DEFINING METABOLIC REQUIREMENTS OF THE THI5 PATHWAY FOR THIAMINE BIOSYNTHESIS

by

MICHAEL DANIEL PAXHIA

(Under the Direction of Diana M. Downs)

ABSTRACT

The metabolic network within organisms is comprised of individual modules that are interconnected. Although the overall architecture of individual networks can be predicted due to the advent of affordable whole genome sequencing, subtle differences in network structure lead to differences in function of individual modules between organisms. In addition, moving metabolic modules between organisms in an effort to build new pathways is encumbered by the lack of knowledge about the functional relationships within and between metabolic networks. Thus, metabolic engineering instead relies on building heterologous pathways by tedious trial and error. Thiamine biosynthesis has been used as a tool to dissect connections within metabolic networks and can be used to extend our understanding about the functional relationships found within metabolism. This thesis uses Thi5 as a model metabolic node to understand the differences between the metabolic networks of *Saccharomyces cerevisiae*, *Legionella pneumophila* and *Salmonella enterica* and determine the new metabolic requirements for 4-amino-5- hydroxymethyl-2-methylpyrimidine-phosphate (HMP-P) biosynthesis when Thi5 is incorporated into the metabolic network of *S. enterica*. The first part of this work identifies a requirement of

additional pyridoxal-5'-phosphate (PLP) synthases for *Sc*Thi5 function in the native network of *S. cerevisiae*. The second investigation extends the Thi5-dependent pathway for HMP-P biosynthesis into *Legionella pneumophila* and identified functional differences between *Sc*Thi5 and *Lp*Thi5. Finally, the third investigation identifies that both alpha-ketoglutarate and PLP are required to integrate *Sc*Thi5 into the *S. enterica* metabolic network and suggests that structural differences between *Lp*Thi5 and *Sc*Thi5 may explain differences in function.

INDEX WORDS: *THI5*, Thiamine, Pyridoxal 5'-phosphate, *Saccharomyces cerevisiae*, *Legionella pneumophila*, Metabolism

DEFINING METABOLIC REQUIREMENTS OF THE THI5 PATHWAY FOR THIAMINE BIOSYNTHESIS

by

MICHAEL DANIEL PAXHIA

B.S., Michigan State University, 2013

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

DOCTER OF PHILOSOPHY

ATHENS, GEORGIA

2021

DEFINING METABOLIC REQUIREMENTS OF THE THI5 PATHWAY FOR THIAMINE BIOSYNTHESIS

by

MICHAEL DANIEL PAXHIA

Major Professor:	Diana M. Downs
Committee:	Michael W. Adams
	Jorge C. Escalante
	David J. Garfinkel
	Timothy R. Hoover

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
May 2021

ACKNOWLEDGEMENTS

I am thankful to many people who supported and guided me during my education. First, I would like to thank my advisor, Dr. Diana Downs, for her mentorship and guidance. Her enthusiasm for understanding metabolism and scientific rigor both contributed toward making me a better scientist. I also appreciate the feedback on this work from my committee members Dr. Jorge Escalante, Dr. Michael Adams, Dr. David Garfinkel and Dr. Timothy Hoover.

The past and current members of the Downs and Escalante labs were very helpful scientifically and personally during graduate school. I would especially like to thank Dr. Dustin Ernst, Dr. Norbert Tavares, Dr. Andrew Borchert, Dr. Flavia Costa, Dr. Tomokazu Ito, Dr. Jessica Irons, and Huong Vu for their feedback and ideas.

I also would like to thank the mentors and teachers who inspired me to pursue scientific research including Walter Erhardt, Mary Lindow, Dr. Chuck Elzinga, Dr. Pankaj Soni, Dr. Jyh-Song Horng and Dr. Yousong Ding. I would also like to thank Dr. Gemma Reguera, Dr. Jenna Young, and Dr. Svetlana Gerdes who encouraged me to pursue a PhD.

Finally, I would like to thank my family and friends for their love and support; especially my partner, Sophia. To my parents, Jeff and Carol, thank you for your encouragement and for planting the seed of my scientific education.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iv

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

THIAMINE BIOSYNTHESIS 4

PYRIDOXAL-5'-PHOSPHATE BIOSYNTHESIS 11

PLP AND THIAMINE BIOSYNTHESIS AS MODEL NODES TO UNDERSTAND METABOLIC NETWORK STRUCTURE 13

MODIFICATION OF THE *S. ENTERICA* METABOLIC NETWORK IS REQUIRED TO ALLOW *Sc*Thi5-DEPENDENT THIAMINE SYNTHESIS 15

DISSERTATION OUTLINE 16

REFERENCES 18

2 *SNZ3* ENCODES A PLP SYNTHASE INVOLVED IN THIAMINE SYNTHESIS IN *SACCHAROMYCES CEREVISIAE* 27

ABSTRACT 28

INTRODUCTION 28

MATERIALS AND METHODS 31

RESULTS AND DISCUSSION 38

CONCLUSIONS 43

REFERENCES 44

3 FUNCTIONAL CHARACTERIZATION OF THE HMP-P SYNTHASE OF *LEGIONELLA PNEUMOPHILA* (Lpg1565).....58

ABSTRACT.....59

INTRODUCTION 60

MATERIALS AND METHODS.....62

RESULTS AND DISCUSSION.....72

CONCLUSIONS 83

REFERENCES 84

4 EVIDENCE FOR A ROLE OF α-KETOGLUTARATE LEVELS IN THE FUNCTION OF THE HETEROLOGOUS Thi5 PATHWAY IN *SALMONELLA ENTERICA* 100

ABSTRACT.....101

INTRODUCTION 102

RESULTS AND DISCUSSION.....105

CONCLUSIONS 112

MATERIALS AND METHODS.....114

REFERENCES 119

5 CONCLUSIONS 133

FUTURE DIRECTIONS 134

REFERENCES 139

APPENDICES

A INDUCTION OF THE SUGAR-PHOSPHATE STRESS RESPONSE ALLOWS *SACCHAROMYCES CEREVISIAE* 2-METHYL-4-AMINO-5-HYDROXYMETHYLPYRIMIDINE PHOSPHATE SYNTHASE TO FUNCTION IN *SALMONELLA ENTERICA* 141

ABSTRACT.....142

INTRODUCTION 143

MATERIALS AND METHODS.....146

RESULTS AND DISCUSSION.....152

CONCLUSIONS 158

REFERENCES 160

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

An integrated metabolism is feature of all cells in which all pathways extend from or lead to central carbon metabolism that provides a set of 12 common intermediates for biomass (1). The intermediates include ribose-5-phosphate, erythrose-4-phosphate, glucose-6-phosphate, fructose-6-phosphate, glyceraldehyde-3-phosphate, glycerate-3-phosphate, phosphoenolpyruvate, pyruvate, acetyl-CoA, α-ketoglutarate, succinyl-CoA and oxaloacetate. Central carbon metabolism comprises of glycolysis, the pentose phosphate pathway, and the TCA cycle. Beyond these conserved pathways, organisms also contain a multitude of primary metabolic pathways for amino acid, nucleotide, vitamin, lipid, and cell wall biosynthesis. Some organisms also encode specialized metabolic pathways to produce secondary metabolites that enhance fitness under certain environmental conditions. These pathways can be regulated at many levels by several mechanisms including transcriptional activation/repression, translational regulation (e.g. riboswitches), post-translational modifications, feed-forward metabolic flux activation, and feedback inhibition (2, 3). These levels of regulation combine to generate an effect where changes in metabolite abundance coordinate efficient growth as well as changes in life cycle (3–6). The establishment of persister cells is an example of how metabolism can coordinate changes in life cycle.

Recent work in *E. coli* has established that flux through glycolysis can be responsible for the formation of persister cells as an indirect consequence of Cra regulation by fructose-1,6-bisphosphate (5). The study found that only a small subpopulation of *E. coli* that was exponentially growing on glycolytic carbon sources grew on gluconeogenic carbon sources, consistent with the lag phase observed in growth. The investigation showed that this was not due to emergence of spontaneous suppressors, as the cells harvested at the end of the growth experiment retained the bistable phenotype. Additionally, the phenotype correlated with Cra regulation, and the majority of cells in the population were found to be resistant to antibiotics that kill growing cells on gluconeogenic carbon sources. Thus, the investigation suggested that these cells established a metabolically dormant persister state on gluconeogenic substrates in response to metabolic flux through glycolysis.

Despite the complexity of integration of metabolic pathways with these multiple levels of regulation, there has been progress toward modeling and predicting the effects of perturbations in organisms with a minimal metabolism (7). In these studies, *in silico* networks were trained on biochemical data from *Mycoplasma genitalium* to simulate and predict how perturbations may affect cell growth. These networks were based on a genome scale metabolic model and kinetic models of transcription, translation and cell division. Interestingly, an emergent property of the network was that cell division was coordinated in response to metabolic changes. However, in more complex model organisms like *Salmonella enterica* and *Escherichia coli* and in organisms with unique lifestyles such as *Shewanella oneidensis*, it is difficult to predict the result of genetic or metabolic perturbations due to the integrated nature of these more complex metabolic networks and subtle changes in their network structure (8–12).

Our lack of understanding of how metabolic network structure is related to function is demonstrated when attempts to integrate new pathways into organisms lead to incompatibilities (13–15). These errors can include depletion of shared intermediates, production of inhibitory intermediates, and unbalanced redox pathways to produce products. Additionally, the accurate prediction of the function of modules in heterologous hosts based on phylogenetic relationships between organisms have not been reliable. In one study, when testing function of a gene encoding a dichloromethane dehalogenase (*dcmA*) acquired by horizontal gene transfer in *Methylobacterium extorquens* DM4 as a heterologous pathway in other *M. extorquens* isolates and more distant *Methylobacterium* hosts, the heterologous pathway performed best in the distant *Methylobacterium* hosts (16). In a second study, heterologous expression of a biosynthetic gene cluster for violacein production from *Pseudoalteromonas luteoviolacea* 2ta16 led to higher titers in more distantly related organisms (*Pseudomonas putida*, *Agrobacterium tumefaciens*) over the more closely related *Escherichia coli* production strains (17). Together, these and other studies emphasize that research into fundamental principles regarding metabolic network structure and its relationship to the integration of heterologous metabolism may improve efforts to construct new pathways in synthetic biology. This has the potential to improve processes and reduce failure rates rather than relying on "brute-force" methods.

To dissect more complex systems, physiological studies of model nodes in metabolism like thiamine biosynthesis can be used to gain a deeper understanding of metabolic network structure and function (18, 19). Several examples described in more detail later in this chapter include production of phosphoribosylamine (PRA) through divergent mechanisms, establishing connections between iron sulfur cluster biogenesis and

thiamine biosynthesis, and determining the enzymatic function encoded by a gene of unknown function (*yjgF/ridA*). Additionally, models of metabolic connectivity obtained from basic research are important to dissect differences in conserved nodes of metabolism. These models may have future applications in streamlining development of organisms to produce value-added products.

Work in this thesis describes connections between two nodes of metabolism, thiamine and pyridoxal-5'-phosphate (PLP), in *Saccharomyces cerevisiae* and *Legionella pneumophila*. Additionally, this thesis describes connections between discrete areas of metabolism required for integration of a heterologous pathway into *S. enterica* metabolism for thiamine biosynthesis.

1.1 THIAMINE BIOSYNTHESIS

Thiamine pyrophosphate (TPP), the cofactor derived from Vitamin B_1, is required for reactions in central carbon metabolism including transketolase, pyruvate dehydrogenase, and α-ketoglutarate dehydrogenase. The cofactor is synthesized *de novo* by Bacteria, Archaea, and some Eukaryotes by independently producing the 4-methyl-5-(2-hydroxyethyl)-thiazole-phosphate (THZ-P) and 4-amino-5- hydroxymethyl-2-methylpyrimidine-pyrophosphate (HMP-PP) moieties, combining them into thiamine monophosphate (TMP), and converting TMP into TPP (20). In many eukaryotes, including mammals, vitamin B_1 is an essential nutrient obtained in the diet, and in humans lack of thiamine leads to the development of beriberi (21). In beriberi the central nervous system as well as the cardiovascular system of humans are affected and if left untreated, can lead to death. Beyond being a required cofactor for human health, thiamine is also an important

micronutrient for marine microbial ecosystems (22). Many marine eukaryotes lack complete thiamine biosynthetic pathways and instead rely on salvage of thiamine or degradation products including HMP, making thiamine an essential nutrient for these microbes within these ecosystems.

The synthesis of the thiazole moiety for thiamine in Bacteria uses a conserved pathway with two variations. The two variations differ in the origin of the 2-iminoacetate for the thiazole synthase ThiG. In *E. coli*, 2-iminoacetate is synthesized from tyrosine via ThiH (E.C. 4.1.99.19) generating 2-iminoacetate and 4-methylphenol using a radical s-adenosylmethionine (SAM) mechanism (23). Aerobic bacteria like *Bacillus subtilis* instead use ThiO (E.C. 1.4.3.19) to generate 2-iminoacetate (24). ThiO directly oxidizes glycine using an FAD-dependent oxidase. Beyond the origin of the 2-iminoacetate moiety, the majority of the pathway for thiazole synthesis is generally conserved. A persulfide is mobilized from cysteine by the sulfurtransferase IscS (E.C. 2.8.1.7) to the sulfur-carrying protein ThiI (E.C. 2.8.1.4) (25). This activated sulfur moiety is then transferred to an adenylated ThiS for subsequent use in the thiazole synthase reaction (E.C. 2.8.1.-) (26). The ThiG synthase (E.C. 2.8.1.10) then generates carboxy-thiazole phosphate (cTHZ-P) from deoxyxyulose-phosphate, 2-iminoacetate and the activated sulfur bound to ThiS (23).

Another pathway to cTHZ-P is found in yeast and Archaea and is initiated using the enzyme Thi4 (27). This enzyme converts NAD and glycine directly into adenylated thiazole, which is then processed by an unknown nudix hydrolase and phosphorylated into cTHZ-P. There is some debate in the literature about the function of Thi4, which has been previously described as a suicide enzyme. There is evidence that in *Saccharomyces cerevisiae* a conserved cysteine residue on Thi4 is sacrificed to insert a sulfur atom in the

absence of exogenous sulfide *in vitro* (28). Regeneration of this cysteine moiety in the enzyme has not been investigated further, however, persulfides like L-cysteine hydropersulfide (2.5 μM) or glutathione persulfide (30 μM) and free sulfide in the form of H_2S are abundant sources of sulfide in the cellular milieu of eukaryotic cells (29) and may be used in the enzymatic reaction or used to regenerate the cysteine in the active site. Additionally, in archaeal genomes that associate in environments containing free sulfide, Thi4 homologs have a conserved histidine instead of cysteine coordinating Fe2+. When purified, Thi4 from *Methanococcus jannaschii* uses sulfide directly as the sulfur source (30). Importantly, when alanine and serine substitutions were made at the conserved cysteine in *Sc*Thi4 from yeast, these protein variants were able to generate the adenylated thiazole precursor with the addition of H_2S, suggesting that free sulfide can be used under certain conditions (30, 31). It is unclear if free sulfide was tested as the sulfur donor with wild-type *Sc*Thi4 or if this is the sulfide donor *in vivo*.

Of particular relevance to the work described in this thesis, there are two pathways identified for the synthesis of the HMP-P moiety, the ThiC-dependent pathway and the Thi5-dependent pathway. This thesis focuses on pathways for HMP-P synthesis, which are detailed below.

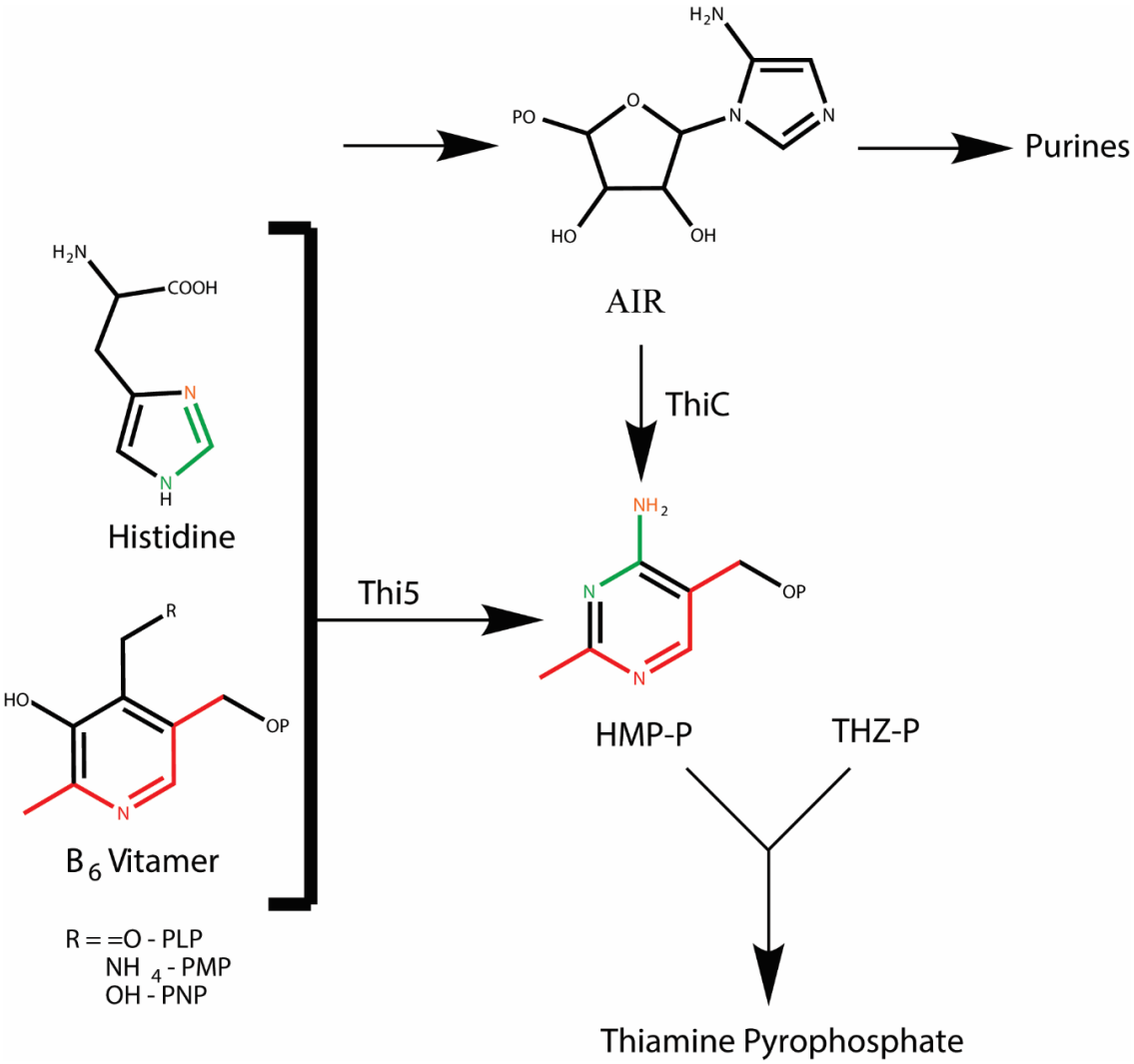


Figure 1.1 – Two pathways for HMP-P biosynthesis. The two routes to HMP-P are shown. Atoms in HMP-P derived from labeling studies in yeast are highlighted in color. Abbreviations: AIR, 5-aminoimidazole ribotide; PLP, pyridoxal-5'-phosphate; PMP, pyridoxamine-5'-phosphate; PNP, pyridoxine-5'-phosphate; HMP-P, 5-hydroxymethyl-2-methylpyrimidine phosphate; THZ-P, 4-methyl-5-(2-hydroxyethyl)-thiazole-phosphate.

ThiC pathway for HMP-P synthesis

A pathway to HMP-P conserved across plants, archaea and the majority of bacteria is the ThiC pathway. ThiC produces HMP-P using 5-aminoimidazole ribotide (AIR) and s-adenosyl-methionine (SAM), generating formate and carbon monoxide as additional byproducts (32). The full reaction from AIR to HMP-P is unknown, however, a proposed

mechanism implicates multiple hydrogen abstractions from the substrate (32). Initial reconstitution of purified ThiC with the addition of crude extracts showed that SAM was required for activity (33). ThiC lacks the classic CX_3CX_2C motif found in other radical SAM enzymes (34, 35). However, physiological studies linked FeS cluster biogenesis to the function of this pathway (36), leading to its reconstitution as a new member of the radical SAM enzyme superfamily (37, 38).

In vitro work with ThiC found that it generates a backbone radical after the addition of SAM, and it was proposed that this radical may facilitate the intermolecular rearrangement to form HMP-P (37, 39). Initially when reconstituted, the reaction conditions produced less than one HMP-P molecule per protein (33, 37, 38). However, after optimization of the anaerobic purification, reconstitution of the [4Fe-4S] cluster and/or metal center before freezing and the addition of S-methyl-5-thioadenosine nucleosidase (MTAN) to degrade 5'-deoxyadenosine, ThiC was shown to catalyze multiple turnovers *in vitro* (40).

There are structural differences among ThiC homologs from diverse organisms, and these differences are found to correlate with aerobic vs anaerobic lifestyle of the respective organism (36). The structure of ThiC from *Arabidopsis thaliana* has been solved, and the active site distinguishes it from other radical SAM enzymes (41). Within canonical radical SAM enzymes, the [4Fe-4S] cluster coordinates the amino, carboxylate and sulfur group of the methionine moiety of SAM. In contrast, the ThiC [4Fe-4S] cluster only coordinates the sulfur atom of SAM, and the amino and carboxylate groups are instead coordinated by a secondary metal binding site defined by two strictly conserved histidines. In the available *At*ThiC structures, these histidine residues bind either a zinc or iron atom.

ThiC contains two active sites which are formed across the dimer interface, where the c-terminal [4Fe-4S] cluster from one monomer extends into the AIR binding site of the adjacent monomer (41).

Thi5 pathway for HMP-P biosynthesis

A second pathway to HMP-P was identified in *S. cerevisiae*, where labeling studies showed that the atoms of the pyrimidine moiety in TPP derive from histidine and a B6 vitamer (presumably PLP) (42–44). *S. cerevisiae* contains four paralogs of *THI5* (*THI5/11/12/13*), and each of the paralogs was sufficient to support HMP-P biosynthesis under aerobic conditions (45). This pathway is conserved in other yeast species, and a *THI5* mutant of *S. pombe* (*nmt1, thi3*) is an HMP auxotroph (46, 47). When overexpressed and purified, Thi5 from *Candida albicans* and *S. cerevisiae* co-purify with PLP (48, 49). *Sc*Thi5 and *Ca*Thi5 structures have been solved, and conserved residues in the active site (N11, W12, K62, H66) near PLP were shown to be important for function of *Sc*Thi5 *in vivo* (49). Additionally, alanine substitutions within the strictly conserved CCCXC motif proposed to be involved in metal binding compromised *Sc*Thi5 function *in vivo* (49). *Sc*Thi5 forms a dimer with a PLP-binding site in each monomer (49). To examine the mechanism of this enzyme, an initial biochemical assay for the Thi5 reaction was developed with *Ca*Thi5 (48). After incubation of the purified protein with Fe(II) and PLP under anaerobic conditions, the protein was exposed to oxygen and HMP-P was detected. Because this reaction was also repeated with N_{15}-labeled protein produced *in vivo* and the resultant HMP-P incorporated the label, it was proposed that Thi5 was a suicide HMP-P synthase *in vitro* (48). It is not clear if these reaction conditions reflect the reaction catalyzed *in vivo*.

The Thi5 protein from *S. cerevisiae* is a structural homolog of the ThiY component of the N-formyl-4-amino-5-(aminomethyl)-2-methylpyrimidine (FAMP) transporter found in bacteria contained in the ThiXYZ complex (50). ThiY was found to share 25 % amino acid identity with Thi5 from *S. cerevisiae*. Importantly, residues that coordinate binding the pyrimidine ring of FAMP are conserved in both ThiY and Thi5 (Asp38, Trp39, Glu192 and His45 in ThiY). This amino acid identity has likely led to several *THI5* orthologs being annotated as components of a putative ABC transporter in several bacterial species (51).

