 (± 0.03)	0.08 (± 0.05)	**Not Significant**
Threonine	1.319, 1.330 [d]	Gly	0.19 (± 0.05)	0.23 (± 0.08)	**Not Significant**
		Min	0.16 (± 0.02)	0.49 (± 0.15)	**3.13 (2.82 x 10^{-6})**
		Ile	0.23 (± 0.07)	0.21 (± 0.05)	**Not Significant**
Uracil	5.784, 5.798 [d]	Gly	0.19 (± 0.05)	0.19 (± 0.05)	**Not Significant**
		Min	0.17 (± 0.04)	0.22 (± 0.08)	**Not Significant**
		Ile	0.21 (± 0.06)	0.20 (± 0.05)	**Not Significant**
Valine	1.028 [d]	Gly	0.20 (± 0.07)	0.26 (± 0.12)	**Not Significant**
		Min	0.15 (± 0.06)	0.52 (± 0.13)	**3.47 (1.72 x 10^{-7})**
		Ile	0.08 (± 0.04)	0.08 (± 0.04)	**Not Significant**

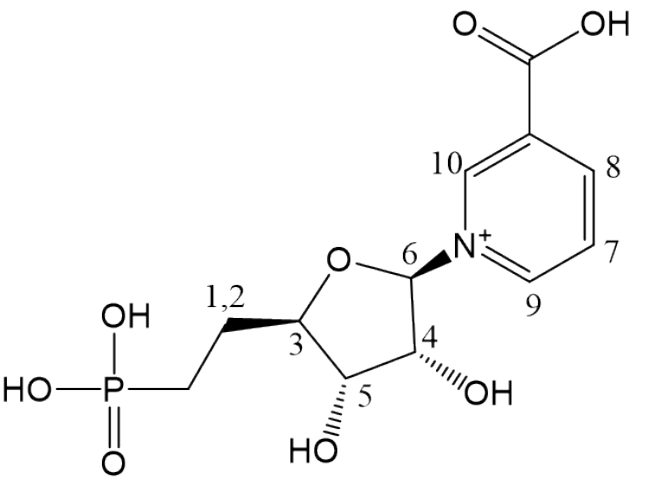
[A]N/A- Peak not detectable above baseline signal

[B]Not Significant - P-value < 0.05;

Undetermined - Fold-change could not be calculated since one or both conditions did not show peak

Blue shading indicated significantly more metabolite [test condition/strain vs. control





Nicotinate mononucleotide (NMN)

Atom	^{1}H peak	^{13}C peak
1	4.05	65.87
2	4.17	65.87
3	4.47	73.59
4	4.57	90.34
5	4.6	80.36
6	6.18	102.75
7	8.27	131.26
8	8.93	149.55
9	9.25	144.39
10	9.45	143.93

Figure A.1. Nicotinate mononucleotide identification by sample spiking. The ^{1}H NMR spectrum was obtained for the pooled metabolite extraction samples with (blue line) and without (red line) the inclusion of 150 μM pure NMN standard. The ^{1}H NMR spectra of NMN alone is provided below (black line). Labeled peaks correspond to labeled atoms on the structural diagram. Exact ^{1}H NMR and ^{13}C peak assignments are listed in the adjacent table.