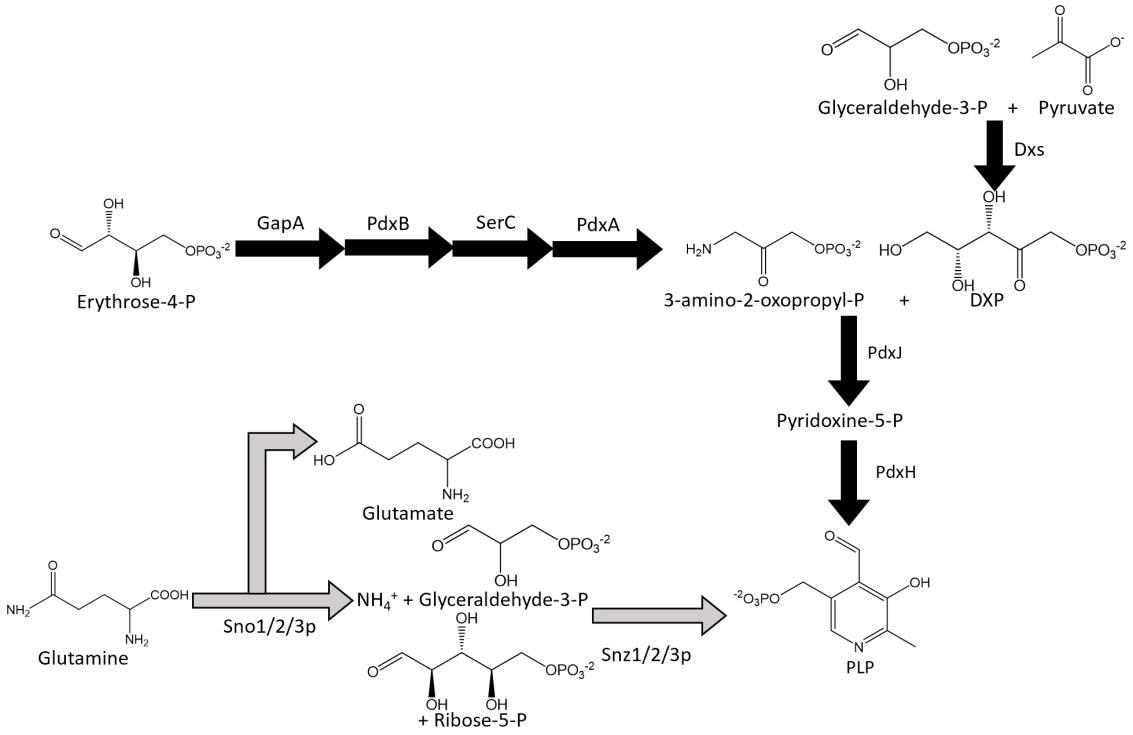


Figure 1.2 – Pathways for PLP biosynthesis. Shown are the pathways for PLP biosynthesis as described in *E. coli* (GapA, PdxB, SerC, PdxA, Dxs, PdxJ, PdxH) and *S. cerevisiae* (Sno1/2/3 and Snz1/2/3). The structure of key intermediates described in the text are shown. Abbreviations: DXP, deoxy-xylulose-5-phosphate; PLP, pyridoxal-5'-phosphate.

1.2 PYRIDOXAL-5'-PHOSPHATE BIOSYNTHESIS

PLP is an essential cofactor found in all free-living organisms that can facilitate a diverse set of biochemical reactions including transamination, elimination, and racemization reactions (52). In total, PLP-dependent enzymes account for ~4 % of all classified biochemical reactions, making it an important cofactor across metabolism. Although PLP enzymes fall into seven structural folds (53), PLP enzymes have catalytic promiscuity and have been shown to share activities catalyzed by other fold types (54, 55). This is likely because all PLP-dependent reactions stem from the formation of an external aldimine intermediate. Reactions then proceed based on the chemistry inherent in the activated substrate, the net charge on PLP, and the location of catalytic residues within the active site. The net charge of PLP is influenced by the amino acids coordinating the cofactor in the active site (56). Similar to the several routes to thiamine described above, there are two biochemical pathways described for the formation of PLP: the DXP-dependent and -independent pathways.

DXP-dependent PLP biosynthesis

S. enterica and other γ-proteobacteria use the deoxy-xylulose-5-phosphate (DXP)-dependent pathway for PLP biosynthesis (Figure 1.2) (57). This pathway in *E. coli* uses seven enzymatic steps to convert erythrose-4-phosphate, pyruvate and glyceraldehyde-3-phosphate into PLP (Figure 1.2). Deoxy-xyulose-5-phosphate is synthesized from glyceraldehyde-3-phosphate and pyruvate by Dxs (E.C. 2.2.1.7) (58). The transformation of erythrose-4-phosphate to 3-amino-1-hydroxyacetone-phosphate proceeds as follows. E4P is oxidized to 4-phospho-erythronate by Epd (E.C. 1.2.1.72) (59). Then, PdxB (E.C.

1.1.1.290) converts 4-phospho-erythronate into 2-oxo-3-hydroxy-4-phosphobutanoate (59). SerC (E.C. 2.6.1.52) then converts 2-oxo-3-hydroxy-4-phosphobutanoate into 4-phosphohydroxy-threonine (60) and PdxA (E.C. 1.1.1.262) then produces 3-amino-1-hydroxyacetone-phosphate (61, 62). The pyridoxine-5'-phosphate (PNP) synthase encoded by *pdxJ* (E.C. 2.6.99.2) synthesizes PNP from DXP and 3-amino-1-hydroxyacetone-phosphate (61). Finally, PNP is oxidized to PLP by PdxH (E.C. 1.4.3.5) (63).

The full DXP-dependent pathway as described in *E. coli* is not strictly conserved amongst organisms that contain homologs to *pdxJ*, which has led to the hypothesis that the pathway emerged by co-opting reactions found throughout metabolism in a patchwork-like manner (57). This is consistent with the finding that serendipitous pathways can bypass separate parts of the pathway in *E. coli* using promiscuous activities (64). Additionally, other closely related organisms like *S. enterica* derive some flux to PLP through these "serendipitous" pathways that branch off the defined pathway in *E. coli* (65).

DXP-independent PLP synthesis

An alternative pathway for PLP biosynthesis exists throughout bacteria, archaea, plants, and yeast that does not rely on DXP as an intermediate. Instead, this pathway synthesizes PLP from glyceraldyhyde-3-phosphate, ribose-5-phosphate and an ammonia source (66). This reaction is catalyzed by a two-component glutamine-hydrolyzing PLP synthase, where a glutaminase subunit funnels ammonia derived from glutamine into the synthase subunit (Figure 1.2) (66–68). Initially characterized from *Bacillus subtilis*, PLP synthases have also been characterized from *S. cerevisiae*, *Thermotoga maritima*, *Arabidopsis thaliana*, and *Plasmodium falciparum* (66, 69–75). Intermediates in the PLP synthase

reaction have been detected and characterized, defining a unique lysine relay mechanism used to facilitate PLP formation (71, 73–76). There are two distinct active sites within the PLP synthase, and active site 1 binds ribose-5'phosphate, facilitates the formation of a chromophoric intermediate (I_{320}) with ammonia, which is then transferred by a swinging lysine into the second active site. The second active site then transforms the imine into PLP through a condensation and cyclization reaction with glyceraldehyde-3'phosphate. Although several eukaryotes contain multiple orthologs of the PLP synthase, the redundancy of this pathway in *S. cerevisiae* had not been dissected previously.

1.3 PLP AND THIAMINE BIOSYNTHESIS AS MODEL NODES TO UNDERSTAND METABOLIC NETWORK STRUCTURE

Vitamins have been used to study the metabolic network structure in both *E. coli* and *S. enterica* (18, 64, 77). This is because the pathways for vitamin biosynthesis form nodes of metabolism that are essential for growth on a minimal medium, and most intermediates can be salvaged providing a simple way to monitor deficiencies within the pathway of interest. Previous work has used pathways to PLP and thiamine to extend our knowledge of metabolic pathways.

PLP biosynthesis in *E. coli* has been used as a model node to uncover subtle activities found in overexpressed enzymes, and some of these pathways exist natively in other organisms (64, 65, 78). In the first study it was found that three new pathways to intermediates in PLP biosynthesis were identified while overexpressing HisB, Php, YjbQ, AroB, YeaB, and ThrB (64). Overexpressing HisB, Php or YjbQ likely feeds into the 2-oxo-3-hydroxy-4-phosphobutanoate branch point in the DXP-dependent pathway for PLP.

Overexpression of AroB, by contrast, likely feeds into the 3-amino-1-hydroxyacetone-phosphate branch point. Finally, overexpressing YeaB or ThrB can convert 3-phosphohydroxypyruvate into 4-hydroxythreonine through a glycoaldehyde intermediate. The 4-hydroxythreonine pathway has also been shown to contribute flux for PLP biosynthesis in *S. enterica*, where an operon found in *S. enterica* but not *E. coli* contains *STM0162* and *pdxA2* (65). *STM0162* encodes a kinase that belongs to the DUF1537 family, and this enzyme can convert 4HT into 4-phosphohydroxy-threonine. Additionally, this pathway provides flux to PLP in the absence of *pdxA1*, and the addition of either 4HT or glycoaldehyde improves PLP biosynthesis under these conditions (65). In another study, it was found that ThiG has an active site structure that likely catalyzes the PLP synthase reaction when overexpressed and grown under high ammonia conditions (78).

Work in *S. enterica* derived from thiamine biosynthesis studies identified multiple routes to phosphoribosylamine (PRA), defined the role of a gene of unknown function (*yjgF*/*ridA*), and defined requirements for ThiC within the thiamine biosynthetic pathway (18, 19). Initially it was found that purines and thiamine share five biosynthetic steps, but an alternative route to PRA was identified when alleles of *purF* compromise purine but not thiamine synthesis (79). Later it was discovered that this PRA formation was due to flux to ribose-5-phosphate (R5P) during growth with ammonia as a nitrogen source, and when cells were grown in the absence of ammonia thiamine synthesis was compromised (80)0]. Accumulation of R5P through compromising either *trpC* or *prsA* can also lead to PRA formation through different metabolic intermediates (80, 81). Elimination of *trpC* leads to R5P accumulation from the degradation of N-(5-phosphoribosyl)-anthranilate, while compromised *prsA* diverts R5P from PRPP synthesis to PRA synthesis.

Research into new routes of PRA formation also found that lack of the gene of unknown function *yjgF* restored PRA synthesis in an IlvA-TrpD dependent manner (82, 83). Once YjgF/RidA was characterized to be an enamine deaminase, this pathway identified that a consequence of the lack of *ridA* in *S. enterica* led to the accumulation of 2-aminocrotanate which led to PRA formation from phosphoribosylpyrophosphate via TrpD (84, 85). It was later found that this pathway exists natively in *E. coli*, where the IlvA-TrpD pathway contributes on the order of 50 % of PRA for thiamine biosynthesis (9).

1.4 MODIFICATION OF THE *S. ENTERICA* METABOLIC NETWORK IS REQUIRED TO ALLOW *Sc*Thi5-DEPENDENT THIAMINE SYNTHESIS

Previous work toward understanding the requirements for Thi5 function *in vivo* found that *ScTHI5* could complement a *thiC* strain of *S. enterica* under some but not all growth conditions (Appendix A) (86). *ScTHI5* complemented on ribose, xylose and mannose as sole carbon sources. However, it failed to complement on carbon sources including glucose, galactose, gluconate and several other carbon sources that fed into glycolysis and the TCA cycle. The addition of histidine and/or PL failed to restore HMP biosynthesis on glucose, suggesting that another component was required for *Sc*Thi5 function under this growth condition. The context-dependent complementation raised questions about 1) the requirements for Thi5 function if the known precursors for HMP synthesis failed to allow complementation 2) differences in metabolic network structure between hosts and carbon source that allowed *Sc*Thi5 function in *S. enterica* under certain growth conditions.

To gain insight into these differences, suppressor mutations that restored Thi5-dependent thiamine synthesis were isolated on glucose. Two of these suppressors were in *sgrR*, which encodes a transcriptional regulator associated with the glucose phosphate stress response, and a previous student (Dr. Lauren Palmer) determined that these alleles encoded variants that were constitutively active. This led to expression of *sgrS*, which in turn led to a reduction in glucose transport via PtsG. Consistent with this model, a null mutation in *ptsG* allowed *Sc*Thi5 function. Through use of a β-galactosidase reporter assay, I determined that metabolic perturbations including deletions of genes encoding glycolytic enzymes *pgi* and *pfkA* led to constitutive expression of *sgrS* during growth on glucose, allowing *Sc*Thi5 function. Additionally, growth on glucose in the presence of 1 % alpha-methyl glucoside also allowed for Thi5 function and led to constitutive expression of *sgrS*. This suggests that the metabolic network of *S. enterica* needs to be modified to support function of this heterologous pathway, likely through reducing flux through glycolysis. The questions raised about potential differences in metabolic network structure between *S. enterica* and *S. cerevisiae*, and an interest in understanding the requirements for function of the Thi5 pathway led to the investigations detailed herein.

1.5 DISSERTATION OUTLINE

Based on previous work using both PLP and thiamine biosynthesis as model nodes of metabolism, this thesis dissects the putative redundancy within the PLP biosynthetic pathway in *S. cerevisiae*, describes functional differences between Thi5 orthologs of *Legionella pneumophila* and *S. cerevisiae*, and connects α-ketoglutarate to Thi5 function in a heterologous host. The work described in this thesis represents an important

contribution to understanding the distribution of the Thi5 pathway amongst bacteria and eukaryotes, and details how this pathway is integrated into the metabolic network in different organisms.

In Chapter 2, I investigate the function of the three PLP synthase paralogs of *S. cerevisiae* and dissect their contribution to thiamine synthesis. Each of the three paralogs of *SNZ* are required for PLP biosynthesis in *S. cerevisiae*. Either *SNZ2* or *SNZ3* are required for thiamine biosynthesis, and Snz3p is a PLP synthase with similar kinetic parameters as Snz1p. We also describe the contribution of *SNO3* to Snz3p function in a heterologous system.

In Chapter 3, I characterize a bacterial Thi5 from *Legionella pneumophila* and identify residues that modulate function *in vivo*. Bacterial Thi5 orthologs were identified in the Legionellaceae clade of γ-proteobacteria, and a *thi5* mutant of *L. pneumophila* is an HMP auxotroph. *Lpthi5* complements *S. enterica*, and when purified binds PLP without an exposed Schiff base. The protein when incubated with iron releases HMP *in vitro*, and several conserved residues are required for function. A conserved motif that is implicated in PLP binding modulates *in vivo* function of *Lp*Thi5.

In Chapter 4, I describe a connection between α-ketoglutarate and *Sc*Thi5 function in *S. enterica*. Two classes of suppressors that restore *Sc*Thi5 function are correlated with potential elevated levels of α-ketoglutarate and PLP. Modifying nitrogen assimilation partially restores *Sc*Thi5 function on multiple carbon sources, and the further addition of pyridoxal improves *Sc*Thi5-dependent thiamine synthesis. This leads to a model supported by nutritional supplementation that both elevated α-ketoglutarate and PLP levels are required for *Sc*Thi5 function in *S. enterica*.

In Chapter 5, I summarize the work presented in this thesis and discuss future directions for this research. Key extensions of this work include expanding knowledge of how Thi5 is integrated into central carbon metabolism, implications for its putative reconstitution *in vitro*, and other organisms in which to work to expand our understanding about the differences between ThiC and Thi5.

Appendix A describes the initial study to integrate the *Sc*Thi5 pathway for HMP-P biosynthesis into *S. enterica*. My contribution to this work included performing Western blots to detect *Sc*Thi5, quantifying *sgrS* expression with a β-galactosidase reporter in suppressor backgrounds and growth conditions, and determining which suppressors rely on *sgrR* for *Sc*Thi5 function.

1.6 REFERENCES

1. Noor E, Eden E, Milo R, Alon U. 2010. Central Carbon Metabolism as a Minimal Biochemical Walk between Precursors for Biomass and Energy. Mol Cell 39:809–820.

2. Chubukov V, Gerosa L, Kochanowski K, Sauer U. 2014. Coordination of microbial metabolism. Nat Rev Microbiol 12:327–340.

3. Sander T, Farke N, Diehl C, Kuntz M, Glatter T, Link H. 2019. Allosteric Feedback Inhibition Enables Robust Amino Acid Biosynthesis in E. coli by Enforcing Enzyme Overabundance. Cell Syst 8:66-75.e8.

4. Fonseca M V., Swanson MS. 2014. Nutrient salvaging and metabolism by the intracellular pathogen Legionella pneumophila. Front Cell Infect Microbiol 4:1–14.

5. Kotte O, Volkmer B, Radzikowski JL, Heinemann M. 2014. Phenotypic bistability in Escherichia coli's central carbon metabolism. Mol Syst Biol 10:736.

6. Litsios A, Ortega ÁD, Wit EC, Heinemann M. 2018. Metabolic-flux dependent regulation of microbial physiology. Curr Opin Microbiol 42:71–78.

7. Karr JR, Sanghvi JC, MacKlin DN, Gutschow M V., Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW. 2012. A whole-cell computational model predicts phenotype from genotype. Cell 150:389–401.

8. Yang H, Krumholz EW, Brutinel ED, Palani NP, Sadowsky MJ, Odlyzko AM, Gralnick JA, Libourel IGL. 2014. Genome-Scale Metabolic Network Validation of Shewanella oneidensis Using Transposon Insertion Frequency Analysis. PLoS Comput Biol 10.

9. Bazurto J V., Farley KR, Downs DM. 2016. An unexpected route to an essential cofactor: Escherichia coli relies on threonine for thiamine biosynthesis. MBio 7:1–9.

10. Bazurto J V., Downs DM. 2016. Metabolic network structure and function in bacteria goes beyond conserved enzyme components. Microb Cell 7:260–262.

11. Borchert AJ, Downs DM. 2017. The Response to 2-Aminoacrylate Differs in Escherichia coli and Salmonella enterica, despite Shared Metabolic Components. J Bacteriol 199:1–13.

12. Long CP, Antoniewicz MR. 2019. Metabolic flux responses to deletion of 20 core enzymes reveal flexibility and limits of E. coli metabolism. Metab Eng 55:249–257.

13. Cardinale S, Arkin AP. 2012. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. Biotechnol J 7:856–866.

14. Kittleson JT, Wu GC, Anderson JC. 2012. Successes and failures in modular genetic engineering. Curr Opin Chem Biol 16:329–336.

15. Close DM, Cooper CJ, Wang X, Chirania P, Gupta M, Ossyra JR, Giannone RJ, Engle N, Tschaplinski TJ, Smith JC, Hedstrom L, Parks JM, Michener JK. 2019. Horizontal transfer of a pathway for coumarate catabolism unexpectedly inhibits purine nucleotide biosynthesis. Mol Microbiol 112:1784–1797.

16. Michener JK, Vuilleumier S, Bringel F, Marx J. 2014. Phylogeny Poorly Predicts the Utility of a Challenging Horizontally Transferred Gene in Methylobacterium Strains. J Bacteriol 196:2101–2107.

17. Zhang JJ, Tang X, Zhang M, Nguyen D, Moore BS. 2017. Broad-host-range expression reveals native and host regulatory elements that influence heterologous antibiotic production in Gram-negative bacteria. MBio 8:1–16.

18. Downs DM. 2006. Understanding Microbial Metabolism. Annu Rev Microbiol 60:533–559.

19. Koenigsknecht MJ, Downs DM. 2010. Thiamine biosynthesis can be used to dissect metabolic integration. Trends Microbiol 18:240–247.

20. Jurgenson CT, Begley TP, Ealick SE. 2009. The structural and biochemical foundations of thiamin biosynthesis. Annu Rev Biochem 78:569–603.

21. Braun R. 2011. Accessory Food Factors: Understanding the Catalytic Function. J Hist Biol 44:483–504.

22. Gutowska MA, Shome B, Sudek S, McRose DL, Hamilton M, Giovannoni SJ, Begley TP, Worden AZ. 2017. Globally important haptophyte algae use exogenous pyrimidine compounds more efficiently than thiamin. MBio 8.

23. Kriek M, Martins F, Leonardi R, Fairhurst SA, Lowe DJ, Roach PL. 2007. Thiazole synthase from Escherichia coli: an investigation of the substrates and purified proteins required for activity in vitro. J Biol Chem 282:17413–17423.

24. Settembre EC, Dorrestein PC, Park J-H, Augustine AM, Begley TP, Ealick SE. 2003. Structural and mechanistic studies on ThiO, a glycine oxidase essential for thiamin biosynthesis in Bacillus subtilis. Biochemistry 42:2971–2981.

25. Lauhon CT, Kambampati R. 2000. The iscS gene in Escherichia coli is required for the biosynthesis of 4-thiouridine, thiamin, and NAD. J Biol Chem 275:20096–20103.

26. Taylor S V, Kelleher NL, Kinsland C, Chiu HJ, Costello CA, Backstrom AD, McLafferty FW, Begley TP. 1998. Thiamin biosynthesis in Escherichia coli. Identification of ThiS thiocarboxylate as the immediate sulfur donor in the thiazole formation. J Biol Chem 273:16555–16560.

27. Chatterjee A, Jurgenson CT, Schroeder FC, Ealick SE, Begley TP. 2006. Thiamin biosynthesis in eukaryotes: Characterization of the enzyme-bound product of thiazole synthase from Saccharomyces cerevisiae and its implications in thiazole

biosynthesis. J Am Chem Soc 128:7158–7159.

28. Chatterjee A, Abeydeera ND, Bale S, Pai P-J, Dorrestein PC, Russell DH, Ealick SE, Begley TP. 2011. Saccharomyces cerevisiae THI4p is a suicide thiamine thiazole synthase. Nature 478:542–6.

29. Ida T, Sawa T, Ihara H, Tsuchiya Y, Watanabe Y, Kumagai Y, Suematsu M, Motohashi H, Fujii S, Matsunaga T, Yamamoto M, Ono K, Devarie-Baez NO, Xian M, Fukuto JM, Akaike T. 2014. Reactive cysteine persulfides and S-polythiolation regulate oxidative stress and redox signaling. Proc Natl Acad Sci U S A 111:7606–7611.

30. Eser BE, Zhang X, Chanani PK, Begley TP, Ealick SE. 2016. From suicide enzyme to catalyst: the iron-dependent sulfide transfer in *Methanococcus jannaschii* thiamin thiazole biosynthesis. J Am Chem Soc jacs.6b00445.

31. Zhang X, Eser BE, Chanani PK, Begley TP, Ealick SE. 2016. Structural basis for iron-mediated sulfur transfer in archeal and yeast thiazole synthases. Biochemistry jacs.6b00445.

32. Chatterjee A, Hazra AB, Abdelwahed S, Hilmey DG, Begley TP. 2010. A Radical dance in thiamin biosynthesis: mechanistic analysis of the bacterial hydroxymethylpyrimidine phosphate synthase. Angew Chemie - Int Ed 49:8653–8656.

33. Lawhorn BG, Mehl R a, Begley TP. 2004. Biosynthesis of the thiamin pyrimidine: the reconstitution of a remarkable rearrangement reaction 2538–2546.

34. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE. 2001. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: Functional characterization using new analysis and information visualization methods. Nucleic Acids Res 29:1097–1106.

35. Galambas A, Miller J, Jones M, McDaniel E, Lukes M, Watts H, Copié V, Broderick JB, Szilagyi RK, Shepard EM. 2019. Radical S-adenosylmethionine maquette chemistry: Cx3Cx2C peptide coordinated redox active [4Fe–4S] clusters. J Biol Inorg Chem 24:793–807.

36. Dougherty MJ, Downs DM. 2006. A connection between iron-sulfur cluster metabolism and the biosynthesis of 4-amino-5-hydroxymethyl-2-methylpyrimidine pyrophosphate in Salmonella enterica. Microbiology 152:2345–2353.

37. Martinez-Gomez NC, Downs DM. 2008. ThiC is an [Fe-S] cluster protein that requires AdoMet to generate the 4-amino-5-hydroxymethyl-2-methylpyrimidine moiety in thiamin synthesis. Biochemistry 47:9054–9056.

38. Chatterjee A, Li Y, Zhang Y, Grove TL, Lee M, Krebs C, Booker SJ, Begley TP, Ealick SE. 2008. Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. Nat Chem Biol 4:758–765.

39. Martinez-Gomez NC, Poyner RR, Mansoorabadi SO, Reed GH, Downs DM. 2009. Reaction of AdoMet with ThiC generates a backbone free radical. Biochemistry 48:217–219.

40. Palmer LD, Downs DM. 2013. The thiamine biosynthetic enzyme ThiC catalyzes multiple turnovers and is inhibited by S-adenosylmethionine (AdoMet) metabolites. J Biol Chem 288:30693–9.

41. Fenwick MK, Mehta AP, Zhang Y, Abdelwahed SH, Begley TP, Ealick SE. 2015. Non-canonical active site architecture of the radical SAM thiamin pyrimidine synthase. Nat Commun 6:6480.

42. Ishida S, Tazuya-Murayama K, Kijima Y, Yamada K. 2008. The direct precursor of the pyrimidine moiety of thiamin is not urocanic acid but histidine in Saccharomyces cerevisiae. J Nutr Sci Vitaminol (Tokyo) 54:7–10.

43. Tazuya K, Yamada K, Kumaoka H. 1989. Incorporation of histidine into the pyrimidine moiety of thiamin in Saccharomyces cerevisiae. Biochim Biophys Acta - Gen Subj 990:73–79.

44. Tazuya K, Azumi C, Yamada K, Kumaoka H. 1995. Pyrimidine moiety of thiamin is biosynthesized from pyridoxine and histidine in Saccharomyces cerevisiae. Biochem Mol Biol Int 36:883–888.

45. Wightman R, Meacock P a. 2003. The THI5 gene family of Saccharomyces cerevisiae: Distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine. Microbiology 149:1447–1460.

46. Maundrell K. 1990. nmt1 of Fission Yeast. J Biol Chem 265:10857–10864.

47. Schweingruber A, Dlugonski J, Edenharter E, Schweingruber M. 1991. Thiamine in Schizosaccharomyces pombe: dephosphorylation, intracellular pool, biosynthesis and transport. Curr Genet 22:153.

48. Lai RY, Huang S, Fenwick MK, Hazra A, Zhang Y, Rajashankar K, Philmus B, Kinsland C, Sanders JM, Ealick SE, Begley TP. 2012. Thiamin pyrimidine biosynthesis in candida albicans: A remarkable reaction between histidine and pyridoxal phosphate. J Am Chem Soc 134:9157–9159.

49. Coquille S, Roux C, Fitzpatrick TB, Thore S. 2012. The last piece in the vitamin B1 biosynthesis puzzle: Structural and functional insight into yeast 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase. J Biol Chem 287:42333–42343.

50. Bale S, Rajashankar KR, Perry K, Begley TP, Ealick SE. 2010. HMP binding protein ThiY and HMP-P synthase THI5 are structural homologues. Biochemistry 49:8929–8936.

51. Paxhia MD, Swanson MS, Downs DM. 2020. Functional characterization of the HMP-P synthase of Legionella pneumophila (Lpg1565). Mol Microbiol 5:1–15.

52. Percudani R, Peracchi A. 2003. A genomic overview of pyridoxal-phosphate-dependent enzymes. EMBO Rep 4:850–854.

53. Percudani R, 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.

54. Soo VWC, Yosaatmadja Y, Squire CJ, Patrick WM. 2016. Mechanistic and evolutionary insights from the reciprocal promiscuity of two pyridoxal phosphate-dependent enzymes. J Biol Chem 291:19873–19887.

55. Ferla MP, Brewster JL, Hall KR, Evans GB, Patrick WM. 2017. Primordial-like enzymes from bacteria with reduced genomes. Mol Microbiol 105:508–524.

56. Toney MD. 2011. Controlling reaction specificity in pyridoxal phosphate enzymes. Biochim Biophys Acta - Proteins Proteomics 1814:1407–1418.

57. Tanaka T, Tateno Y, Gojobori T. 2005. Evolution of vitamin B6 (pyridoxine) metabolism by gain and loss of genes. Mol Biol Evol 22:243–250.

58. Sprenger GA, Schörken U, Wiegert T, Grolle S, De Graaf AA, Taylor S V., Begley TP, Bringer-Meyer S, Sahm H. 1997. Identification of a thiamin-dependent synthase in Escherichia coli required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. Proc Natl Acad Sci U S A 94:12857–12862.

59. Zhao G, Pease AJ, Bharani N, Winkler ME. 1995. Biochemical characterization of gapB-encoded erythrose 4-phosphate dehydrogenase of Escherichia coli K-12 and its possible role in pyridoxal 5'- phosphate biosynthesis. J Bacteriol 177:2804–2812.

60. Drewke C, Klein M, Clade D, Arenz A, Mtiller R, Leistner E. 1996. 4-O-Phosphoryl-l-threonine, a substrate of the pdxC(serC) gene product involved in vitamin B6 biosynthesis. FEBS Lett 390:179–182.

61. Laber B, Maurer W, Scharf S, Stepusin K, Schmidt FS. 1999. Vitamin B6 biosynthesis: Formation of pyridoxine 5′-phosphate from 4-(phosphohydroxy)-L-threonine and I-deoxy-D-xylulose-5-phosphate by PdxA and PdxJ protein. FEBS Lett 12:920–924.

62. Banks J, Cane DE. 2004. Biosynthesis of vitamin B6: Direct identification of the product of the PdxA-catalyzed oxidation of 4-hydroxy-L-threonine-4-phosphate using electrospray ionization mass spectrometry. Bioorganic Med Chem Lett 14:1633–1636.

63. Di Salvo M, Yang E, Zhao G, Winkler ME, Schirch V. 1998. Expression, purification, and characterization of recombinant Escherichia coil pyridoxine 5'-phosphate oxidase. Protein Expr Purif 13:349–356.

64. Kim J, Kershner JP, Novikov Y, Shoemaker RK, Copley SD. 2010. Three serendipitous pathways in E. coli can bypass a block in pyridoxal-5'-phosphate synthesis. Mol Syst Biol 6:1–13.

65. Thiaville JJ, Flood J, Yurgel S, Prunetti L, Elbadawi-Sidhu M, Hutinet G, Forouhar F, Zhang X, Ganesan V, Reddy P, Fiehn O, Gerlt JA, Hunt JF, Copley SD, de Crécy-Lagard V. 2016. Members of a Novel Kinase Family (DUF1537) Can Recycle Toxic Intermediates into an Essential Metabolite. ACS Chem Biol acschembio.6b00279.

66. Raschle T, Amrhein N, Fitzpatrick TB. 2005. On the two components of pyridoxal 5'-phosphate synthase from Bacillus subtilis. J Biol Chem 280:32291–32300.

67. Burns KE, Xiang Y, Kinsland CL, McLafferty FW, Begley TP. 2005. Reconstitution and biochemical characterization of a new pyridoxal-5'-phosphate biosynthetic pathway. J Am Chem Soc 127:3682–3683.

68. Strohmeier M, Raschle T, Mazurkiewicz J, Rippe K, Sinning I, Fitzpatrick TB, Tews I. 2006. Structure of a bacterial pyridoxal 5'-phosphate synthase complex. Proc Natl Acad Sci U S A 103:19284–19289.

69. Dong YX, Sueda S, Nikawa JI, Kondo H. 2004. Characterization of the products of the genes SNO1 and SNZ1 involved in pyridoxine synthesis in Saccharomyces cerevisiae. Eur J Biochem 271:745–752.

70. Zein F, Zhang Y, Kang YN, Burns K, Begley TP, Ealick SE. 2006. Structural insights into the mechanism of the PLP synthase holoenzyme from Thermotoga maritima. Biochemistry 45:14609–14620.

71. Hanes JW, Keresztes I, Begley TP. 2008. 13C NMR snapshots of the complex reaction coordinate of pyridoxal phosphate synthase. Nat Chem Biol 4:425–30.

72. Neuwirth M, Strohmeier M, Windeisen V, Wallner S, Deller S, Rippe K, Sinning I, Macheroux P, Tews I. 2009. X-ray crystal structure of Saccharomyces cerevisiae Pdx1 provides insights into the oligomeric nature of PLP synthases. FEBS Lett 583:2179–2186.

73. Moccand C, Kaufmann M, Fitzpatrick TB. 2011. It takes two to Tango: Defining an essential second active site in pyridoxal 5'-phosphate synthase. PLoS One 6.

74. Robinson GC, Kaufmann M, Roux C, Fitzpatrick TB. 2016. Structural definition of the lysine swing in Arabidopsis thaliana PDX1: Intermediate channeling facilitating vitamin B6 biosynthesis. Proc Natl Acad Sci 113:E5821–E5829.

75. Rodrigues MJ, Windeisen V, Zhang Y, Guédez G, Weber S, Strohmeier M, Hanes JW, Royant A, Evans G, Sinning I, Ealick SE, Begley TP, Tews I. 2017. Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis. Nat Chem Biol 6–11.

76. Raschle T, Arigoni D, Brunisholz R, Rechsteiner H, Amrhein N, Fitzpatrick TB. 2007. Reaction mechanism of pyridoxal 5′-phosphate synthase: Detection of an enzyme-bound chromophoric intermediate. J Biol Chem 282:6098–6105.

77. Lin S, Hanson RE, Cronan JE. 2010. Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nat Chem Biol 6:682–8.

78. Oberhardt MA, Zarecki R, Reshef L, Xia F, Duran-Frigola M, Schreiber R, Henry CS, Ben-Tal N, Dwyer DJ, Gophna U, Ruppin E. 2016. Systems-Wide Prediction of Enzyme Promiscuity Reveals a New Underground Alternative Route for Pyridoxal 5'-Phosphate Production in E. coli. PLoS Comput Biol 12:1–19.

79. Downs DM, Roth JR. 1991. Synthesis of thiamine in Salmonella typhimurium independent of the purF function. J Bacteriol 173:6597–6604.

80. Koenigsknecht MJ, Fenlon LA, Downs DM. 2010. Phosphoribosylpyrophosphate synthetase (PrsA) variants alter cellular pools of ribose 5-phosphate and influence thiamine synthesis in Salmonella enterica. Microbiology 156:950–959.

81. Ramos I, Vivas EI, Downs DM. 2008. Mutations in the tryptophan operon allow purF-independent thiamine synthesis by altering flux in vivo. J Bacteriol 190:815–822.

82. Enos-Berlage JL, Langendorf MJ, Downs DM. 1998. Complex metabolic phenotypes caused by a mutation in yjgF, encoding a member of the highly conserved YER057c/YjgF family of proteins. J Bacteriol 180:6519–6528.

83. Browne BA, Ramos AI, Downs DM. 2006. PurF-independent phosphoribosyl amine formation in yjgF Mutants of Salmonella enterica utilizes the tryptophan biosynthetic enzyme complex anthranilate synthase-phosphoribosyltransferase. J Bacteriol 188:6786–6792.

84. Lambrecht JA, Browne BA, Downs DM. 2010. Members of the YjgF/YER057c/UK114 family of proteins inhibit phosphoribosylamine synthesis in vitro. J Biol Chem 285:34401–34407.

85. Lambrecht JA, Downs DM. 2013. Anthranilate phosphoribosyl transferase (TrpD) generates phosphoribosylamine for thiamine synthesis from enamines and phosphoribosyl pyrophosphate. ACS Chem Biol 8:242–248.

86. Palmer LD, Paxhia MD, Downs DM. 2015. Induction of the Sugar-Phosphate Stress Response Allows Saccharomyces cerevisiae 2-Methyl-4-Amino-5-Hydroxymethylpyrimidine Phosphate Synthase To Function in Salmonella enterica. J Bacteriol 197:3554–3562.

CHAPTER 2

SNZ3 ENCODES A PLP SYNTHASE INVOLVED IN THIAMINE SYNTHESIS IN *SACCHAROMYCES CEREVISIAE*[1]

[1] Paxhia MD, Downs DM. 2019. *G3: Genes, Genomes, Genetics (Bethesda)*. 9(2):335-344.
Reprinted here with permission of the publisher

2.1 ABSTRACT

Pyridoxal 5'-phosphate (the active form of vitamin B6) is a cofactor that is important for a broad number of biochemical reactions and is essential for all forms of life. Organisms that can synthesize pyridoxal 5'-phosphate use either the deoxyxylulose phosphate-dependent or -independent pathway, the latter is encoded by a two-component pyridoxal 5'-phosphate synthase. *Saccharomyces cerevisiae* contains three paralogs of the two-component *SNZ/SNO* pyridoxal 5'-phosphate synthase. Past work identified the biochemical activity of Snz1p, Sno1p and provided in vivo data that *SNZ1* was involved in pyridoxal 5'-phosphate biosynthesis. Snz2p and Snz3p were considered redundant isozymes and no growth condition requiring their activity was reported. Genetic data herein showed that either *SNZ2* or *SNZ3* are required for efficient thiamine biosynthesis in *Saccharomyces cerevisiae*. Further, *SNZ2* or *SNZ3* alone could satisfy the cellular requirement for pyridoxal 5'-phosphate (and thiamine), while *SNZ1* was sufficient for pyridoxal 5'-phosphate synthesis only if thiamine was provided. qRT-PCR analysis determined that *SNZ2,3* are repressed ten-fold by the presence thiamine. In total, the data were consistent with a requirement for PLP in thiamine synthesis, perhaps in the Thi5p enzyme, that could only be satisfied by *SNZ2* or *SNZ3*. Additional data showed that Snz3p is a pyridoxal 5'-phosphate synthase *in vitro* and is sufficient to satisfy the pyridoxal 5'-phosphate requirement in *Salmonella enterica* when the medium has excess ammonia.

2.2 INTRODUCTION

Pyridoxal-5'-phosphate (PLP, the active form of vitamin B_6) is an essential cofactor that is used for diverse reactions including α/β eliminations, retro-aldol cleavages,

transaminations and racemizations [1]. Two pathways for PLP biosynthesis have been described. The 1-deoxy-D-xylulose-5-phosphate (DXP)-independent pathway is found in most bacteria, archaea and eukaryotes, while a DXP-dependent pathway is found in some bacteria including the proteobacteria, firmicutes, chlorobi, cyanobacteria and aquificae [2-4]. The DXP-dependent pathway synthesizes PLP from erythrose-4-phosphate, glyceraldehyde-3-phosphate and pyruvate in a series of seven enzymatic steps. Formation of the pyridine heterocyclic ring in this pathway is catalyzed by the pyridoxine-5'-phosphate synthase (E.C. 2.6.99.2), which is encoded by *pdxJ* in *Escherichia coli* (Figure 1). In contrast, the DXP-independent pathway uses two enzymes to create the heterocyclic pyridine ring of PLP from glutamine, glyceraldehyde-3-phosphate and ribose-5-phosphate [4] (Figure 2.1). Salvage of B_6 vitamers requires enzymes that are conserved across organisms. Notably, pyridoxine phosphate oxidase (PNPO, PdxH; E.C.:1.4.3.5) is required for *de novo* synthesis in the DXP-dependent pathway and is required for salvage in organisms using either pathway.

The two enzymes unique to the DXP-independent pathway to PLP form a complex comprised of a glutaminase and a PLP synthase subunit (EC 4.3.3.6) [5-7]. In general, the glutaminase subunit liberates ammonia from glutamine and delivers it to the PLP synthase subunit, where it combines with glyceraldehyde-3-phosphate and ribose-5-phosphate to form PLP [5-7]. Similar ammonia tunneling is a feature of several multi-subunit synthase enzymes that use glutamine as a source of ammonia including, carbamoyl-phosphate synthetase (E.C. 6.3.5.5), anthranilate synthase (E.C. 4.1.3.27), aminodeoxychorismate synthase (E.C. 2.6.1.85) and imidazole glycerol phosphate synthase (E.C. 4.3.2.-) [8-11].

In cases where it has been tested, the glutaminase subunit is dispensable both *in vitro* and *in vivo* when there are high levels of ammonia [12, 13].

Several PLP synthase enzymes from bacteria, archaea, yeast and plants, have been characterized biochemically with and without the associated glutaminase [5-7, 14-23]. A variety of names have been used for the genes encoding PLP synthase and glutaminase enzymes, e.g., *SNZ/pdxS/pdx1, SNO/pdxT/pdx2*, respectively. The varied nomenclature has complicated the analyses and comparison of these enzymes across organisms. For simplicity the names for the glutaminase and synthase proteins coined in yeast, Sno and Snz, respectively are used herein. Organisms can encode multiple paralogs of *SNZ* and *SNO*, and in most cases the value of the redundancy remains unclear [24-30]. An exception is *Arabidopsis thaliana,* where each of the three isozymes of Snz have been characterized *in vivo* and *in vitro* [26, 28]. Two of the isozymes (Pdx1.1, Pdx1.3) had PLP synthase activity *in vitro,* and either was sufficient for PLP synthesis *in vivo* [28]. The third isozyme (Pdx1.2) had weak homology, lacked key residues involved in catalysis, and failed to generate a detectable phenotype when absent [28].

Saccharomyces cerevisiae has three glutaminase paralogs (*SNO1-SNO3*) and three PLP synthase paralogs (*SNZ1-SNZ3*) [24]. *SNZ1* was required for growth in the absence of exogenous pyridoxine, but the need for *SNO1 in vivo* was less clear [24, 31]. Snz1p is a PLP synthase that uses glyceraldehyde-3-phosphate, ribose-5-phosphate and free ammonia as substrates *in vitro* [19], and Sno1p has glutaminase activity *in vitro* [14]. The inability of an *snz1* mutant to grow on dropout medium lacking PN was attributed to the presence of thiamine. The *SNZ2/3* genes were regulated by thiamine, in a manner completely dependent on *THI2* and partially dependent *on THI3* [24]. The pattern of regulation

suggested *SNZ2/3* could have a role in thiamine synthesis and/or metabolism, although to our knowledge this idea was not directly tested experimentally.

A long-time interest in metabolic network structure and robustness, and redundancy in vitamin biosynthesis, with a focus on thiamine (reviewed in [32, 33]), prompted us to explore the functional roles of the Sno/Snz proteins in *S. cerevisiae.* Herein we assimilate the genetic and biochemical characterization of the *SNZ* paralogs to lay the groundwork for future work on the integration of the biosynthesis of two essential cofactors, thiamine and PLP, in *Salmonella enterica* and *S. cerevisiae*. The data presented confirmed that *SNZ3* encodes a PLP synthase, quantified the transcriptional regulation of *SNZ2* and *SNZ3* by thiamine, and showed that strains lacking *SNZ2* and *SNZ3* required thiamine supplementation for growth. Together these data provide the first report of a functional role for these genes *in vivo.*

2.3 MATERIALS AND METHODS

Strains, Media and Chemicals

Yeast. S. cerevisiae strains used in this work were derived from YJF153 (*MATa HO::dsdAMX4*, a haploid derivative of YPS163) [34], and the relevant genotypes are listed in Table 2.1. *S. cerevisiae* strains were routinely grown on rich medium containing 10 g/L yeast extract, 20 g/L peptone, 20 g/L dextrose and 20 g/L agar (YPD). Two variations on defined medium were used to monitor vitamin requirements. Synthetic defined media (SD, SG) contained 1.7 g/L yeast nitrogen base without amino acids or nitrogen (YNB, Sunrise Science catalog no. 1500-100) or the respective drop-out as indicated (YNB-Pyridoxine, YNB-Thiamine, YNB-Pyridoxine-Thiamine) (Sunrise Science), 5 g/L ammonium sulfate,

20 g/L agar and either 20 g/L dextrose or 30 g/L glycerol as carbon source. A minimal medium (Minimal Vitamin Dextrose; MVD) contained 1.7 g/L YNB-Vitamins (Sunrise Science), biotin (0.002 mg/L), and D-pantothenic acid hemicalcium salt (0.4 mg/L), 5 g/L ammonium sulfate, 20 g/L agar and 20 g/L dextrose as carbon source. Thiamine (0.4 mg/L) and/or pyridoxine (0.4 mg/L) were added as indicated. Antibiotics used for deletion marker selection were added to the following final concentrations in YPD: 400 mg/L geneticin (G-418 sulfate), 200 mg/L Hygromycin B, and 100 mg/L nourseothricin sulfate (clonNAT) (Gold Biotechnology). A lower concentration of 200 μg/mL geneticin was used for maintenance of strains with G-418 resistance. 2-methyl-4-amino-5-hydroxymethylpyrimidine (HMP) was purchased from LabSeeker, Inc.

Bacteria. Media for bacterial growth were Nutrient Broth (NB) containing 8 g/L Difco Nutrient broth and 5 g/L NaCl, lysogeny broth (LB), or superbroth (SB; 32 g/L tryptone (Fisher Scientific), 20 g/L yeast extract (Fisher Scientific), 5 g/L NaCl with 0.05 N NaOH). Solid media contained 1.5 % agar. Antibiotics were added at the following concentrations in rich media, unless otherwise indicated: kanamycin (Kn), 50 mg/L; chloramphenicol (Cm), 20 mg/L; ampicillin (Ap), 100 mg/L. Minimal media was no-carbon E medium (NCE) [35] with 1 mM $MgSO_4$, 0.1x trace minerals [36], with either glucose (11 mM) or glycerol (22 mM) (Fisher Scientific) as a sole carbon source. Minimal medium with low nitrogen was no-carbon and nitrogen (NCN) [37] with 1 mM glutamine and glucose or glycerol as sole carbon source. All strains of *S. enterica* are derived from strain LT2 and their relevant genotypes are described in Table 2.1. Chemicals were purchased from Sigma-Aldrich, St. Louis, MO unless otherwise indicated.

Genetic Techniques

In-frame deletions of genes in *S. enterica* were created with Lambda-Red recombineering as described [38]. Insertions were reconstructed by transduction into DM7080 (*araCBAD*) with the high-frequency generalized transducing mutant of bacteriophage P22 (HT105/1, int-201) [39]. Primers used to generate these deletions are listed in Table 2.1.

Gene disruptions in *S. cerevisiae* were made using a described gene replacement method [40]. Antibiotic cassettes were amplified from the appropriate plasmid using primers listed in Table 2.1. Five µg of purified DNA was transformed into *S. cerevisiae* by incubating cells suspended in a mixture of 33 % polyethylene glycol 3350 (PEG 3350), 100 mM lithium acetate, and 0.28 mg/mL salmon sperm DNA at 42 °C for 90 minutes. The transformed cells were recovered in YPD for 3 hours with shaking at 30 °C and plated to YPD with the appropriate antibiotic. Colonies that grew on the plates after three days were streaked onto selective media and insertions were confirmed by colony PCR. Insertions in *SNZ2* and *SNZ3* were distinguished by PCR using gDNA as a template and AGP3R, RPD3R, and CM primers listed in Table 2.1.

Molecular Techniques

Plasmids were constructed using standard molecular techniques. Plasmid DNA was isolated using the PureYield Plasmid MiniPrep System (Promega, Madison, WI). Q5 DNA polymerase (New England Biolabs, Ipswich, MA) was used to amplify DNA with primers synthesized by Integrated DNA Technologies, Coralville, IA or Eton Bioscience, Inc., Research Triangle Park, NC. PCR products were purified using the PCR purification kit

(Qiagen, Venlo, Limburg, The Netherlands). Restriction endonucleases were purchased from New England Biolabs, Ipswich, MA, and ligase was purchased from ThermoScientific, Waltham, MA.

Growth Analysis

Bacteria. Growth of *S. enterica* strains were monitored at OD_{650} in 96 well plates with a BioTek ELx808 plate reader. Strains were grown overnight in NB with Cm or Ap as indicated and inoculated at 1 % into 200 µL of media indicated. Plates were incubated at 37 °C with medium shaking and data were plotted using Prism 7 (GraphPad).

Yeast. Growth of *S. cerevisiae* strains was followed by dilution plating or liquid growth. Liquid growth was monitored at OD_{650} in 96 well plates with a BioTek ELx808 plate reader. Strains were grown for 24 hours in SD before pelleting and resuspending twice in saline. Washed cells were inoculated in 100 µL of the appropriate medium at 1 %. Plates were incubated at 30 °C with fast shaking and data were plotted using Prism 7 (GraphPad). To monitor growth via dilution plating, *S. cerevisiae* strains were grown for 24 hours in YPD before pelleting, resuspending in saline, and 5 µL of serial dilutions in saline from 10^{-2} to 10^{-7} were plated onto the respective media. Plates were incubated at 30 °C for two or three days, as indicated.

Reverse transcription-quantitative-PCR

RNA from three biological replicates of YJF153 was prepared and extracted as follows. Independent cultures of YJF153 were grown for 24 hours in SD with shaking at 30 °C

before pelleting and resuspending twice in saline. Washed cells were used as a 1 % inoculum into 5 mL MVD with and without thiamine and/or PN as indicated and grown at 30 °C with shaking for 12 hours. Cells were pelleted at 10,000×g for 15 seconds and frozen in liquid nitrogen. RNA was extracted using the RNA*snap*™ method modified for yeast [41]. Briefly, 100 µL of glass beads (425-600 µm) and 110 µL of RNA extraction solution (95 % molecular biology grade formamide, 0.025 % SDS, 18 mM EDTA and 1 % β-mercaptoethanol) were added to the cell pellets and vortexed with a bead-beating adaptor for five minutes. Tubes were then incubated at 95 °C for 7 minutes and centrifuged at 16,000×g for 5 minutes. The supernatant was transferred into another tube and RNA was concentrated by sodium acetate/ethanol precipitation, treated with RNase-free Turbo DNase (Ambion), precipitated by sodium acetate/ethanol precipitation and stored at -80 °C.

Quality and concentration of total RNA was evaluated at the Georgia Genomics and Bioinformatics Core (GGBC) using the RNA nano 6000 kit for the Agilent Bioanalyzer 2100. cDNA was generated from 800 ng of total RNA using the iScript cDNA synthesis kit (Bio-Rad Laboratories) by following the manufacturer's protocol. Real-time PCR reactions (20 µL) were prepared with 10 µL Fast SYBR™ Green Master Mix (Applied Biosystems), 8 ng cDNA and 500 nM gene-specific primers (Table 2.1). Real-time PCR was performed using the Applied Biosystems 7500 Fast real-time PCR system. Expression of *SNZ2* and *SNZ3* were treated as one transcriptional response, due to their 99 % identity at the nucleotide level. Relative expression of *SNZ2/3* was calculated using the comparative cycle threshold method ($\Delta\Delta C_T$) with *UBC6* as an internal control and fold change (treated/untreated) was calculated with the equation $2^{-\Delta\Delta C_T}$ [42, 43]. The standard error of

the mean (SEM) was calculated for the test condition ΔC_T, using Guassian error propagation, and this was used to calculate the 95 % confidence intervals for each $\Delta\Delta C_T$ calculation. To ensure the effectiveness of using *UBC6* as an internal standard, *ALG9* was used as an alternative internal control [42]. Under the conditions tested, no significant differences in expression were observed for either internal control.

Protein Purification

SNZ3 was cloned into pTEV5 [44] at the NheI/NcoI sites, the plasmid was purified and transformed into *E. coli* BL21-AI. The resulting strain was grown overnight at 37 °C in 100 mL NB Ap and inoculated into six liters of SB Ap (1 %), and grown at 37 °C with shaking (200 rpm) to an OD_{650} of 0.6. The temperature was lowered to 30 °C, arabinose added to a final concentration of 0.2 % and cells were incubated for 19 hours prior to harvesting by centrifugation. The cell pellet was resuspended in Buffer A (50 mM HEPES, 300 mM NaCl, 20 mM Imidizole, pH 7.5 at 4 °C) with DNAse (0.025 mg/mL), lysozyme (1 mg/mL) and phenylmethylsulfonyl fluoride (0.1 mg/mL) and kept on ice for one hour. The cell suspension was lysed at 20 kpsi using a Constant Systems Limited One Shot (United Kingdom), and cell lysate was cleared at 48,000 xg for 50 minutes at 4 °C. The cell-free extract was passed through a 0.45 µM PVDF filter (Millipore) and injected onto a pre-equilibrated 5 mL HisTrap HP Ni-sepharose column. The column was washed with 5 column volumes of Buffer A, followed by 5 column volumes of 4 % Buffer B (50 mM HEPES, 300 mM NaCl, 500 mM Imidazole, pH 7.5 at 4 °C) and finally a gradient of Buffer B from 4 % to 100 % over 10 column volumes. Fractions containing Snz3p were combined, rTEV protease was added at a 50:1 protein to rTEV ratio and the mixture sat for 3.5 hrs at

room temperature before it was dialyzed into Buffer A overnight at 4 °C with 3 buffer changes. The tagless protein was separated from His_6-Snz3p and His_6-rTEV by gravity column chromatography with HisPur Ni-NTA resin. Snz3p was concentrated by centrifugation using a 10 kDa filter (Millipore), exchanged into a 50 mM HEPES buffer, pH 7.5, with 10 % glycerol using a PD10 column (GE Healthcare), flash-frozen in liquid nitrogen and stored at -80 °C until use. Protein concentration was determined by bicinchoninic acid (BCA) assay (Pierce) with bovine serum albumin as a standard. The Snz3p preparation was > 98 % pure based on densitometery.

PLP synthase assay

Snz3p was thawed and dialyzed into assay buffer (50 mM Tris-HCl, pH 8.0). Reactions were performed in a buffer of 50 mM Tris-HCl, pH 8.0 at 37 °C with ammonium sulfate (10 mM), D/L glyceraldehyde-3-phosphate, ribose-5-phosphate and Snz3p [6, 19]. Assays were performed in triplicate with 85 μM Snz3p and the formation of PLP was followed spectrophotometrically at 414 nm using Spectramax 398-Plus plate reader. The extinction coefficient for PLP in assay buffer was determined to be 7.57×10^3 M^{-1} cm^{-1} at 414 nm with a five-point standard curve measured in duplicate from 0 – 0.1 M PLP ($R^2 = 0.9996$). Kinetic parameters for Snz3p with D/L glyceraldehyde-3-phosphate as a substrate were determined with 1 mM ribose-5-phosphate and reactions were initiated with concentrations of D/L glyceraldehyde-3-phosphate from 62.5 μM to 2 mM. Activity of Snz3p with ribose-5-phosphate as a substrate was determined with 2 mM D/L glyceraldehyde-3-phosphate and reactions were initiated with concentrations of ribose-5-phosphate from 25 μM to 0.8 mM. Data were plotted and analyzed using Prism 7 (Graph Pad).

2.4 RESULTS AND DISCUSSION

SNZ1, 2, 3 paralogs have distinguishable roles *in vivo*

Mutants of *S. cerevisiae* YJF153 lacking one or more of the *SNZ* paralogs were constructed to evaluate the role of these proteins *in vivo* (Table 2.1). The YJF153 strain of *S. cerevisiae* was chosen for use because it i) has a wild type allele of the transcriptional regulator *THI3*, and ii) has no auxotrophy. The former point is relevant in that the status of *THI3* impacts expression of genes in the thiamine regulon, which includes *SNZ2* and *SNZ3,* due to its activity as a co-activator with Thi2p [24, 45-47]. The standard laboratory strain of *S. cerevisiae* (S288C) has a mutant allele of *THI3* which results in abhorrent expression of the thiamine regulon [47]. The lack of any auxotrophy indicated the strain had the functional metabolic network needed to dissect metabolic interactions and detecting often subtle connections between pathways.

Eight strains (single, double and triple mutants) were constructed to query the role of the *SNZ* paralogs. Standard yeast drop-out media with glucose or glycerol were used and the data for seven of the mutants are shown in Figure 2.2. Growth of an *snz3* strain was indistinguishable from the *snz2* strain under these conditions (data not shown). Consistent with previous observations, *SNZ1* was required for growth in dextrose medium lacking only pyridoxine (SD-PN). However, when thiamine was also excluded (SD-PN-Thiamine), a single functional copy of any of the *SNZ* paralogs allowed growth. These data were consistent with a model where each PLP synthase had the capacity to generate sufficient PLP for growth, but the regulation of *SNZ2,3* by thiamine prevented them from contributing to PLP synthesis in its presence. The expression of *SNZ2* or *SNZ3 in trans* had

been shown to complement an *snz1* mutant, but this is the first demonstration that at chromosomal levels, either *SNZ2* or *SNZ3* was sufficient for PLP synthesis that allowed optimal growth.

A similar analysis was done on media with glycerol rather than dextrose to compare fermentative vs. respiratory lifestyles (Figure 2.2B). The data were generally similar, with two differences noted. First, deletion of *SNZ1* severely decreased, but did not eliminate growth on SG-PN medium. These data suggested that either the PLP requirement was lower during glycerol respiration, or the repression of *SNZ2,3* in the presence of thiamine was weaker on glycerol. Secondly, we noted that the *snz2,3* double mutant grew poorly on SD-PN-Thiamine, while it showed robust growth on SG-PN-thiamine, suggesting *SNZ1* was not always sufficient for PLP synthesis.

Minimal vitamin medium clarifies the role of *SNZ2* and *SNZ3*.

It was formally possible that additional nutrients in the drop out medium were complicating the interpretation of the nutritional phenotypes due to unanticipated metabolic interactions or regulation. The results above were readdressed using medium with supplementation, rather than drop-out. Control experiments showed that *S. cerevisiae* strain YJF153 grew well on synthetic dextrose medium with YNB-vitamins if both biotin and pantothenate provided. This medium was designated as Minimal Vitamin Dextrose medium (MVD) and used in the subsequent experiments. The growth of the eight strains described above was quantified in liquid MVD, with pyridoxine and/or thiamine added exogenously (Figure 2.3). Several points were taken from the resulting data. The parental wild type strain (YJF153), DMy53 (*snz2*), and DMy57 (*snz3*) strains had full growth on

each medium (Figure 2.3A,B,C), while the triple mutant (DMy52) failed to grow in the absence of PN (Figure 2.3D). The other three strains that lacked *SNZ1* grew in MVD, supporting the conclusion that either isozyme was sufficient for PLP synthesis in the absence of repression by thiamine. However, the strains that depended on *SNZ2* and/or *SNZ3* for PLP synthesis, failed to grow if thiamine was present in the medium (Figure 2.3 E,F,G). To verify the explanation that the lack of growth was due to transcriptional repression caused by thiamine, transcript levels of *SNZ2* and *SNZ3* were determined by qRT-PCR. The relative expression level of *SNZ2/3* in YJF153 grown in MVD compared to MVD containing thiamine, PN, or both was measured. The data in Figure 2.4 showed that the addition of thiamine repressed transcription of *SNZ2/3* approximately 10-fold. The presence of PN did not affect this repression and had no detectable effect by itself. These data validated the model that the conditional auxotrophy of strains lacking *SNZ1* in the presence of thiamine is due to the transcriptional repression of *SNZ2* and *SNZ3*.

The data in Figure 2.3 shed new light on the role and limitation of *SNZ1* function that was only hinted at by the data in Figure 2.2. The *snz2,3* double mutant failed to grow on MVD medium (Figure 2.3H). These data showed the PLP synthase encoded by *SNZ1* was not able to provide sufficient PLP synthesis for growth. Growth was restored by the addition of either thiamine or pyridoxine. The growth behavior of the *snz2,3* double mutant was consistent with a scenario in which thiamine synthesis required PLP and Snz1p alone could not synthesize sufficient PLP to satisfy this requirement. In fact, in *S. cerevisiae*, synthesis of the 2-methyl-4-amino-5-hydroxymethylpyrimidine (HMP) moiety of thiamine involves use of PLP as a substrate by a poorly characterized HMP-P synthase enzyme, Thi5p [48]. Consistently, exogenous addition of HMP (but not the thiazole moiety) restored

growth to *snz2,3* double mutant. Together, these data suggest the synthesis of thiamine, potentially via Thi5p, has unique PLP requirements which could help explain the presence of multiple *SNZ/SNO* paralogs in *S. cerevisiae*.

SNO1 is required for PLP synthesis when ammonia is limiting *in vivo*

In *S. cerevisiae* each *SNZ* paralog has a corresponding *SNO* glutaminase subunit [24]. The need for the glutaminase activity in PLP synthesis has not been clearly demonstrated *in vivo* [24, 31]. The role of *SNO1* on MVD was tested (Figure 2.5). The strain lacking *SNO1* had a small but reproducible growth defect when thiamine was added, that was corrected with the addition of PN. This result suggested that *SNO1* was required for optimal PLP synthesis. This interpretation was complicated by the presence of other *SNO* paralogs, despite the assumption their expression was completely repressed by the thiamine [24].

The *SNZ/SNO* pathway for PLP synthesis is less complex than the multi enzyme DXP-dependent pathway used by *S. enterica* and other organisms. If functional, introduction of these enzymes into *S. enterica* would i) allow characterization of single paralogs, and ii) provide a heterologous system that could be used to probe *S. enterica* with a simplified metabolic network. A plasmid expressing *SNZ3* from the *lac* promoter on pSU18 (pDM1595) was introduced into a *S. enterica* strain lacking the pyridoxine-5'-phosphate (PNP) synthase (E.C. 2.6.99.20) encoded by *pdxJ* (Figure 2.1). *SNZ3* provided *in trans* supported growth of the resulting strain in minimal (NCE) media with glycerol as a carbon source. These data showed that *SNZ3* was necessary and sufficient to synthesize PLP in the *pdxJ* mutant of *S. enterica* (Figure 2.6A), and were generally consistent with

the ability of *SNZ1* from *Cercospora nicotianae* to complemented a *pdxJ* strain of *E. coli* [49].

NCE medium has high levels of ammonia that could bypass a need for the putative glutaminase activity of Sno3p. In fact, *SNZ3* failed to allow growth of a *pdxJ* mutant when glutamine (1 mM) was provided as sole nitrogen source (Figure 2.6B). When pDM1595 was present and *SNO3* was provided on a compatible plasmid (pDM1596), growth of the *pdxJ* mutant was restored to the level allowed by exogenous pyridoxal. In total, these data showed that *SNZ3* encodes a functional PLP synthase and that *SNO3* is needed only when ammonium concentrations are low.

Snz3p has PLP synthase activity *in vitro*

The PLP synthase activity of *SNZ3* suggested by sequence identity and the *in vivo* data above was confirmed *in vitro.* The *SNZ3* gene was cloned into a pTEV5 vector and His_6-Snz3p was purified to >95 % homogeneity prior to cleaving the His_6 Tag to generate native protein. The purified protein was assayed for PLP synthase activity *in vitro*. An assay reaction with ammonium sulfate, glyceraldehyde-3-phosphate and ribose-5-phosphate as substrates was used to define basic kinetic parameters of Snz3p. The K_m for ribose-5-phosphate was 0.09 ± 0.02 mM while the K_m for glyceraldehyde-3-phosphate was 0.29 ± 0.03 mM (Figure 2.7). The K_{cat} for ribose-5-phosphate was 0.044 min^{-1} while the K_{cat} for glyceraldehyde-3-phosphate was 0.042 min^{-1}. These data were similar to those reported for Snz1p (K_m = 0.11 mM and 0.3 mM, K_{cat} = 0.036 min^{-1} and 0.039 min^{-1} for ribose-5-phosphate and glyceraldehyde-3-phosphate, respectively), when assayed in the absence of Sno1p [19] . Kinetic constants available for the *Bacillus subtilis* PLP synthase in the

absence of the glutaminase subunit demonstrate that while this synthase has a higher affinity for ribose-5-phosphate and glyceraldehyde-3-phosphate (K_m = 0.068 mM and 0.077 mM, respectively), its catalytic turnover is similar (K_{cat} = 0.02 min^{-1}) [6].

2.5 CONCLUSIONS

SNZ3 encodes a functional PLP synthase that uses glyceraldehyde-3-phosphate, ribose-5-phosphate and ammonia as substrates. Despite the near identical kinetic constants of Snz1p and Snz3p, results here demonstrate the two isozymes have different roles *in vivo*. The data showed that *SNZ2* or *SNZ3*, but not *SNZ1*, was sufficient to generate PLP for growth on MVD medium (i.e., in the absence exogenous PN or Thi). *SNZ1* supported growth only when PN and/or thiamine were provided. This result suggested that *SNZ1* was unable to satisfy the PLP requirement for thiamine synthesis. The finding that *SNZ2* and/or *SNZ3* are important for thiamine, specifically HMP, synthesis supports a connection between *SNZ2/3* and the Thi5p family of enzymes. The poorly characterized Thi5p enzymes use PLP as a substrate rather than a co-factor to generate the HMP-P moiety used for thiamine synthesis [50, 51]. The finding that the lack of a specific *SNZ* paralog impacts thiamine biosynthesis, out of many metabolic pathways that use PLP as a cofactor, suggests that there are unique requirements for PLP in this pathway, likely involving the Thi5p family of enzymes.

To our knowledge these data provided the first evidence of distinct roles for the *SNZ* paralogs *in vivo* that was not due to regulation of gene expression. It is worth noting that the phenotypes key to the above conclusions were not obvious from past studies using dropout media [24, 31]. In fact, the previous studies led to the conclusion that *SNZ1* encoded the primary PLP synthase, a conclusion that the results herein bring into doubt.

Although not conclusive from the results with *S. cerevisiae*, studies with the heterologous host *S. enterica* showed that the glutaminase subunit *SNO3* is dispensable for PLP synthesis in the presence of excess ammonia. These data further showed that the DXP-dependent pathway for PLP synthesis could be replaced by a single gene (*SNZ3*) in *S. enterica*, and defined a heterologous system that will be valuable in studies to probe network structure with a simplified B_6 metabolism.

ACKNOWLEDGEMENTS

We thank Jorge Escalante-Semerena for rTEV protease and use of the Applied Biosystems 7500 Fast Real-Time PCR System. We also thank Andrew Borchert for technical assistance with qRT-PCR. This work was supported by an award from the competitive grants program at the NIH (GM095837) to DMD and a Graduate Research Fellowship Grant (DGE-1443117) from the NSF to MDP.

2.6 REFERENCES

1. Toney, M.D., *Controlling reaction specificity in pyridoxal phosphate enzymes.* Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 2011. **1814**(1): p. 1407-1418.

2. Tanaka, T., Y. Tateno, and T. Gojobori, *Evolution of vitamin B6 (pyridoxine) metabolism by gain and loss of genes.* Mol Biol Evol, 2005. **22**(2): p. 243-50.

3. Fitzpatrick, T.B., et al., *Two independent routes of de novo vitamin B6 biosynthesis: not that different after all.* The Biochemical journal, 2007. **407**(1): p. 1-13.

4. Mukherjee, T., et al., *Pyridoxal phosphate: biosynthesis and catabolism.* Biochim Biophys Acta, 2011. **1814**(11): p. 1585-96.

5. Burns, K.E., et al., *Reconstitution and biochemical characterization of a new pyridoxal-5'-phosphate biosynthetic pathway.* Journal of the American Chemical Society, 2005. **127**(11): p. 3682-3683.

6. Raschle, T., N. Amrhein, and T.B. Fitzpatrick, *On the two components of pyridoxal 5'-phosphate synthase from Bacillus subtilis.* Journal of Biological Chemistry, 2005. **280**(37): p. 32291-32300.

7. Strohmeier, M., et al., *Structure of a bacterial pyridoxal 5'-phosphate synthase complex.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(51): p. 19284-19289.

8. Makoff, A.J. and A. Radford, *Genetics and biochemistry of carbamoyl phosphate biosynthesis and its utilization in the pyrimidine biosynthetic pathway.* Microbiol. Mol. Biol. Rev., 1978. **42**(2): p. 307-328.

9. Klem, T.J. and V.J. Davisson, *Imidazole Glycerol Phosphate Synthase: The Glutamine Amidotransferase in Histidine Biosynthesis.* Biochemistry, 1993. **32**(19): p. 5177-5186.

10. Romero, R.M., M.F. Roberts, and J.D. Phillipson, *Anthranilate synthase in microorganisms and plants.* Phytochemistry, 1995. **39**(2): p. 263-276.

11. Viswanathan, V.K., J.M. Green, and B.P. Nichols, *Kinetic Characterization of 4-Amino 4-Deoxychorismate Synthase from Escherichia coli.* Journal of Bacteriology, 1995. **177**(20): p. 5918-5923.

12. Huang, X., H.M. Holden, and F.M. Raushel, *Channeling of Substrates and Intermediates in Enzyme-Catalyzed Reactions.* New York, 2001.

13. Belitsky, B.R., *Physical and Enzymological Interaction of Bacillus subtilis Proteins Required for de Novo Pyridoxal 5'-Phosphate Biosynthesis.* Journal of Bacteriology, 2004. **186**(4): p. 1191-1196.

14. Dong, Y.X., et al., *Characterization of the products of the genes SNO1 and SNZ1 involved in pyridoxine synthesis in Saccharomyces cerevisiae.* European Journal of Biochemistry, 2004. **271**(4): p. 745-752.

15. Gengenbacher, M., et al., *Vitamin B6 biosynthesis by the malaria parasite Plasmodium falciparum: biochemical and structural insights.* J Biol Chem, 2006. **281**(6): p. 3633-41.

16. Zein, F., et al., *Structural insights into the mechanism of the PLP synthase holoenzyme from Thermotoga maritima.* Biochemistry, 2006. **45**(49): p. 14609-14620.

17. Raschle, T., et al., *Reaction mechanism of pyridoxal 5'-phosphate synthase: Detection of an enzyme-bound chromophoric intermediate.* Journal of Biological Chemistry, 2007. **282**(9): p. 6098-6105.

18. Hanes, J.W., I. Keresztes, and T.P. Begley, *13C NMR snapshots of the complex reaction coordinate of pyridoxal phosphate synthase.* Nature chemical biology, 2008. **4**(7): p. 425-30.

19. Neuwirth, M., et al., *X-ray crystal structure of Saccharomyces cerevisiae Pdx1 provides insights into the oligomeric nature of PLP synthases.* FEBS Lett, 2009. **583**(13): p. 2179-86.

20. Raschle, T., et al., *Intersubunit cross-talk in pyridoxal 5'-phosphate synthase, coordinated by the C terminus of the synthase subunit.* Journal of Biological Chemistry, 2009. **284**(12): p. 7706-7718.

21. Moccand, C., M. Kaufmann, and T.B. Fitzpatrick, *It takes two to Tango: Defining an essential second active site in pyridoxal 5???-phosphate synthase.* PLoS ONE, 2011. **6**(1).

22. Robinson, G.C., et al., *Structural definition of the lysine swing in Arabidopsis thaliana PDX1: Intermediate channeling facilitating vitamin B6 biosynthesis.* Proceedings of the National Academy of Sciences, 2016. **113**(40): p. E5821-E5829.

23. Rodrigues, M.J., et al., *Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis.* Nature Chemical Biology, 2017(January): p. 6-11.

24. Rodriguez-Navarro, S., et al., *Functional analysis of yeast gene families involved in metabolism of vitamins B1 and B6.* Yeast, 2002. **19**(14): p. 1261-76.

25. Tambasco-Studart, M., et al., *Vitamin B6 biosynthesis in higher plants.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(38): p. 13687-92.

26. Titiz, O., et al., *PDX1 is essential for vitamin B6 biosynthesis, development and stress tolerance in Arabidopsis.* The Plant Journal, 2006. **48**(6): p. 933-946.

27. Boycheva, S., et al., *Consequences of a deficit in vitamin B6 biosynthesis de novo for hormone homeostasis and root development in Arabidopsis.* Plant physiology, 2015. **167**(1): p. 102-17.

28. Moccand, C., et al., *The pseudoenzyme PDX1.2 boosts vitamin B6 biosynthesis under heat and oxidative stress in Arabidopsis.* J Biol Chem, 2014. **289**(12): p. 8203-16.

29. Leuendorf, J.E., et al., *Arabidopsis thaliana PDX1.2 is critical for embryo development and heat shock tolerance.* Planta, 2014. **240**(1): p. 137-46.

30. Dell'Aglio, E., S. Boycheva, and T.B. Fitzpatrick, *The Pseudoenzyme PDX1.2 Sustains Vitamin B6 Biosynthesis as a Function of Heat Stress.* Plant Physiol, 2017. **174**(4): p. 2098-2112.

31. Stolz, J. and M. Vielreicher, *Tpn1p, the plasma membrane vitamin B6 transporter of Saccharomyces cerevisiae.* J Biol Chem, 2003. **278**(21): p. 18990-6.

32. Downs, D.M., *Understanding microbial metabolism.* Annu Rev Microbiol, 2006. **60**: p. 533-59.

33. Koenigsknecht, M.J. and D.M. Downs, *Thiamine biosynthesis can be used to dissect metabolic integration.* Trends in Microbiology, 2010. **18**(6): p. 240-247.

34. Li, X.C. and J.C. Fay, *Cis-Regulatory Divergence in Gene Expression between Two Thermally Divergent Yeast Species.* Genome biology and evolution, 2017. **9**(5): p. 1120-1129.

35. Vogel, H.J. and D.M. Bonner, *Acetylornithase of Escherichia coli: partial purification and some properties.* Journal of Biological Chemistry, 1956. **218**: p. 97-106.

36. Balch, W.E., et al., *Methanogens: reevaluation of a unique biological group.* Microbiol. Rev., 1979. **43**(2): p. 260-96.

37. Davis, R.W., et al., *Advanced bacterial genetics*. Manual for genetic engineering. 1980, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory. x, 254.

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

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

40. Hegemann, J.H. and S.B. Heick, *Delete and Repeat: A Comprehensive Toolkit for Sequential Gene Knockout in the Budding Yeast Saccharomyces cerevisiae Johannes.* Methods in Molecular Biology, 2011. **765**(3): p. 83-97.

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

42. Teste, M.A., et al., *Validation of reference genes for quantitative expression analysis by real-time RT-PCR in Saccharomyces cerevisiae.* BMC Mol Biol, 2009. **10**: p. 99.

43. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.

44. Rocco, C.J., et al., *Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from Escherichia coli.* Plasmid, 2008. **59**(3): p. 231-7.

45. Nosaka, K., et al., *Genetic regulation mediated by thiamin pyrophosphate-binding motif in Saccharomyces cerevisiae.* Mol Microbiol, 2005. **58**(2): p. 467-79.

46. Mojzita, D. and S. Hohmann, *Pdc2 coordinates expression of the THI regulon in the yeast Saccharomyces cerevisiae.* Mol Genet Genomics, 2006. **276**(2): p. 147-61.

47. Brion, C., et al., *Deciphering regulatory variation of THI genes in alcoholic fermentation indicate an impact of Thi3p on PDC1 expression.* BMC Genomics, 2014. **15**(1): p. 1085-1085.

48. Wightman, R. and P.A. Meacock, *The THI5 gene family of Saccharomyces cerevisiae: distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine.* Microbiology, 2003. **149**(Pt 6): p. 1447-60.

49. Wetzel, D.K., et al., *Functional complementation between the PDX1 vitamin B6biosynthetic gene of Cercospora nicotianae and pdxJ of Escherichia coli.* FEBS Letters, 2004. **564**(1-2): p. 143-146.

50. Lai, R.Y., et al., *Thiamin Pyrimidine Biosynthesis in Candida albicans : A Remarkable Reaction between Histidine and Pyridoxal Phosphate.* J Am Chem Soc, 2012.

51. Coquille, S., et al., *The last piece in the vitamin B1 biosynthesis puzzle: structural and functional insight into yeast 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase.* J Biol Chem, 2012. **287**(50): p. 42333-43.

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

53. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter.* Journal of Bacteriology, 1995. **177**(14): p. 4121-4130.

Table 2.1 – Strains, Plasmids and Primers

Strain Number	**Genotype**
Saccharomyces cerevisiae	
YJF153	WT
DMy49	*snz1*::kanMX-loxP
DMy51	*snz1*::kanMX-loxP *snz3*::hphMX-loxP
DMy52	*snz1*::kanMX-loxP *snz2*::natMX-loxP *snz3*::hphMX-loxP
DMy53	*snz2*::kanMX-loxP
DMy54	*snz2*::kanMX-loxP *snz3*::natMX-loxP
DMy55	*sno1*::kanMX-loxP
DMy56	*snz1*::kanMX-loxP *snz2*::hphMX-loxP
DMy57	*snz3*::kanMX-loxP
Salmonella enterica	
DM7080	Δ*araCBAD*
DM15839	Δ*araCBAD pdxJ662*::Kn / pSU18
DM15840	Δ*araCBAD pdxJ662*::Kn / pDM1595
DM15843	Δ*araCBAD pdxJ662*::Kn / pDM1595 pBAD24
DM15844	Δ*araCBAD pdxJ662*::Kn / pDM1595 pDM1596
Plasmid Name	**Description or Reference**
pSU18	[49]
pBAD24	[50]
pDM1595	pSU18-*SNZ3*
pDM1596	pBAD24-*SNO3*
Primer Name	**Sequence**
CM	CCTCGACATCATCTGCCC
AGP3R	CGTTCCAGAATAGAAGGTCGA
RPD3R	TGTCAACTATGCGGGTGGTTT
pdxJ F	AACGCACAGTAAAAACGAAGAAAGATTAACGAGGAT TGTCGTGTAGGCTGGAGCTGCTTC
pdxJ R	GGGCAATCTCTACAATATCCGTTCCCAGGCCGAGAAT CGCCATATGAATATCCTCCTTAG
5’ SNZ2 SacI	TAGGGAGCTC<u>AGGAGG</u>ACAGCT**ATG**TCAGAATTCAAG GTTAAAACTG
3’ SNZ2 XbaI	TAGGTCTAGACTACCATCCGATTTCAGAAAGTC
SNZ1 F	AGTAAATATACACAGTACTAATATTCAGTTAATTATCA CGCAGCTGAAGCTTCGTACGC
SNZ1 R	GGAAAAGTGTTATAATGCTCAAAATACCTGTTCAAAG AAAGCATAGGCCACTAGTGGATCTG

SNZ2 F	ACTATAATAGAAAAATAAGTATATCGTAAAAAAGACA AAACAGCTGAAGCTTCGTACGC
SNZ2 R	TCGAAGGAAACAAATTAGCGTTGTGTGAGCATCGCTA GTTGCATAGGCCACTAGTGGATCTG
SNO1 F	TTCATTTCGTTAAATAGAAAGAAAAACCATATCTTAA AGTCAGCTGAAGCTTCGTACGC
SNO1 R	AGGTTTTGGTAATATAAAAATGTGGAAAACCGGCGGT ATTGCATAGGCCACTAGTGGATCTG
5’ SNO2 *Nco*I 2	TAGGACC**ATG**GCCGTCGTTATCGGAGT
3’ SNO2 *Pst*I 2	TAGGCTGCAGAGGCGAGTTCAGAATGAACA
5’ SNZ2 *Nhe*I	TAGGGCTAGCATGTCAGAATTCAAGGTTAAAACTG
3’ SNZ2 *Nco*I	TAGGCCATGGCTACCATCCGATTTCAGAAAGTC
ALG9 qRT-PCR F	TCACGGATAGTGGCTTTGGT
ALG9 qRT-PCR R	CATTCACTACCGGTGCCTTC
UBC6 qRT-PCR F	ATCCTGGCTGGTCTGTCTCA
UBC6 qRT-PCR R	ATTGATCCTGTCGTGGCTTC
SNZ2/3 qRT-PCR F	GCAATGATCCGTACCAAAGG
SNZ2/3 qRT-PCR R	CCGCCTTAATCTTGGTGATG

Underlining identifies an added ribosome binding site (RBS), bold letters represent start codon

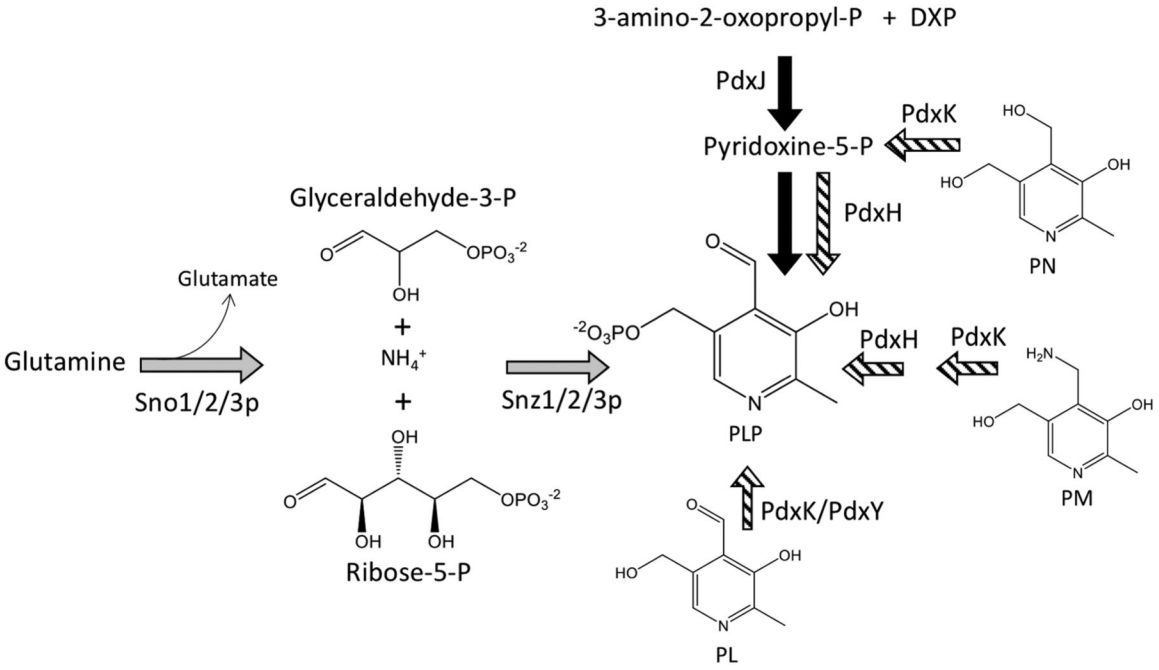


Figure 2.1 – Formation and modification of the pyridine heterocycle in PLP biosynthesis. The enzymatic steps that form and modify the pyridine heterocycle in the DXP-dependent (black arrows), and -independent (gray arrows), pathways for PLP synthesis are depicted. Enzymatic steps involved in Vitamin B_6 salvage are shown with hashed arrows. Proteins catalyzing each step are indicated by the relevant arrows, using *E. coli* (black/hatched) or *S. cerevisiae* (gray) nomenclature. Abbreviations: DXP – deoxyxyulose-5-phosphate; PL – pyridoxal; PM – pyridoxamine; PN – pyridoxine; PLP – pyridoxal-5'-phosphate.

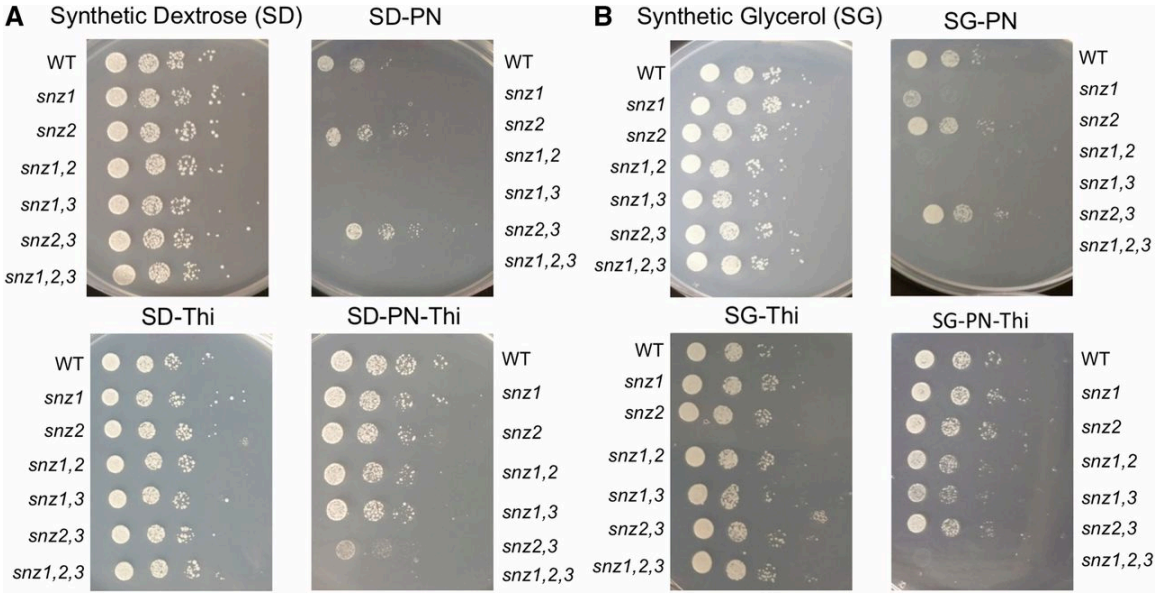


Figure 2.2 – *SNZ1* is required for PLP synthesis in the presence of thiamine. Strains with deletions of *SNZ1*, *SNZ2, SNZ1,2, SNZ1,3*, *SNZ2,3* and *SNZ1,2,3* were grown overnight in YPD and 5 µL each of dilutions from 10^{-2} to 10^{-7} were spotted on synthetic dextrose (**A**) or glycerol (**B**) media or a dropout derivative of this medium lacking pyridoxine and/or thiamine as indicated. Plates were incubated for two (SD) or three (SG) days at 30 °C. Growth of the strain deleted for *SNZ3* was indistinguishable from the *snz2* mutant.

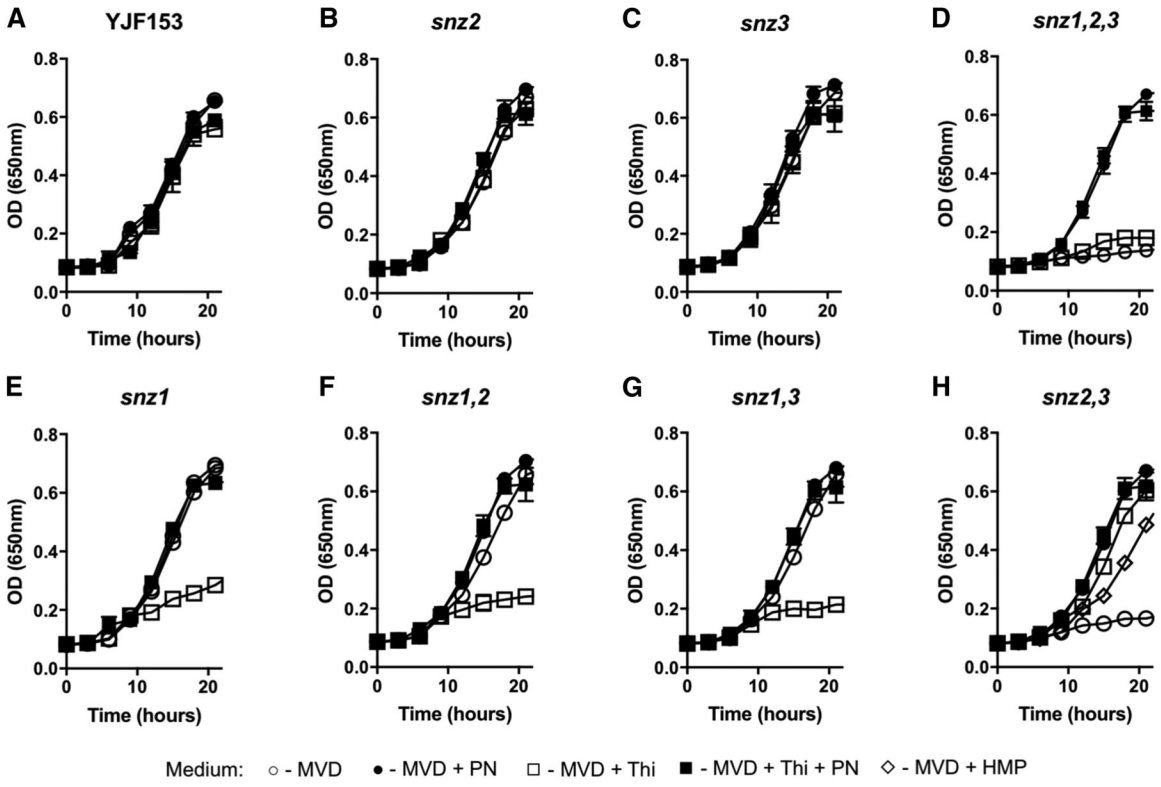


Figure 2.3 – *SNZ2* or *SNZ3* is required for growth on minimal dextrose media. Growth of wildtype *S. cerevisiae* (**A**) or strains with deletions of *SNZ2* (**B**), *SNZ3* (**C**), *SNZ1,2,3* (**D**), *SNZ1* (**E**), *SNZ1,2* (**F**), *SNZ1,2,3* (**G**), or *SNZ2,3* (**H**) was monitoredon minimal vitamin media with dextrose (MVD; circles). The strains were also grown in MVD with added PN (filled symbols), and/or thiamine (squares), and the legend is shown. In the case of the *snz2,3* mutant, growth with added HMP (200 nM) is also shown (diamonds). Error bars indicate the standard deviation of three independent biological replicates.

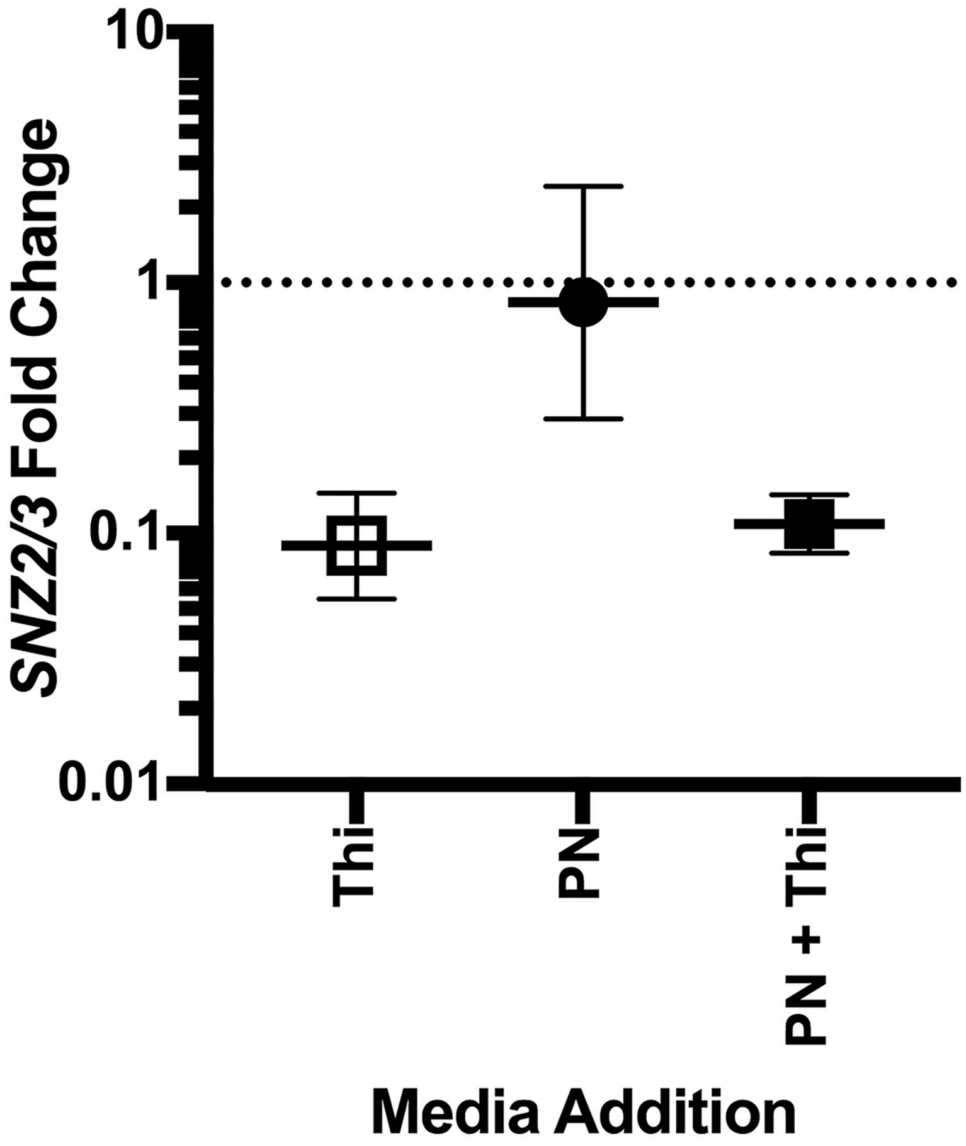


Figure 2.4 – *SNZ2/3* are repressed when thiamine is present in minimal vitamin dextrose medium. Expression of *SNZ2* and *SNZ3* was determined by qRT-PCR when YJF153 was grown in MVD medium with various additions. Fold change represents the ratio of expression on MVD supplemented with thiamine (0.4mg/L), pyridoxine (0.4mg/L), or (as indicated) compared to the expression on MVD with no supplements. Error bars indicate the 95 % confidence interval of three independent biological replicates.

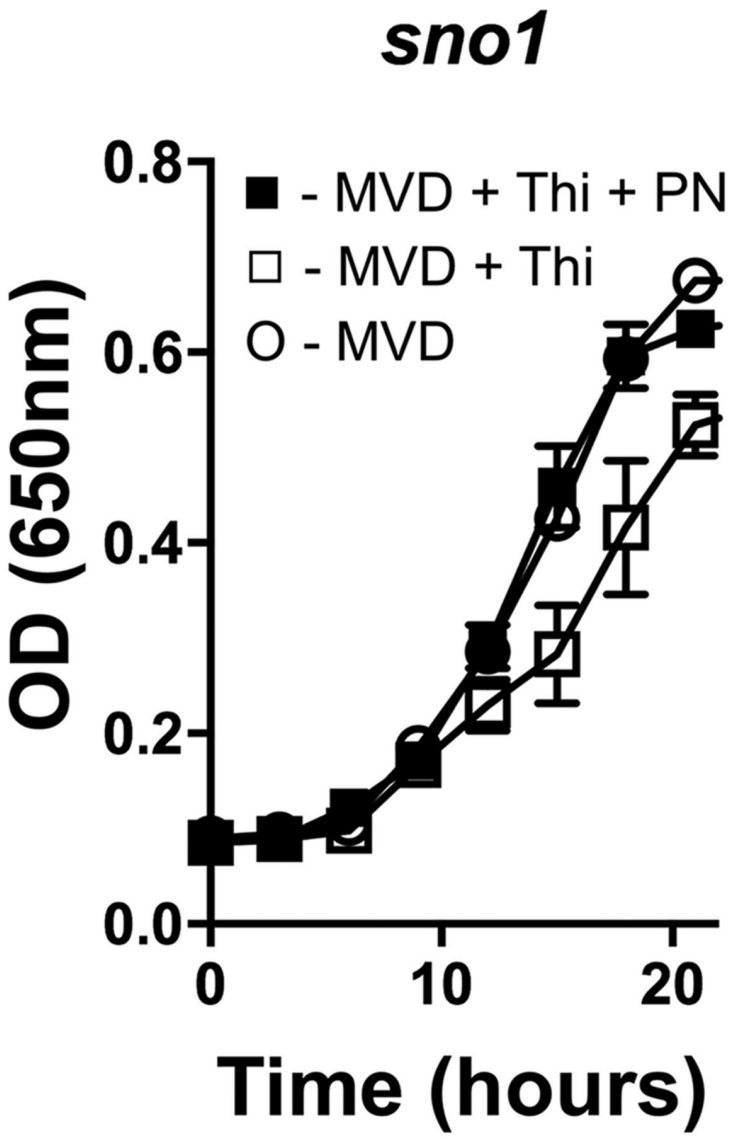


Figure 2.5 – *SNO1* contributes to PLP biosynthesis in *S. cerevisiae*. *A sno1* mutant was grown on minimal vitamin media with dextrose (open circles), with addition of thiamine (open squares), and with addition of thiamine and PN (filled squares). Error bars indicate the standard deviation of three independent biological replicates.

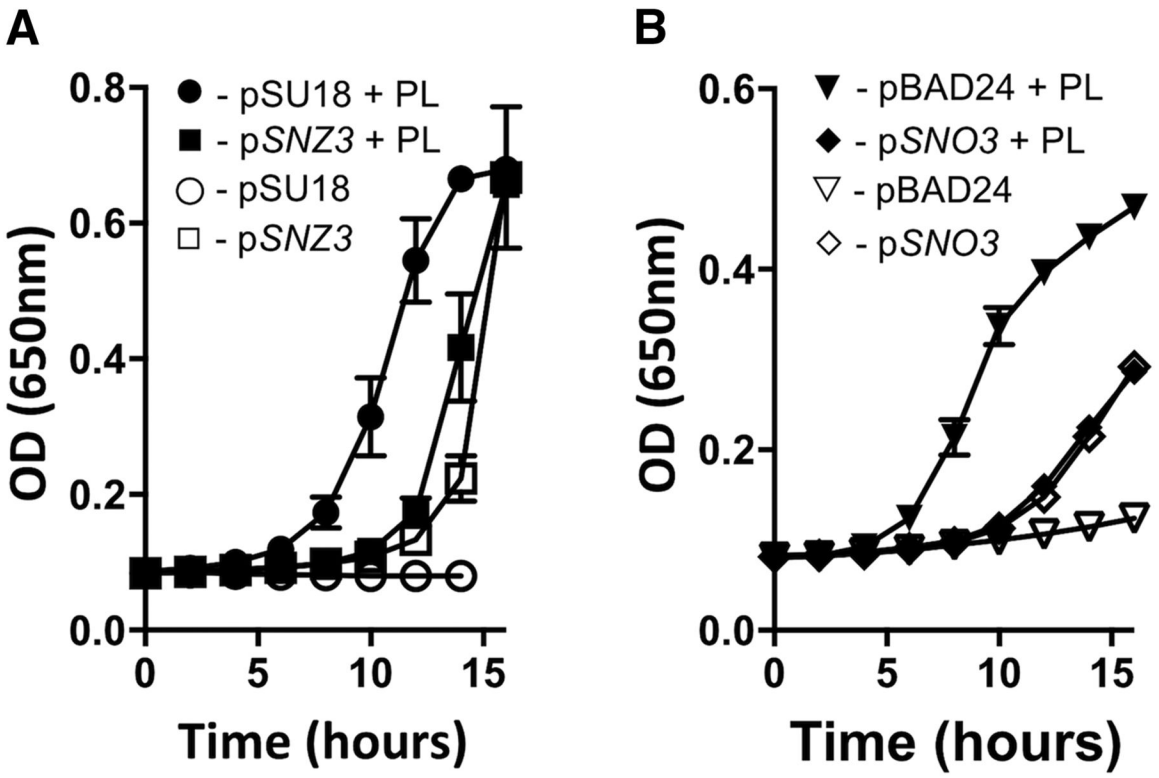


Figure 2.6 – *SNZ3* and *SNO3* can synthesize PLP in *Salmonella enterica*. Growth of a *pdxJ* mutant of *S. enterica* with various plasmids was monitored for growth. In panel (**A**) the *pdxJ* mutant carried an empty vector (pSU18) (circles) or a plasmid expressing *SNZ3* (squares). Each of these two strains were grown in NCE (i.e., high ammonia) minimal medium with glycerol. In panel (**B**) the *pdxJ* mutant carried a plasmid expressing *SNZ3*. In addition, the strain had a compatible empty vector (triangles), or a plasmid expressing p*SNO3* (diamonds). All strains were grown in NCN minimal medium with glycerol and glutamine (i.e., low ammonia). Open symbols represent growth in the absence, while solid symbols represent the presence, of 100 nM pyridoxal. Growth was monitored as a function of optical density at 650 nm with shaking at 37 °C. Error bars indicate the standard deviation of three independent biological replicates.

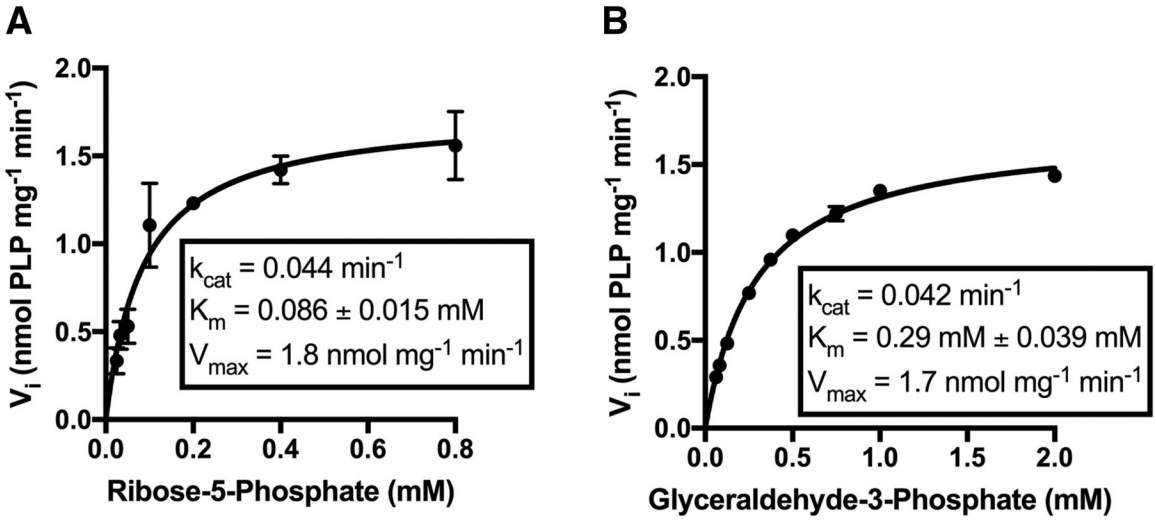


Figure 2.7 – Snz3p is a PLP synthase. Saturation curves for Snz3p were determined by measuring the initial rate of PLP formation vs ribose-5-phosphate **(A)** or D/L-glyceraldehyde-3-phosphate **(B)** as a substrate. Reactions were performed in 50 mM Tris pH 8.0 at 37 °C, containing 85 μM Snz3p. When ribose-5-P was titrated, the reaction mix contained 2 mM D/L glyceraldehyde-3-phosphate and 20 mM NH_4. When D/L glyceraldehyde-3-phosphate was titrated, the reaction mix contained 1 mM ribose-5-phosphate and 20 mM NH_4. All reactions were performed in triplicate, and error bars are shown.

CHAPTER 3

FUNCTIONAL CHARACTERIZATION OF THE HMP-P SYNTHASE OF *LEGIONELLA PNEUMOPHILA* (Lpg1565)[2]

[2] Paxhia MD, Swanson MS, Downs DM. 2020. *Mol Microbiol*. 5:1–15.
Reprinted here with permission of the publisher

3.1 ABSTRACT

The production of the pyrimidine moiety in thiamine synthesis, 2-methyl-4-amino-5-hydroxymethylpyrimidine phosphate (HMP-P), has been described to proceed through the Thi5-dependent pathway in *Saccharomyces cerevisiae* and other yeast. Previous work found that *Sc*Thi5 functioned poorly in a heterologous context. Here we report a bacterial ortholog to the yeast HMP-P synthase (Thi5) was necessary for HMP synthesis in *Legionella pneumophila*. Unlike *Sc*Thi5, *Lp*Thi5 functioned *in vivo* in *Salmonella enterica* under multiple growth conditions. The protein *Lp*Thi5 is a dimer that binds pyridoxal-5'-phosphate (PLP), apparently without a solvent-exposed Schiff base. A small percentage of *Lp*Thi5 protein co-purifies with a bound molecule that can be converted to HMP. Analysis of variant proteins both *in vivo* and *in vitro* confirmed that residues in sequence motifs conserved across bacterial and eukaryotic orthologs modulate the function of *Lp*Thi5.

Importance

Thiamine is an essential vitamin for the vast majority of organisms. There are multiple strategies to synthesize and salvage this vitamin. The predominant pathway for synthesis of the pyrimidine moiety of thiamine involves the Fe-S cluster protein ThiC. An alternative pathway utilizes Thi5, a novel enzyme that uses PLP as a substrate. The Thi5-dependent pathway is poorly characterized in yeast and has not been characterized in Bacteria. Here we demonstrate that a Thi5-dependent pathway is necessary for thiamine biosynthesis in *Legionella pneumophila* and provide biochemical data to extend knowledge of the Thi5 enzyme, the corresponding biosynthetic pathway, and the role of metabolic network architecture in optimizing its function.

3.2 INTRODUCTION

Thiamine pyrophosphate (TPP), the active form of Vitamin B_1, is a cofactor important for several enzymes in central metabolism including transketolase, pyruvate dehydrogenase, and α-ketoglutarate dehydrogenase. The cofactor is comprised of two moieties, 2-methyl-4-amino-5-hydroxymethylpyrimidine diphosphate (HMP-PP) and 4-methyl-5-β-hydroxyethylthiazole phosphate (THZ-P), that are independently synthesized and combined to form TPP (Jurgenson *et al.*, 2009). The pyrimidine precursor HMP-P is synthesized from an intermediate in the purine biosynthetic pathway (aminoimidazole ribotide) by the phosphomethylpyrimidine synthase ThiC (E.C. 4.1.99.17) an enzyme encoded in plants, archaea and most bacteria (Jurgenson *et al.*, 2009). However some organisms, notably *Saccharomyces cerevisiae*, use Thi5 (HMP-P synthase) in place of ThiC to generate HMP-P. Labeling studies showed that the atoms of HMP originate from histidine and a B6 vitamer, presumably pyridoxal-5'-phosphate (PLP) (Tazuya *et al.*, 1989, Tazuya *et al.*, 1995, Ishida *et al.*, 2008). (Figure 3.1A) The hemiascomycetes clade of fungi contain multiple paralogs of *THI5*, and genetic studies in *S. cerevisiae* S288c showed any member of the *THI5/11/12/13* gene family was sufficient for the synthesis of HMP-P (Wightman & Meacock, 2003). Other fungi encode a single ortholog of *THI5*, and a *Schizosaccharomyces pombe* mutant lacking the *THI5* homolog *nmt1* is a thiamine auxotroph that specifically requires HMP for growth (Maundrell, 1990, Schweingruber *et al.*, 1991).

The Thi5 HMP-P synthase from lower eukaryotes (*e.g., Candida albicans* and *S. cerevisiae*) is a member of the periplasmic binding protein superfamily (COG0715), and

three-dimensional crystal structures of these proteins have been solved (Lai *et al.*, 2012, Coquille *et al.*, 2012), and are available in protein databases (PDBs 4H65, 4H67, 4H6D, 4ESX). The *S. cerevisiae* Thi5 protein was purified with PLP in its active site and formed dimers in solution, an oligomeric state confirmed by its crystal structure (Coquille *et al.*, 2012). An *in vitro* study of the *C. albicans* Thi5 protein showed that the protein had HMP-P synthase activity, albeit unexpectedly low. When PLP and iron were provided in a reaction mixture, *Ca*Thi5 produced between 0.2 and 0.5 mol of HMP-P per mol of protein (Lai *et al.*, 2012). The small amount of product formed, in addition to the lack of the stimulation of enzyme activity by histidine, suggested that a necessary molecule was supplied by the protein. *Ca*Thi5 variants H66G and H66N had no activity in the *in vitro* assay (Lai *et al.*, 2012) and other variants that altered residue H66 were inactive *in vivo* (Coquille *et al.*, 2012). In total, the data led authors of one study with *Ca*Thi5 to conclude that the enzyme used a suicide mechanism to donate a moiety of residue H66 to yield HMP-P (Lai *et al.*, 2012). The *in vitro* formation of HMP-P required a high concentration (600 μM) of *Ca*Thi5, the addition of excess ferrous iron (900 μM) and molecular oxygen (Lai *et al.*, 2012). The above conditions of the assay, and the low product yield achieved, make it formally possible that a component or condition relevant for Thi5 activity *in vivo* was not recapitulated in the *in vitro* work.

Despite the presumed *in vivo* availability of both histidine and pyridoxine substrates, *Sc*Thi5 fails to support the growth of a *thiC* mutant of *S. enterica* on minimal glucose medium (Palmer *et al.*, 2015). The ease with which suppressors that restored growth were isolated suggested that the architecture of the metabolic network was impacting Thi5 activity. An additional study found that the Thi5-dependent thiamine

synthesis in *S. cerevisiae* required an ammonium-dependent PLP synthase (E.C. 4.3.3.6; *SNZ2, SNZ3*) (Paxhia & Downs, 2019). Thus, it was formally possible that the lack of activity in a bacterial host reflected the need for a PLP synthesis/delivery system specific for the Thi5 enzyme. Such enzymes are not present in *S. enterica*, which generates PLP via a DXP-dependent pathway (Figure 3.1B). The study herein was initiated to address this possibility, and to better understand the cellular function and biochemical properties of a bacterial Thi5 protein. To gain insight into the impact of network structure on thiamine synthesis, we identified and focused on a *THI5* homolog encoded within the *Legionella pneumophila* thiamine biosynthesis operon (Rodionov *et al.*, 2002, Sahr *et al.*, 2012). This homolog is regulated by a thiamine-pyrophosphate riboswitch in the 5' untranslated region, and a study in *L. pneumophila* Paris demonstrated that transcription from this region is repressed by thiamine-pyrophosphate and activated by CsrA (Sahr *et al.*, 2017). The work herein verified the role of the *THI5* ortholog *lpg1565* (redesignated *thi5*) in thiamine synthesis in *L. pneumophila* and analyzed its ability to function in *S. enterica.* Further, data herein showed that *Lp*Thi5 co-purifies with a molecule that can be converted to HMP, and identified variants that have altered activity *in vivo.* In total, this work expands our understanding of the function of Thi5, and its integration in the metabolic network of the cell.

3.3 MATERIALS AND METHODS

Strains

Strains used in this study are derivatives of strain *Salmonella enterica* Serovar Typhimurium strain LT2, *Escherichia coli* K12*,* or *Legionella pneumophila* strain Lp02.

Strains, plasmids and their source are listed in Table 3.1. An *lpg1565* deletion mutant in *L. pneumophila* strain Lp02, a thymidine auxotroph derived from Philadelphia-1 (Berger & Isberg, 1993), was constructed by recombineering as described previously (Bryan *et al.*, 2013). In brief, the *lpg1565* gene and its ~ 750 bp 5' and 3' flanking sequences were amplified using primers lpg1565F and lpg1565R (Table 3.1). The DNA product was cloned by standard methods into vector pGEM T-easy (Promega), creating pGEM-*lpg1565*. An *FRT*-flanked *cat* cassette encoding chloramphenicol resistance was amplified from pKD3 (Datsenko & Wanner, 2000) using primers lpg1565P0 and lpg1565P2 (Table 3.1), and the product purified. Replacement of the *lpg1565* allele with the *cat* cassette was accomplished by co-transforming *E. coli* strain DY330, which encodes λ-red recombinase, by electroporation with the purified *cat* fragment and pGEM-*lpg1565*. Candidate *E. coli* DY330 colonies harboring pGEM-*lpg1565::cat* were screened by PCR, and the corresponding recombinant plasmids then transformed into *E. coli* DH5α. Next, the recombinant allele *lpg1565::cat* was amplified by PCR using primers lpg1565F and lpg1565R (Table 3.1) and transferred to *L. pneumophila* strain Lp02 by natural transformation and chloramphenicol selection. Replacement of the Lp02 *lpg1565* locus with *cat* was confirmed by DNA sequencing.

Media and Chemicals

E. coli and *S. enterica* strains were routinely grown on Nutrient Broth (NB) containing 8 g/L Difco Nutrient broth and 5 g/L NaCl. For protein purification the cultures were grown in superbroth (SB; 32 g/L vegetable tryptone, 20 g/L yeast extract (Fisher Scientific), 5 g/L NaCl with 0.05 N NaOH). Solid media contained 1.5 % agar. Kanamycin (Kn) and

Ampicillin (Ap) were added to rich media at 50, or 100 mg/L, respectively. Minimal media was No-carbon E salts (NCE) (Vogel & Bonner, 1956) with 1 mM $MgSO_4$, 0.1X trace minerals (Balch *et al.*, 1979) and 11 mM glucose, gluconate, galactose or 13.2 mM ribose as a sole carbon source as indicated. 2-methyl-4-amino-5-hydroxymethylpyrimidine (HMP) was purchased from LabSeeker, Inc. (Wujiang City, China).

L. pneumophila strains were grown in ACES-buffered Feeley-Gorman (FG) broth (10 g/L ACES, 17.5 g/L Casein enzymatic hydrolysate, 3 g/L Beef extract, 0.4 g/L L-cysteine HCl•H_2O, 0.25 g/L ferric pyrophosphate, adjusted to pH 6.9 with KOH), and on ACES-buffered Yeast Extract plates containing 10 g/L ACES, 10 g/L yeast extract, 0.4 g/L L-cysteine HCl•H_2O, 0.25 g/L ferric pyrophosphate, adjusted to pH 6.9 with KOH, 1.7 % agar and 0.2 % activated charcoal (Feeley *et al.*, 1978, Pasculle *et al.*, 1980). Defined media for growth of *L. pneumophila* was Modified Ristroph Medium (MRM) (Ristroph *et al.*, 1981, Sauer *et al.*, 2005). 100 µg/mL thymidine was added to all media when growing thymine auxotrophs. When divalent cations were present in the incubation of *Lp*Thi5, they were added as the following salts; $MgSO_4$, $NiSO_4$, $MnCl_2$, $ZnSO_4$, $CoCl_2$, $CaCl_2$, $CdCl_2$, $CuSO_4$, $Fe(II)(NH_4)_2(SO_4)_2$, $Fe(III)_2(SO_4)_3$. Chemicals were purchased from Sigma-Aldrich, St. Louis, MO unless otherwise indicated.

Growth Analysis

Growth of bacterial strains was monitored at OD_{650} in a 96-well plate with a BioTek ELx808 plate reader. For *L. pneumophila*, strains were grown overnight in ACES-buffered FG (5 mL) to an OD_{650} = 1.5 (late-log) and pelleted (9400 x g for 5 min). The cell pellet was resuspended in an equal volume of ddH_2O and used to inoculate (5 %) MRM medium

with indicated supplements. Plates were incubated at 37 °C with fast shaking, and data were plotted using Prism 7 (Graph Pad). *S. enterica* strains were grown overnight in NB Ap (2 mL) prior to pelleting and resuspension in an equal volume of 0.85 % NaCl. The cell suspension was used to inoculate (1 %) the indicated medium. Plates were incubated at 37 °C with medium shaking, and data were plotted using Prism 7 (Graph Pad).

Bioinformatics analyses

A BLAST-P search of the RefSeq Protein database (09/20/19) used the amino acid sequence encoded by *lpg1565* as a query. The default BLAST parameters were used, and the top 1000 hits were examined. Bacterial homologs with a conserved CCCXC motif were identified in the *Endozoicomonas*, *Fluorobacter* and *Legionella* genera. Based on this result, finished and permanent draft genomes on the IMG database of *Endozoicomonas*, *Fluorobacter* and *Legionella* genera of Bacteria and all available Eukaryotic genomes were queried to find *THI5* homologs (Chen *et al.*, 2019). Iterative BLAST-P searches using the default settings and the predicted amino acid sequences of *THI5* homologs from *L. pneumophila* (*lpg1565*), *Endozoicomonas elysicola* (B144DRAFT_03762) and *S. cerevisiae* (*THI5*; *YFL058W*) as the query identified 587 genomes that contained *THI5* homologs with a CCCXC motif. The amino acid sequences of the 35 non-redundant bacterial *THI5* homologs were downloaded from the IMG database in FASTA format, and the headers of the files were modified to place the organism name first. Representative eukaryotic homologs from *S. cerevisiae* and *S. pombe* were also included in the phylogenetic analysis. Geneious Prime 2019.2.1 was used to align these 41 homologs. A MUSCLE Alignment with 100 iterations was used to generate a PHYLIP alignment file

which was analyzed via Smart Model Selection PhyML to generate a phylogenetic tree using the Le-Gascuel (LG) substitution model (Edgar, Guindon *et al.*, 2010, Lefort *et al.*, 2017). The phylogenetic tree was annotated using the interactive Tree Of Life (iTOL) and Adobe Illustrator (Letunic & Bork, 2019). The presence or absence of homologs of genes encoding key enzymes involved in PLP biosynthesis and salvage (*pdxJ*, *pdxS*, and *pdxH*) were identified using the IMG Pathway Profile for the 215 non-redundant genomes with *THI5* homologs (Chen *et al.*, 2019). In cases where multiple genomes were found within the same species, only genomes from the original type strain were used to annotate putative PLP biosynthesis homologs.

Molecular Techniques

Plasmids were constructed and modified using standard molecular techniques. pTac85 (Marsh, 1986), pJB98 (Hammer & Swanson, 1999), and pET-28b(+) (Novagen) were isolated using the PureYield Plasmid MiniPrep System (Promega, Madison, WI). Q5 DNA polymerase (New England Biolabs, Ipswich, MA) was used to amplify DNA with primers synthesized by Integrated DNA Technologies, Coralville, IA or Eton Bioscience, Inc., Research Triangle Park, NC. PCR products were purified using the PCR purification kit (Qiagen, Venlo, Limburg, The Netherlands). Restriction endonucleases and ligase were purchased from New England Biolabs, Ipswich, MA.

The *THI5* homolog *lpg1565* was amplified by PCR with *L. pneumophila* gDNA as a template using primers LpTHI5 *Nco*I F and LpTHI5 *Sal*I R, listed in Table 3.1. The amplified product was purified, digested with *Nco*I and *Sal*I and ligated into pTac85 (Marsh, 1986), resulting in pDM1486, which was confirmed by sequencing. The *lpg1565*

coding sequence was cloned with additional nucleotides encoding a C-terminal His_6 tag into the *Nco*I/*Sal*I sites of pTac85. The relevant insert was then cloned into pET-28b(+) at the *Nco*I/*Sal*I sites to generate plasmid pDM1630, which expressed *Lp*Thi5-His_6 fusion protein for purification. *ScTHI5,* with flanking *Nco*I and *Sal*I sites, was codon-optimized for translation in *E. coli*, purchased from Genscript and ligated into pTac85 to generate pDM1625. Primers for site-directed mutagenesis were designed using the Agilent QuikChange Primer Design webtool and are listed in Table 3.1. Variants were created following instructions from the QuikChange II Site Directed Mutagenesis kit (Agilent Technologies, Inc., Santa Clara, CA), and confirmed by sequencing (Eton Bioscience, Inc, Research Triangle Park, NC).

To construct pDM1631, the *lpg1565* locus was amplified by PCR from gDNA using primers lpg1565 comp F and lpg1565 comp R (Table 3.1). The fragment was purified, digested with *Bam*HI and *Sac*I and ligated into pJB98. The *lpg1565-lpg1569* genes that constitute an operon (Sahr *et al.*, 2012) were PCR amplified with gDNA as a template, using primers ThiOperon_Fwd and ThiOperon_Rev (Table 3.1). The amplified product was purified, digested with *Kpn*I and *Xba*I and ligated into pJB98 to generate pDM1632. The cloned operon was modified to remove *lpg1565* using the Q5 Site-directed mutagenesis kit (New England Biolabs, Ipswitch, MA) with primers NEBQC-thiOF and NEBQC-thi5R (Eton Bioscience, Inc, Research Triangle Park, NC). This manipulation placed the starting codon for *lpg1566* where the annotated starting codon for *lpg1565* is found and generated pDM1633, which was confirmed by sequencing.

Protein Purification

A culture (100 mL) of *E. coli* BL21-AI carrying pDM1630 was grown overnight at 30 °C in NB Kn, and 4 flasks of 1.5 L of SB Kn + pyridoxine (1 mM) were inoculated (1 %). Each of four cultures were grown at 37 °C with shaking (200 rpm). When the OD_{650} reached 0.6, temperature was lowered to 22 °C and arabinose and IPTG were added to a final concentration of 0.2 % and 1 mM, respectively. Incubation continued for 19-20 hours prior to harvesting by centrifugation. Typical cell yield was 8 g/L under these conditions. 100 g of cells were resuspended to a total volume of 200 mL in Buffer A [50 mM HEPES (Fisher Scientific), 300 mM NaCl, 20 mM Imidazole (Fisher Scientific), 1 mM TCEP (Gold Biotechnology), pH 7.5 at 4 °C] with DNAse (0.025 mg/mL), lysozyme (1 mg/mL) and phenylmethylsulfonyl fluoride (0.1 mg/mL) and incubated on ice for one hour. The cell suspension was lysed at 20 kpsi using a Constant Systems Limited One Shot (United Kingdom), and cell lysate was cleared at 48,000 x g (50 min, 4 °C). The cell-free extract was passed through a 0.45 μM PVDF filter (Millipore) and injected onto two pre-equilibrated 5 mL HisTrap HP Ni-Sepharose columns connected in sequence. The protein was washed with 5 column volumes of Buffer A, 5 column volumes of 4 % Buffer B (Buffer A + 480 mM Imidazole (Fisher Scientific), pH 7.5 at 4 °C) and finally eluted from the column with a gradient of Buffer B from 4 % to 100 % over 10 column volumes. Fractions containing *Lp*Thi5 as determined by SDS-PAGE were combined and concentrated by centrifugation using a 10 kDa filter (Millipore), exchanged into 50 mM HEPES buffer with 10 % glycerol, 1 mM TCEP, pH 7.5, using a PD10 column following the manufacturer's instructions (GE Healthcare), flash-frozen in liquid nitrogen and stored at -80 °C until use. Protein concentration was determined by extinction coefficient using

the theoretical molecular weight and A_{280} extinction coefficient of *Lp*Thi5-His_6 as determined by the ExPASy Protparam database (ε_{280} = 30870 M^{-1} cm^{-1}) (Gasteiger *et al.*, 2003). A typical purification yielded *Lp*Thi5-His_6 that was > 85 % pure as determined by densitometry (Figure 3.S1).

Determination of PLP and iron content

The PLP was released from *Lp*Thi5 by denaturing the protein. Fifteen nmol of *Lp*Thi5 at a purity of 95 % was treated in 50 mM HEPES, pH 7.5 with 0.1 M NaOH in a total volume of 150 µL, the protein removed using a Nanosep 10 kDa spin-filter (PALL), and the absorbance of the supernatant measured at 392 nm. PLP was quantified using the empirically determined extinction coefficient of PLP (ε_{392}) of 1.965 x 10^6 M^{-1} cm^{-1} in 50 mM HEPES pH 7.5. Occupancy was calculated based on the percent purity from the starting protein preparation.

Iron associated with *Lp*Thi5 was determined by Ferene assay (Kennedy *et al.*, 1984, Palmer & Downs, 2013). Briefly, purified *Lp*Thi5 was concentrated using an Amicon Ultra 0.5 mL 30 kDa desalting filter at 14,000 x g for five minutes and resuspending the concentrated protein with 400 µL of 50 mM HEPES, pH 7.5 treated with 2 % Chelex. After three cycles of concentration and dilution with the Chelex-treated HEPES, *Lp*Thi5 concentration was determined by extinction coefficient, ε_{280} = 30870 M^{-1} cm^{-1}. HCl (0.06 N, 50 µL) was added to 19 nmol *Lp*Thi5 in 25 mM HEPES (25 µL) and incubated at 80 °C for 10 minutes. The following were added sequentially with mixing by vortex between additions: 0.96 M ammonium acetate (125 µL), 0.2 M ascorbic acid (25 µL), 87 mM SDS (25 µL), and 30 mM Ferene (25 µL). Precipitated protein was removed by centrifugation

at 9000 x g for 5 minutes and absorbance of the supernatant was measured at 593 nm. Iron content was determined based on the extinction coefficient of Ferene complexed with Fe^{2+} (ε_{593} = 35.5 x 10^3 M^{-1} cm^{-1}).

Size exclusion chromatography

The quaternary structure of purified *Lp*Thi5 (which released bioactive molecule) was assessed by size exclusion chromatography (Hong *et al.*, 2012). A BioRad NGC Chromatography System with an SEC 650 column (BioRad) was equilibrated with 2 column volumes of HEPES (50 mM), TCEP (1 mM) pH 7.5 at 4 °C. Absorbance at 280 nm was monitored to detect when the protein was eluting. Injection of blue dextran (50 µL, 3 mg/mL) in duplicate with a flow rate of 1 mL/min determined the void volume (V_0) was 8.9 mL. The elution volumes (V_e) of duplicate samples of BioRad Gel Filtration Standards (50 µL) were used to establish a standard curve of K_{av} ($\frac{V_e - V_0}{V_{column} - V_0}$) vs $\log_{10}$ (molecular weight) (Figure 3.5). Three samples, each containing 17 nmol *Lp*Thi5, were injected at a flow rate of 1 mL/min, and the elution volume was used to determine the K_{av}. The molecular size of *Lp*Thi5 was determined by interpolation from the standard curve ($K_{AV} = -0.231 \log(MW) + 0.732;\ R^2 = 0.992$), and standard deviation was determined from the deviation in retention time.

Detection of HMP

Bioassay. Purified *Lp*Thi5 (20 nmol) in 200 µL HEPES (50 mM, pH 7.5) was incubated at 37 °C. At T=0, 4, 16 hr or indicated time, the sample was shifted to 95 °C for 5 min, and denatured protein was pelleted by centrifugation (17,000 x g, 5 minutes). When indicated,

ferrous ammonium sulfate hexahydrate was added to 200 μM in the protein sample. The supernatant was evaluated for its ability to stimulate growth of a *thiC* or *thiI* mutant strain of *S. enterica* on minimal medium. Samples (5 μL) were spotted on soft agar embedded with the relevant mutant strain overlaid on minimal medium. Growth was assessed after ~16 hr incubation at 37 ºC. A *thiI* mutant responds to thiamine, or TPP, while a *thiC* mutant responds to thiamine, TPP, HMP, HMP-P.

High Performance Liquid Chromatography (HPLC) and MS. Supernatants judged active and inactive by bioassay were subjected to HPLC, monitoring absorbance at 270 nm, which is a lambda max for HMP. Initial control experiments showed that a peak (270 nm) correlated with a bioactive molecule released from an overnight incubation of 3 mg *Lp*Thi5. After a 16 hr incubation of 3 mg *Lp*Thi5 at 37 °C, supernatant (2 mL) resulting from heat treatment was applied to a C18 Sep-Pak Plus cartridge (Waters) for concentration prior to HPLC analysis. Six cartridges were conditioned using 5 mL acetonitrile (ACN) followed by 10 mL 50 mM HEPES, pH 7.5. Five 1 mL fractions were eluted with acetonitrile and collected, evaporated, and tested for biological activity by bioassay. Biologically active fractions (Fractions 1 and 2 from each column) were combined, evaporated, suspended to a total volume of 0.8 mL in water, and 100 μL fractions were sequentially injected onto the HPLC. A 3 mg sample of *Lp*Thi5 (T=0) was treated in parallel. Fractions collected from the C18 Sep-Pak Plus cartridges were dehydrated, resuspended in 100 μL water and injected onto the HPLC for analysis.

Concentrated samples were separated by reversed-phase HPLC using a Shimadzu LC20-AT delivery system equipped with a 250 x 4.6 mm Luna C_{18} column (Phenomenex). The UV-Vis spectrum was monitored over time using a photodiode array detector

(Shimadzu). Samples were eluted with a flow rate of 1 mL/min with a gradient of water and ACN; 100 % water for 10 minutes, a gradient from 0 – 15 % ACN over 10 minutes, and 15 % ACN for 10 minutes. Fractions (0.5 mL) collected over 30 min were dehydrated and resuspended in 50 μL water. Fractions that contained biological activity were determined by bioassay (spotting 1 μL onto a soft-agar overlay containing a *thiC* strain as an indicator and observing growth). The relevant fractions, eluting from 19.5 – 22 min, were pooled and analyzed by MALDI-TOF MS. Fractions over the same retention time from control samples, which had no activity in the bioassay, were independently pooled. Concentrated fractions and an HMP-P standard were analyzed by MALDI-TOF MS using a Bruker Autoflex (TOF) mass spectrometer at the Proteomics and Mass Spectrometry facility at UGA. Fractions containing biological activity from 25 mg of *Lp*Thi5 that had been incubated for 16 hours were combined from several HPLC runs to facilitate identification by MALDI-TOF MS.

3.4 RESULTS AND DISCUSSION

THI5 orthologs co-occur with DXP-dependent PLP synthesis in prokaryotes

In order to test the simple hypothesis that lack of Thi5 function in *S. enterica* was due to an absence of an Snz/PdxS pathway for PLP biosynthesis, genomes in the IMG database containing *THI5* orthologs were found using iterative BLAST-P searches, and the phylogeny of those homologs as well as the presence of homologs of key PLP biosynthetic genes were assessed. A CCCXC motif distinguishes *Sc*Thi5p from the structural homolog *N*-formyl-4-amino-5-aminomethyl-2-methylpyrimidine-binding ThiY, and so this motif was used to identify putative *THI5* homologs that had been annotated as a substrate binding

protein for an ABC transport system (Bale *et al.*, 2010). Bacterial homologs with the CCCXC motif were identified in the *Endozoicomonas*, *Fluorobacter* and *Legionella* genera. Orthologs identified without the N-terminal 60 amino acids were not considered, since this region of the protein contains residues important for coordinating PLP in the *Sc*Thi5p active site (Coquille *et al.*, 2012). Presence of homologs for PNP synthase (PdxJ; E.C. 2.6.99.2), or a subunit of the glutamine-hydrolyzing PLP synthase (PdxS, Snz; E.C. 4.3.3.6) were used to define the presence of the DXP-dependent or -independent pathway for PLP biosynthesis, respectively (Figure 3.2). Among Eukaryotes *THI5* homologs were only found in fungi. Each of the 180 fungal genomes with a putative *THI5* gene had a homolog of both *pdxS* and *pdxH*, indicating that they used the DXP-independent pathway for PLP biosynthesis and could salvage B_6 vitamers, similar to *S. cerevisiae*. Two clades of γ-proteobacteria with *THI5* homologs were identified: four *Endozoicomonas* species and 31 species within the Legionellaceae family. Each of these bacteria had homologs to *pdxJ* and *pdxH* indicating the synthesis of PLP was by the DXP-dependent pathway and suggesting a potential for vitamer salvage. It was noted that many of the Legionellaceae species contained multiple paralogs of *pdxH* (29 of 31 species have 2-4 paralogs), a feature that was not investigated further. Thus, genomic analysis suggested Thi5 enzymes could function in the absence of a PdxS PLP synthase. While this finding appeared to negate the hypothesis that *Sc*Thi5 required this enzyme as a specific PLP delivery system to function *in vivo*, it is worth noting that non-orthologous replacement for PdxS was not ruled out by these analyses.

lpg1565* encodes a Thi5 ortholog and contributes to HMP synthesis in *L. pneumophila

A role for the *THI5* ortholog in a metabolic network using DXP-dependent PLP biosynthesis was investigated using *L. pneumophila. lpg1565* is adjacent to several open reading frames predicted to encode enzymes for thiamine biosynthesis. The operon includes homologs of *thiO, thiG* and *thiF* (*lpg1566, lpg1567, lpg1569,* respectively), and a short ORF with homology to *thiS*. Each of these proteins is involved in the synthesis of the thiazole moiety of thiamine in defined pathways. The operon also had an ORF homologous to a fusion protein between *thiD2* and *thiE* (*lpg1568*) (Figure 3.1C) (Sahr *et al.*, 2012). A mutant of *L. pneumophila* with an in-frame insertion-deletion of *lpg1565* was constructed and subjected to growth analyses. The Δ*lpg1565* strain with an empty vector (pJB98) required exogenous thiamine for growth on Modified Ristroph Medium (Figure 3.3A). A plasmid containing the *lpg1565* gene (pDM1631) was unable to complement the nutritional requirement of the Δ*lpg1565* strain, suggesting the insertion was polar on one or more downstream genes (data not shown). A plasmid carrying the entire locus (*lpg1565* – *lpg1569*; pDM1632) restored growth of the Δ*lpg1565* strain in the absence of thiamine (Figure 3.3B). In contrast, a plasmid derived from pDM1632 that lacked *lpg1565* (pDM1633) failed to complement the nutritional defect of the Δ*lpg1565* strain. Full growth of the Δ*lpg1565* strain carrying pDM1633 was restored by the addition of 100 nM thiamine. Partial growth of this strain was restored with the addition of 2.5 mM HMP (Figure 3.3C). Together these results demonstrate that *L. pneumophila* requires the *lpg1565* gene product for thiamine biosynthesis, specifically to generate the pyrimidine moiety. Based on these data and the bioinformatic analyses, *lpg1565* was renamed *thi5*, and the gene product designated *Lp*Thi5 throughout.

The inability of HMP to restore robust growth to the *thi5* mutant suggested *L. pneumophila* lacked the salvage system characterized in *E. coli*, *S. enterica* and *S. cerevisiae*. In these organisms, exogenous HMP is incorporated into the biosynthetic pathway by the hydroxymethylpyrimidine kinase ThiD (E.C. 2.7.1.49). The presence of a *thiD2-thiE* fusion homolog (*lpg1568*) in *L. pneumophila*, rather than full length *thiD* and *thiE* genes is likely to be the genomic difference that accounts for the weak HMP salvage observed. Other ThiD2-ThiE fusion proteins lack the HMP kinase activity associated with purified full-length ThiD while retaining HMP-P kinase activity. Thus, a non-specific kinase may generate some HMP-P from exogenous HMP, allowing salvage via Lpg1568 in *L. pneumophila* (Thamm *et al.*, 2017). The ability to salvage thiamine reflects the presence of a homolog to a thiamine pyrophosphokinase (E.C. 2.7.6.2 ; ThiN) in *L. pneumophila* (Lpg2497), although a transporter for thiamine has not been identified.

*Lp*Thi5 functions in the *S. enterica* metabolic network

The co-occurrence phylogenetic data above was not consistent with the simple hypothesis that Thi5 proteins require an Snz/PdxS-like protein to be functional. However, the phylogeny of Thi5 proteins clustering with distinct PLP biosynthesis pathways may suggest distinct structural features have evolved in Thi5 orthologs within metabolic networks of similar architecture. Therefore, we considered whether there were structural aspects of Thi5 that differ in concordance with the pathway for PLP biosynthesis in the organism, and whether structural differences could account for functional differences *in vivo*. As a first step, the functional complementation of a *thiC* mutant of *S. enterica* with a protein from a Eukaryote (*Sc*Thi5p) and the ortholog from a Prokaryote (*Lp*Thi5) were

compared. *ScTHI5* was codon-optimized for expression in *S. enterica* and cloned into pTac85 under the regulation of a Tac promoter (pDM1625). *Lpthi5* was similarly inserted into pTac85 to generate pDM1486. Both constructs were introduced into a *thiC* mutant of *S. enterica*, and growth was monitored on several carbon sources to test the response to different metabolisms (Figure 3.4). Neither *Lpthi5* nor *ScTHI5* restored growth on any carbon source without induced expression of the respective genes. However, when transcription was induced with IPTG (100 μM), *Lpthi5* (pDM1486), allowed the *S. enterica* mutant grew on ribose, glucose, gluconate and galactose (Figure 3.4, Panel Ⅰ, A-D). In contrast, the codon-optimized *ScTHI5* conferred growth on ribose but not glucose and supported intermediate growth on gluconate and galactose (Figure 3.4, Panel Ⅱ, E-H).

When the thiazole moiety THZ (100 nM) was added to the growth medium, expression of *ScTHI5* allowed full growth of the *thiC* mutant on glucose, galactose and gluconate, and also increased the growth conferred by *Lpthi5*. Titration experiments on glucose medium showed the addition of THZ reduced the HMP concentration required for full growth of a *thiC* mutant from 25 nM to 10 nM (Supplemental Table 3.1). The growth stimulation by THZ is likely due to the kinetics of thiamine synthase (ThiE), which allows excess THZ-P to drive the reaction when HMP-PP levels are low (Backstrom *et al.*, 1995). Additionally, the *thiC* mutant required less HMP when using ribose as a carbon source than when growing on glucose, also considered to be an indirect effect of metabolic flux differences (Supplemental Table 3.1). In total, the data above support the conclusion that *Lp*Thi5 is an HMP-P synthase that is more active in the *S. enterica* metabolic network than is the *Sc*Thi5 enzyme.

*Lp*Thi5 is a pyridoxal-5'-phosphate binding protein

*Lp*Thi5-His_6 was purified from *E. coli* and characterized for quaternary structure and cofactor occupancy, using size exclusion chromatography and UV-Vis spectroscopy, respectively. Based on protein standards of known molecular weight, *Lp*Thi5 was calculated to be 86 ± 5.8 kDa, reasonably close to the theoretical weight of a *Lp*Thi5 dimer (72.2 kDa) (Figure 3.5) and thus consistent with the designation of *Sc*Thi5p as a dimer based on size exclusion chromatography as well as its crystal structure (Coquille *et al.*, 2012).

After purification in buffer lacking PLP, *Lp*Thi5 had absorbance maxima at 325 and 425 nm, features that are characteristic of a tautomeric equilibrium of the Schiff-base of characterized PLP-binding proteins (Mozzarelli & Bettati, 2006) (Figure 3.6A). Dialysis of *Lp*Thi5 against HEPES buffer (50 mM, pH 7.5) containing PLP (200 µM) increased the absorbance maxima at 325 and 425 nm, indicating the as-purified protein was not fully occupied with PLP (Figure 3.6B). Denaturation of *Lp*Thi5 (100 µM in 200 µL) with NaOH (0.1 M), released PLP that was detected by the absorption maximum at 392 nm, characteristic of free PLP in HEPES (Geders *et al.*, 2012). The PLP occupancy of the as-purified protein was 29 ± 1 %, while dialysis against PLP-containing buffer raised the occupancy to 63 ± 2 %.

In a majority of PLP-binding proteins, treatment with $NaBH_4$ (1 mM) reduces the Schiff-base between the protein and PLP, resulting in a noticeable reduction in absorbance at 425 nm (Toney, 2011, Soniya & Chandra, 2018). Treatment of as-purified *Lp*Thi5 with up to 25 mM $NaBH_4$ did not affect the absorbance at 425 nm. Accordingly, if present, a Schiff-base between PLP and the protein was not solvent accessible (Figure 3.6A).

Surprisingly, when a variant lacking the proposed catalytic residue H70 (*Lp*Thi5$_{H70A}$) was treated with $NaBH_4$ (25 mM), absorbance at 425 nm decreased significantly (Figure 3.6C). These data suggested the catalytic histidine residue might shield access to the Schiff base in the wild-type protein.

*Lp*Thi5 purifies with a bound metabolite

Efforts to detect HMP-forming activity of *Lp*Thi5 with the addition of its presumed substrates, histidine and PLP (Tazuya *et al.*, 1995, Ishida *et al.*, 2008) under multiple conditions were not successful. Specifically, these conditions included incubations of 100 µM *Lp*Thi5 at 37 °C for 16 hours in 50 mM HEPES pH 7.5, aerobically or anaerobically, with the addition of 1-10X histidine, PLP or PLP + histidine. However, control experiments generated an unexpected result. When *Lp*Thi5 (20 nmol) in HEPES (50 mM, pH 7.5) was incubated overnight (16 hr) at 37 °C, a biologically active compound was released when the protein was denatured with heat (Figure 3.7A). The biologically active molecule allowed the growth of a *thiC* mutant, but not a *thiI* mutant strain of *S. enterica*; moreover, this activity was detected in samples with no added substrate. Based on these data the compound was not thiamine, which would have satisfied the requirement of both the *thiC* and the *thiI* mutants. Further, neither histidine, PLP, or histidine and PLP together detectably impacted the biological activity present after the incubation described (data not shown).

A series of experiments identified conditions required for the production of the biologically active compound. First, if a protein sample was pre-treated with acid (10 % TFA), base (0.1 M NaOH), or temperature (95 °C, 5 mins) prior to incubation at 37 °C,

(i.e., T=0), no biological activity was detected. These data demonstrated that time was required for formation, and that HMP was not present in the purification. Second, when the concentration of HEPES was 10, rather than 50 mM, no biological activity was detected. Third, the formation of a biologically active molecule was dependent on buffer: Tris (40 mM, pH 7.5), Bis-Tris-Propane (40 mM, pH 7.5), or HEPES resulted in its formation, but incubation in MOPS (40 mM, pH 7.5) did not. Additionally, since the protein sample was exchanged into a buffer lacking PLP using a PD10 desalting column, the relevant molecule was likely bound to the enzyme. In total, these data suggested the molecule that purified with the protein was not the fully formed biologically active compound. It is formally possible that the protein is partially modified *in vivo*, and that this modified *Lp*Thi5 may serve as the substrate for the *in vitro* reaction as previously proposed for *Ca*Thi5 (Lai *et al.*, 2012).

HMP is released after incubation of Thi5

The bioassay results indicated that the active compound released from *Lp*Thi5 was either HMP, HMP-P or a molecule that the *S. enterica* could transport and convert to HMP-P. Indeed, MALDI-TOF MS identified the active molecule as HMP (Fig 7). The molecule released after an overnight incubation of 690 nmol of *Lp*Thi5 was concentrated with a C18 SepPak and separated by HPLC with a Luna C_{18} column. Fractions (0.5 mL) were collected across 30 minutes. From a sample incubated overnight, biological activity was found in fractions with a retention time of 20-21.5 minutes and the corresponding peak had UV-Vis spectral features similar to HMP. Moreover, a standard of authentic HMP eluted at a similar retention time. A control (T=0) sample generated no peak with a UV-Vis spectrum

consistent with HMP(-P); nor did these collected fractions have biological activity (Figure 3.7B). The active fractions of the sample from an overnight incubation contained a mass of 140.0 Da when ionized by MALDI-TOF MS. This peak, present in the active but not an inactive fraction, was consistent with the monoisotopic mass of ionized HMP (Figure 3.7C). Together these data support the conclusion that *Lp*Thi5 co-purified with a molecule that generated HMP upon incubation in buffer. It was considered possible the relevant molecule was HMP-P since it was not clear the phosphorylated form would be stable during the ionization protocol.

Iron decreases the time required for the release of HMP(-P) from *Lp*Thi5

Initial experiments suggested that more than a 12-hour incubation was needed to generate detectable HMP from a purified Thi5 protein. Based on its putative metal binding motif (CCCXC) (Bale *et al.*, 2010), we tested the hypothesis that the reaction is catalyzed by metals. To do so, purified Thi5 was pre-incubated with a 10-fold excess of a series of metals, and reaction time was quantified. *Lp*Thi5 samples supplemented with metals were denatured after eight-hours at 37 °C and product formation was determined by bioassay. Of the metals tested (Mg^{2+}, Ni^{2+}, Mn^{2+}, Zn^{2+}, Co^{2+}, Ca^{2+}, Cd^{2+}, Cu^{2+}, and Fe), only Fe stimulated the release of HMP by *Lp*Thi5. The presence of either Fe (III) or Fe (II) salts increased the efficiency of HMP formation with *Lp*Thi5. Titration of 100μM protein with 50, 100, 200, 500 μM, or 1 mM iron showed that a Fe:protein ratio of 2:1 was optimal for product formation. When *Lp*Thi5 (100μM) and iron ammonium sulfate (200 μM) were present in buffer, HMP was detectable by bioassay after 2 hr at 37°C. In these conditions, HMP release increased over time plateauing at 4 hr at a level similar to that released with

an overnight incubation in buffer alone (Figure 3.7D). Based on the ability of iron to increase the rate of product release, it was plausible that the generation of HMP in an overnight incubation was facilitated by iron that was in the *Lp*Thi5 protein preparation and was acting as a general electron sink. The requirement for buffer suggests that the exchange of ions between the protein and the solvent, perhaps accelerated by iron, could be important for the release of HMP. Bound iron in purified *Lp*Thi5 was quantified by Ferene assay (Kennedy *et al.*, 1984, Palmer & Downs, 2013). A sample of *Lp*Thi5 (70.8 μM) contained 545 ± 43 nM Fe, while a buffer-only control contained 140 nM Fe. These data indicated that 18 nmol *Lp*Thi5 contained only 0.10 ± 0.008 nmol of Fe, minimizing the possibility that iron was specifically bound to the purified protein. While not quantifiable precisely, analysis by bioassay suggested 5%-8% of purified *Lp*Thi5 protein released HMP. The stoichiometry of HMP(-P) to *Lp*Thi5 cannot effectively distinguish between the release of a previously bound molecule and the proposed suicide mechanism that was previously suggested (Lai *et al.*, 2012).

Conserved residues are involved in the function of *Lp*Thi5 *in vivo* and *in vitro*

Thi5 variants that have individually altered conserved residues were surveyed for their functional properties. Previous work with *Sc*Thi5 showed alanine substitutions at the PLP-binding lysine residue binding (K62) or the adjacent histidine residue (H66) eliminated function *in vivo* (Coquille *et al.*, 2012). Likewise, neither *Lp*Thi5$_{K66A}$ or *Lp*Thi5$_{H70A}$ complemented a *thiC* strain of *S. enterica* (data not shown). Further, these purified variants did not to produce a biologically active compound when incubated in the presence of iron (data not shown). Previous work with *Sc*Thi5 showed that alanine substitutions in the

CCCXC motif compromised function *in vivo* (Coquille *et al.*, 2012). To corroborate the importance of this motif in *Lp*Thi5, each of the cysteines in the conserved CCCXC motif was changed to Ala, His or Ser by site-directed mutagenesis in pDM1486. None of the twelve variants of *Lp*Thi5 complemented a *thiC* mutant when tested only with ribose as the sole carbon source (data not shown). Further, four of the variants (*Lp*Thi5$_{C191A}$, *Lp*Thi5$_{C192A}$, *Lp*Thi5$_{C193A}$, *Lp*Thi5$_{C195A}$) were purified, and none released a biologically active molecule after a 4 hr incubation in the presence of iron, or overnight in its absence, implying that these conserved residues are essential for Thi5 activity *in vitro* and *in vivo* (data not shown).

The crystal structure of *Sc*Thi5p revealed a GEFG motif involved in hydrogen bonding with the phosphate group of PLP in the active site (Coquille *et al.*, 2012). This motif is strictly conserved between diverse homologs of Thi5 (Figure 3.S2) and was modeled in *Lp*Thi5 (Figure 3.8B). Because of the conservation of this motif as well as residues lining the active site across Bacteria and Eukaryotes (K66, H70 in *Lp*Thi5) it is highly likely the proteins have the same mechanism. If the mechanism proposed for *Ca*Thi5 is correct, it would suggest H70 was a substrate for *Lp*Thi5. *Lp*Thi5 variants of the GEFG loop were tested for activity *in vivo*. A G121A variant failed to complement a *S. enterica thiC* mutant grown with glucose or ribose as carbon source. However, when the expression of this protein was induced, the variant was not soluble, suggesting that the alanine substitution of the G121 residue impacted the folding of *Lp*Thi5. Interestingly, alleles of *Lpthi5* encoding individual alanine substitutions in each position of the GEF motif supported the growth of the *thiC* mutant on ribose medium (Figure 3.9 A-C) but eliminated complementation on glucose (Figure 3.9 D-F). Taken together with previous

results demonstrating a reduced requirement for HMP on ribose compared to glucose (Table 3.S1), the data suggest that these variants had reduced activity *in vivo*. When overexpressed and purified, both *Lp*Thi5$_{G118A}$ and *Lp*Thi5$_{F120A}$ variants produced a product that supported the growth of an HMP auxotroph when incubated in buffer with iron (data not shown). These data indicated that both *Lp*Thi5$_{G118A}$ and *Lp*Thi5$_{F120A}$ variants co-purified with the same molecule as wild-type *Lp*Thi5, consistent with detection of the enzyme's *in vivo* activity.

3.5 CONCLUSIONS

Prior to this work, the Thi5 pathway for the synthesis of HMP was characterized only in fungi. The data herein show that *L. pneumophila* encodes a Thi5 pathway for HMP-P synthesis but lacks an efficient salvage pathway for HMP. *Lp*Thi5 appears to outperform the *S. cerevisiae* Thi5 ortholog in the heterologous metabolic network of *S. enterica*, even when the fungal locus was codon optimized for expression in this host. These differences may suggest structural variances within the PLP biosynthetic pathways of fungi vs. bacteria. When purified, *Lp*Thi5-His6 binds PLP, but the histidine residue in the active site restricts the solvent accessibility of the Schiff base. *Lp*Thi5 purifies with a molecule that when incubated with iron produces HMP. Neither the identity of this molecule, nor the formal possibility that it is a partially modified His70 residue was addressed here. We identified a conserved GEF motif that modulates function *in vivo* and corroborated the importance of other conserved amino acids predicted to be in the active site for the function of this homolog both *in vitro* and *in vivo* (CCCXC, K66, H70). The results herein extend

the understanding of the Thi5 pathway for HMP-P synthesis and confirm its presence beyond Eukaryotes into the Legionellaceae family of Bacteria.

Acknowledgments

We thank Thomas Knight for the construction of the *lpg1565* mutant and Vincent Starai for the gift of *Legionella pneumophila* genomic DNA. This work was supported by an award from the competitive grants program at the NIH (GM095837) to DMD and a Graduate Research Fellowship Grant (DGE-1443117) from the NSF to MDP. The authors state that they have no conflicts of interest in presenting this work.

3.6 REFERENCES

Backstrom, A.D., McMordie, R.A.S., and Begley, T.P. (1995) Biosynthesis of Thiamin I: The Function of the thiE Gene Product. *Journal of the American Chemical Society* **117**: 2351-2352.

Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., and Wolfe, R.S. (1979) Methanogens: reevaluation of a unique biological group. *Microbiol Rev* **43**: 260-296.

Bale, S., Rajashankar, K.R., Perry, K., Begley, T.P., and Ealick, S.E. (2010) HMP binding protein ThiY and HMP-P synthase THI5 are structural homologues. *Biochemistry* **49**: 8929-8936.

Berger, K.H., and Isberg, R.R. (1993) Two distinct defects in intracellular growth complemented by a single genetic locus in *Legionella pneumophila*. *Mol Microbiol* **7**: 7-19.

Bryan, A., Abbott, Z.D., and Swanson, M.S. (2013) Constructing unmarked gene deletions in *Legionella pneumophila*. *Methods Mol Biol* **954**: 197-212.

Chen, I.A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., Huntemann, M., Varghese, N., White, J.R., Seshadri, R., Smirnova, T., Kirton, E., Jungbluth, S.P., Woyke, T., Eloe-Fadrosh, E.A., Ivanova, N.N., and Kyrpides, N.C. (2019) IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* **47**: D666-D677.

Coquille, S., Roux, C., Fitzpatrick, T.B., and Thore, S. (2012) The last piece in the vitamin B1 biosynthesis puzzle: structural and functional insight into yeast 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase. *J Biol Chem* **287**: 42333-42343.

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 U S A* **97**: 6640-6645.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-1797.

Feeley, J.C., Gorman, G.W., Weaver, R.E., Mackel, D.C., and Smith, H.W. (1978) Primary isolation media for Legionnaires disease bacterium. *J Clin Microbiol* **8**: 320-325.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**: 3784-3788.

Geders, T.W., Gustafson, K., and Finzel, B.C. (2012) Use of differential scanning fluorimetry to optimize the purification and crystallization of PLP-dependent enzymes. *Acta crystallographica. Section F, Structural biology and crystallization communications* **68**: 596-600.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**: 307-321.

Hammer, B.K., and Swanson, M.S. (1999) Co-ordination of legionella pneumophila virulence with entry into stationary phase by ppGpp. *Mol Microbiol* **33**: 721-731.

Hong, P., Koza, S., and Bouvier, E.S. (2012) Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *J Liq Chromatogr Relat Technol* **35**: 2923-2950.

Ishida, S., Tazuya-Murayama, K., Kijima, Y., and Yamada, K. (2008) The direct precursor of the pyrimidine moiety of thiamin is not urocanic acid but histidine in *Saccharomyces cerevisiae*. *J Nutr Sci Vitaminol (Tokyo)* **54**: 7-10.

Jurgenson, C.T., Begley, T.P., and Ealick, S.E. (2009) The structural and biochemical foundations of thiamin biosynthesis. *Annu Rev Biochem* **78**: 569-603.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**: 845-858.

Kennedy, M.C., Kent, T.A., Emptage, M., Merkle, H., Beinert, H., and Munck, E. (1984) Evidence for the formation of a linear [3Fe-4S] cluster in partially unfolded aconitase. *J Biol Chem* **259**: 14463-14471.

Lai, R.Y., Huang, S., Fenwick, M.K., Hazra, A., Zhang, Y., Rajashankar, K., Philmus, B., Kinsland, C., Sanders, J.M., Ealick, S.E., and Begley, T.P. (2012) Thiamin pyrimidine biosynthesis in *Candida albicans* : a remarkable reaction between histidine and pyridoxal phosphate. *J Am Chem Soc* **134**: 9157-9159.

Lefort, V., Longueville, J.E., and Gascuel, O. (2017) SMS: Smart Model Selection in PhyML. *Mol Biol Evol* **34**: 2422-2424.

Letunic, I., and Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**: W256-W259.

Marsh, P. (1986) Ptac-85, an E. coli vector for expression of non-fusion proteins. *Nucleic Acids Res* **14**: 3603.

Maundrell, K. (1990) nmt1 of fission yeast. A highly transcribed gene completely repressed by thiamine. *J Biol Chem* **265**: 10857-10864.

Mozzarelli, A., and Bettati, S. (2006) Exploring the pyridoxal 5′-phosphate-dependent enzymes. *The Chemical Record* **6**: 275-287.

Palmer, L.D., and Downs, D.M. (2013) The thiamine biosynthetic enzyme ThiC catalyzes multiple turnovers and is inhibited by S-adenosylmethionine (AdoMet) metabolites. *J Biol Chem* **288**: 30693-30699.

Palmer, L.D., Paxhia, M.D., and Downs, D.M. (2015) Induction of the Sugar-Phosphate Stress Response Allows Saccharomyces cerevisiae 2-Methyl-4-Amino-5-Hydroxymethylpyrimidine Phosphate Synthase To Function in *Salmonella enterica*. *J Bacteriol* **197**: 3554-3562.

Pasculle, A.W., Feeley, J.C., Gibson, R.J., Cordes, L.G., Myerowitz, R.L., Patton, C.M., Gorman, G.W., Carmack, C.L., Ezzell, J.W., and Dowling, J.N. (1980) Pittsburgh pneumonia agent: direct isolation from human lung tissue. *J Infect Dis* **141**: 727-732.

Paxhia, M.D., and Downs, D.M. (2019) SNZ3 Encodes a PLP Synthase Involved in Thiamine Synthesis in *Saccharomyces cerevisiae*. *G3 (Bethesda)* **9**: 335-344.

Ristroph, J.D., Hedlund, K.W., and Gowda, S. (1981) Chemically defined medium for *Legionella pneumophila* growth. *J Clin Microbiol* **13**: 115-119.

Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., and Gelfand, M.S. (2002) Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. *J Biol Chem* **277**: 48949-48959.

Sahr, T., Rusniok, C., Dervins-Ravault, D., Sismeiro, O., Coppee, J.Y., and Buchrieser, C. (2012) Deep sequencing defines the transcriptional map of *L. pneumophila* and identifies growth phase-dependent regulated ncRNAs implicated in virulence. *RNA Biol* **9**: 503-519.

Sahr, T., Rusniok, C., Impens, F., Oliva, G., Sismeiro, O., Coppee, J.Y., and Buchrieser, C. (2017) The *Legionella pneumophila* genome evolved to accommodate multiple regulatory mechanisms controlled by the CsrA-system. *PLoS Genet* **13**: e1006629.

Sauer, J.D., Bachman, M.A., and Swanson, M.S. (2005) The phagosomal transporter A couples threonine acquisition to differentiation and replication of *Legionella pneumophila* in macrophages. *Proc Natl Acad Sci U S A* **102**: 9924-9929.

Schweingruber, A.M., Dlugonski, J., Edenharter, E., and Schweingruber, M.E. (1991) Thiamine in *Schizosaccharomyces pombe*: dephosphorylation, intracellular pool, biosynthesis and transport. *Curr Genet* **19**: 249-254.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D., and Higgins, D.G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.

Soniya, K., and Chandra, A. (2018) Free energy landscapes of prototropic tautomerism in pyridoxal 5'-phosphate schiff bases at the active site of an enzyme in aqueous medium. *J Comput Chem* **39**: 1629-1638.

Tazuya, K., Azumi, C., Yamada, K., and Kumaoka, H. (1995) Pyrimidine moiety of thiamin is biosynthesized from pyridoxine and histidine in *Saccharomyces cerevisiae*. *Biochem Mol Biol Int* **36**: 883-888.

Tazuya, K., Yamada, K., and Kumaoka, H. (1989) Incorporation of histidine into the pyrimidine moiety of thiamin in *Saccharomyces cerevisiae*. *Biochim Biophys Acta* **990**: 73-79.

Thamm, A.M., Li, G., Taja-Moreno, M., Gerdes, S.Y., de Crécy-Lagard, V., Bruner, S.D., and Hanson, A.D. (2017) A strictly monofunctional bacterial hydroxymethylpyrimidine phosphate kinase precludes damaging errors in thiamin biosynthesis. *Biochem J* **474**: 2887-2895.

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

Vogel, H.J., and Bonner, D.M. (1956) Acetylornithase of *Escherichia coli*: partial purification and some properties. *Journal of Biological Chemistry* **218**: 97-106.

Wightman, R., and Meacock, P.A. (2003) The THI5 gene family of *Saccharomyces cerevisiae*: distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine. *Microbiology* **149**: 1447-1460.

Table 3.1 – Plasmids, Bacterial strains and primers used in this study

Plasmid Name	**Description**
pDM1486	pTac85-*lpg1565*
pDM1625	pTac85-*ScTHI5* (codon-optimized)
pDM1630	pET28b-lpg1565-His_6
pJB98	Amp^R Thy^+ (Hammer & Swanson, 1999)
pDM1631	pJB98-*lpg1565*
pDM1632	pJB98-*lpg1565-69*
pDM1633	pJB98-*lpg1566-69*
Strain	**Genotype**
S. enterica	
DM15269	*ΔthiC1225 ΔaraCBAD* / pDM1486
DM16449	*ΔthiC1225 ΔaraCBAD* / pDM1625
L. pneumophila	
DMLp6	*thyA rpsL hsdR lpg1565::Cm* / pJB98
DMLp7	*thyA rpsL hsdR lpg1565::Cm* / pDM1632
DMLp8	*thyA rpsL hsdR lpg1565::Cm* / pDM1633
Primer Name	**Sequence**
LpTHI5 *Nco*I F	TAGGCCATGGCGATGTCATCACTAAAATCC
LpTHI5 *Sal*I R	TAGGGTCGACTTAATTTTCAAGACAACAGGCG
LpTHI5-HisTag *Sal*I R	TAGGGTCGACTTAGTGATGGTGATGGTGATGATT TTCAAGACAACAGGCG
LpTHI5 K66A F	tgcaacagtatgaatcattgccgcaacaccgaaatctactgttcccagg
LpTHI5 H70A	ttttgctttggctgcaacagtagcaatcattgctttaacaccgaaatc
LpTHI5 C191A	cgatcaattagctggtttaggcgcttgttgtttctgctcaatacaa
LpTHI5 C192A	cgatcaattagctggtttaggctgtgcttgtttctgctcaatacaatttatt

LpTHI5 C193A	tcaattagctggtttaggctgttgtgctttctgctcaatacaatttattgtt
LpTHI5 C195A	ctggtttaggctgttgttgtttcgcctcaatacaatttattgttcctga
LpTHI5 G118A	ggaaacgagtaggttatatcgccgaattcggcaaaaaaattat
LpTHI5 E119A	gaaacgagtaggttatatcggcgcattcggcaaaaaaattattgatg
LpTHI5 F120A	aataatttttttgccggcttcgccgatataacctactcgtttcccaaca
LpTHI5 G121A	gtaggttatatcggcgaattcgccaaaaaaattattgatgatttgg
LpTHI5 C191S	cgatcaattagctggtttaggcagctgttgtttctgctcaatacaa
LpTHI5 C192S	cgatcaattagctggtttaggctgtagctgtttctgctcaatacaatttatt
LpTHI5 C193S	tcaattagctggtttaggctgttgtagcttctgctcaatacaatttattgtt
LpTHI5 C195S	ctggtttaggctgttgttgtttcagctcaatacaatttattgttcctga
LpTHI5 C191H	cgatcaattagctggtttaggccattgttgtttctgctcaatacaa
LpTHI5 C192H	cgatcaattagctggtttaggctgtcattgtttctgctcaatacaatttatt
LpTHI5 C193H	tcaattagctggtttaggctgttgtcatttctgctcaatacaatttattgtt
LpTHI5 C195H	ctggtttaggctgttgttgtttccattcaatacaatttattgttcctga
lpg1565F	AAATTAAGCGGGAATCGAAGTGTAGC
lpg1565R	AAATATGAGGTAAAAATTTCCAGGTCT
lpg1565P0	TTTAAAATAATAACATAAGGAGTTATGGCGATGT CATGTGTAGGCTGGAGCTGCTTC
lpg1565P2	AATGCCTGCTGCCATAATTAATTTTCAAGACAAC AGCATATGAATATCCTCCTTAGTTCC
lpg1565 comp F (*Bam*HI)	AAAGGATCCCCTCCATATATCCAATCTCGCAAG
lpg1565 comp R (*Sac*I)	AAAGAGCTCCCGCATTATGTAATACAAAAGCCA
thiOperon_fwd	attcttcgagctcggtacccCCTCCATATATCCAATCTC
thiOperon_rev	gtcgactctagaggatccccTATCTGCTAAATAATATTGCCG
NEBQC-thiOF	ATGCGAGCAGGCATTGTA
NEBQC-thi5R	AACTCCTTATGTTATTATTTTAAATATTATGGAG

* Underlining designates nucleotides encoding six additional histidine residues at the c-terminus

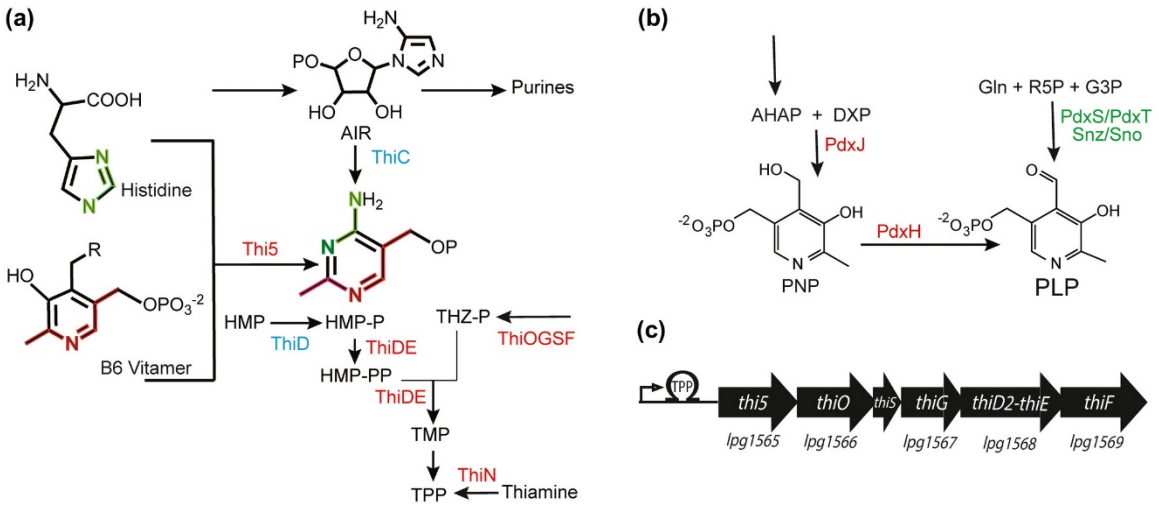


Figure 3.1 – Pathways for HMP-P, PLP biosynthesis. A) Two biosynthetic routes for the synthesis of thiamine pyrophosphate (TPP) are schematically represented with a focus on the source of HMP-P. Gene products are indicated by the reactions they catalyze. Those depicted in blue are present in *S. enterica* and not *L. pnuemophila.* Those in red, are found in *L. pnuemophila* and other organisms as described in the text. In *S. enterica,* ThiC converts 5'-aminoimidizole ribotide (AIR) into HMP-P, while Thi5 synthesizes HMP-P from histidine and a B_6 vitamer. Incorporation of atoms derived from histidine and the B_6 vitamer into HMP-P are shown (Ishida *et al.*, 2008, Tazuya *et al.*, 1995, Tazuya *et al.*, 1989). B) Two pathways for PLP biosynthesis are schematically represented with emphasis on the final steps. Enzymes belonging to the DXP-dependent pathway are highlighted in red, and exemplified by *S. enterica*. PdxH is present in many organisms as part of a salvage pathway and its presence does not indicate the DXP-dependent pathway is present. The enzymes of the DXP-independent pathway are highlighted in green and exemplified by *S. cerevisiae* (Snz/Sno) and *B. subtilis* (PdxS/PdxT). C) Organization of the thiamine biosynthetic locus in *L. pneumophila*. The putative annotations and locus tags for coding regions homologous with other thiamine biosynthetic enzymes are shown. The TPP riboswitch in the 5'UTR of *lpg1565* is indicated (Sahr *et al.*, 2017). Abbreviations: 3-amino-1-hydroxyacetone phosphate, AHAP; deoxyxylulose-5-phosphate, DXP; glutamine, Gln; ribose-5'-phosphate, R5P; glyceraldehyde-3-phosphate, G3P; pyridoxine-5'-phosphate, PNP; pyridoxal-5'-phosphate, PLP; 5'-aminoimidizole ribotide, AIR; 2-methyl-4-amino-5-hydroxymethylpyrimidine, HMP; 4-amino-5-hydroxymethylpyrimidine phosphate, HMP-P; 4-amino-5-hydroxymethylpyrimidine diphosphate, HMP-PP; 4-methyl-5-β-hydroxyethylthiazole, THZ; 4-methyl-5-β-hydroxyethylthiazole phosphate, THZ-P; thiamine, Thi; thiamine diphosphate, TPP; thiamine phosphate, TMP.

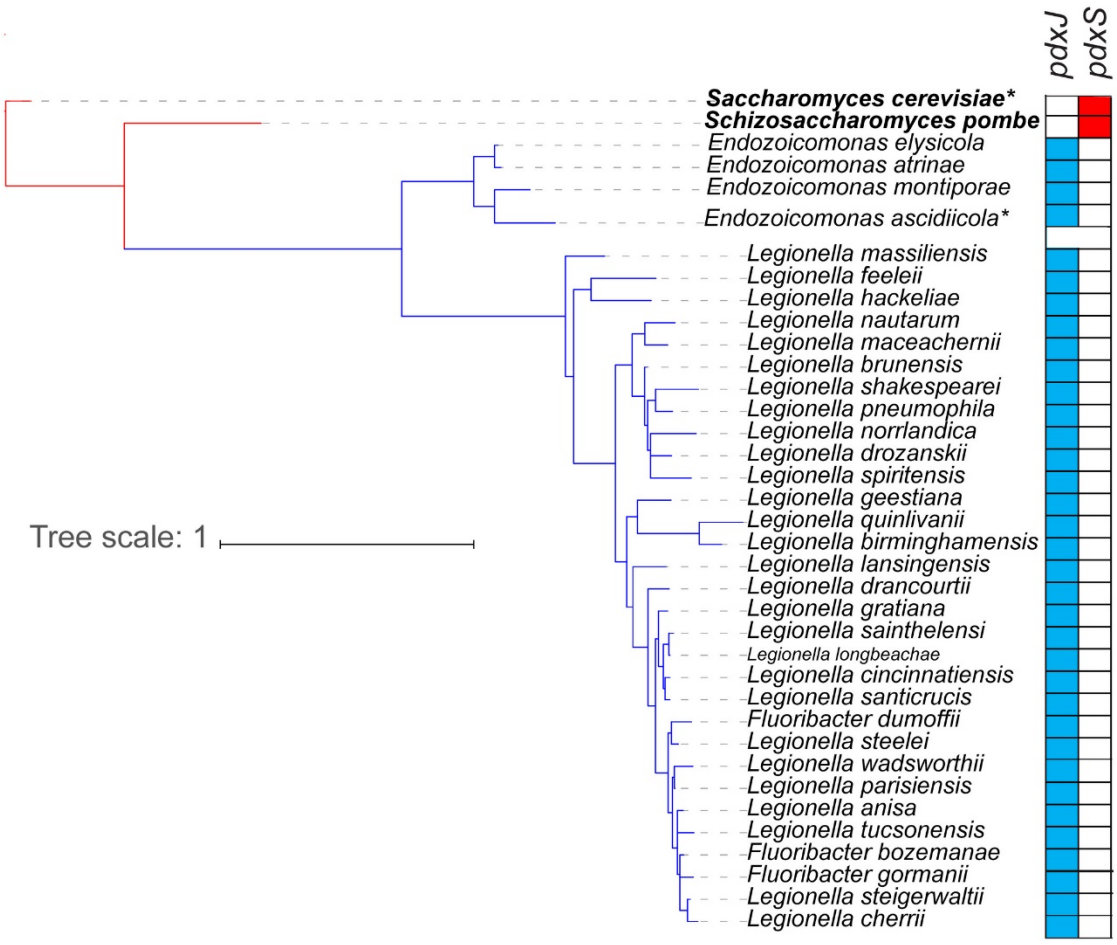


Figure 3.2 –Phylogeny of Thi5 homologs and associated PLP biosynthetic pathways. The phylogenetic relationship of Thi5p homologs among Bacteria (blue lines) and select Eukaryotes (red lines) in the IMG database containing the CCCXC motif. Annotated KEGG functions for *pdxJ* (K03474; E.C. 2.6.99.2) and *pdxS* (K06215; E.C. 4.3.3.6) in each genome are highlighted. Organisms in which mutants lacking the homolog are demonstrable HMP auxotrophs are highlighted in bold. Organisms which contain multiple Thi5 homologs are annotated with an asterisk. The scale represents percent genetic distance between homologs.

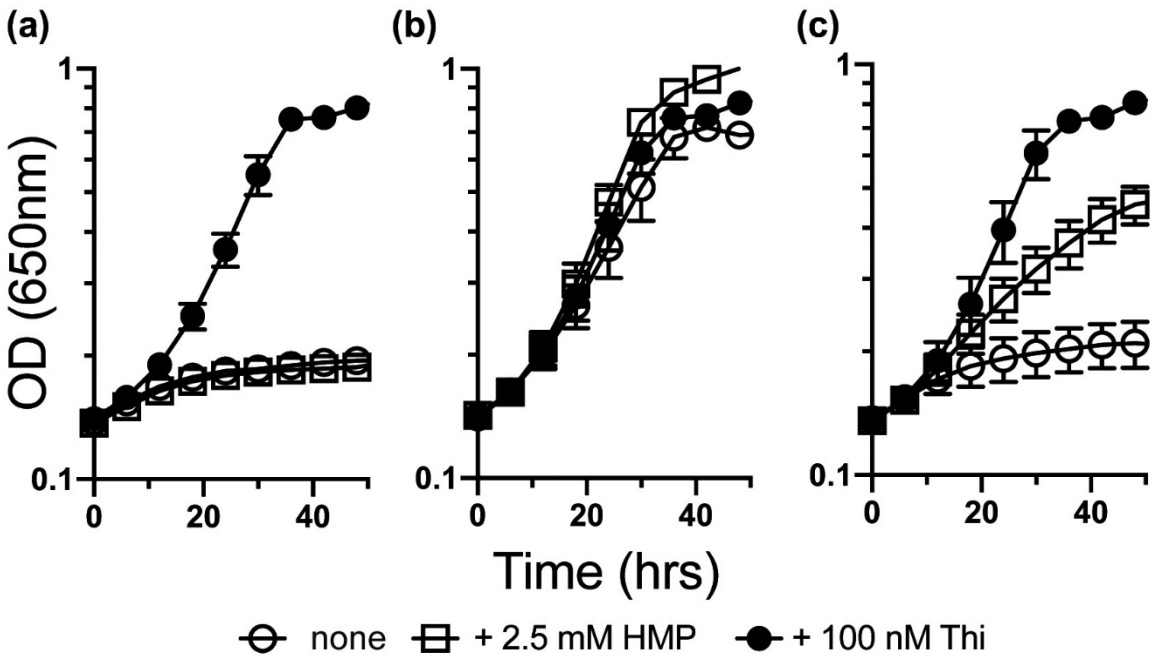


Figure 3.3 – *lpg1565* contributes to HMP synthesis in *L. pneumophila.* A mutant of *L. pneumophila* with an insertion-deletion of *lpg1565* containing (A) pJB98 (vector only), (B) pDM1632 (pJB98-*lpg1565-9*), or (C) pDM1633 (pJB98-*lpg1566-9*) was grown on MRM supplemented with i) no additions, i) HMP (2.5 mM) or iii) thiamine (100 nM) as indicated. Error bars indicate the standard deviation of three independent biological replicates.

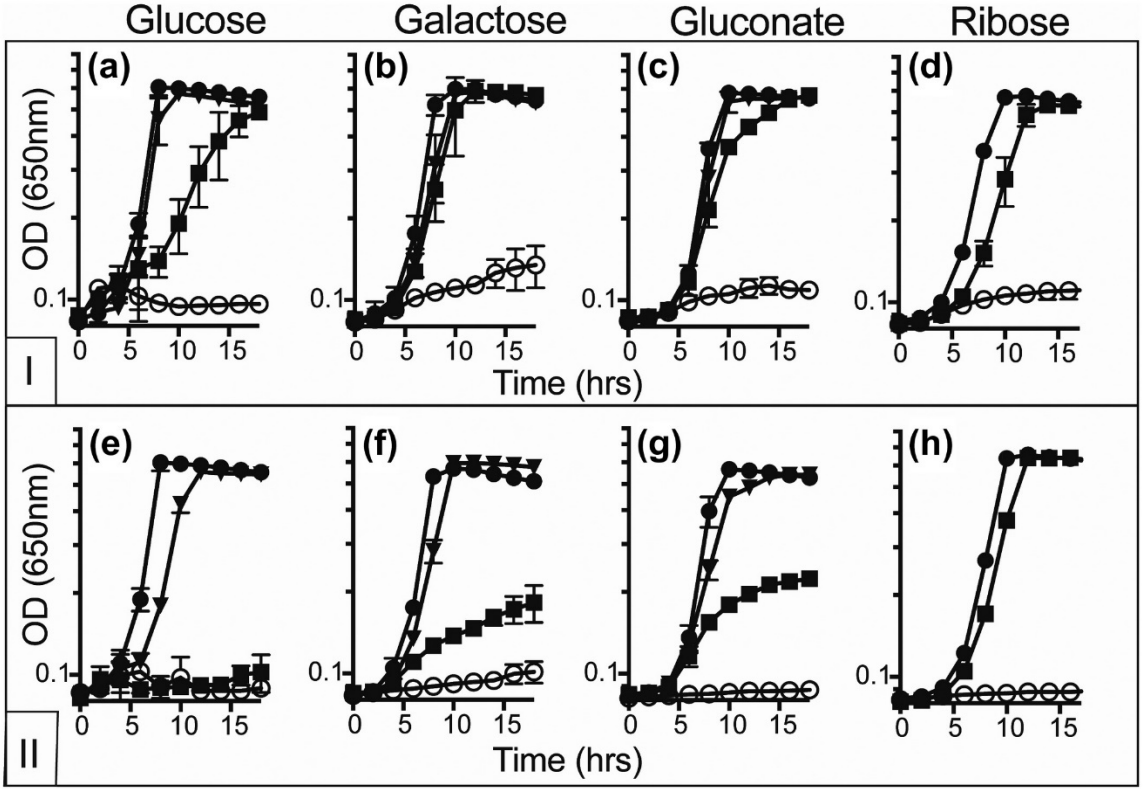


Figure 3.4 – *Lp*Thi5 can generate HMP in *S. enterica.* Strains of *S. enterica* lacking *thiC* and containing either *Lpthi5* (Panel I) or a codon-optimized version of *ScTHI5* (Panel II) were grown on a minimal medium containing either 11 mM glucose (A, E), 11 mM galactose (B, F), 11 mM gluconate (C, G) or 13.2mM ribose (D, H). Medium was otherwise supplemented with nothing (open circles), 100 μM IPTG (solid squares), 100 μM IPTG and 100 nM THZ (solid inverted triangles) or 100 nM thiamine (solid circles). Growth was measured by optical density at 650 nm over time. Error bars indicate the standard deviation of three independent biological replicates.

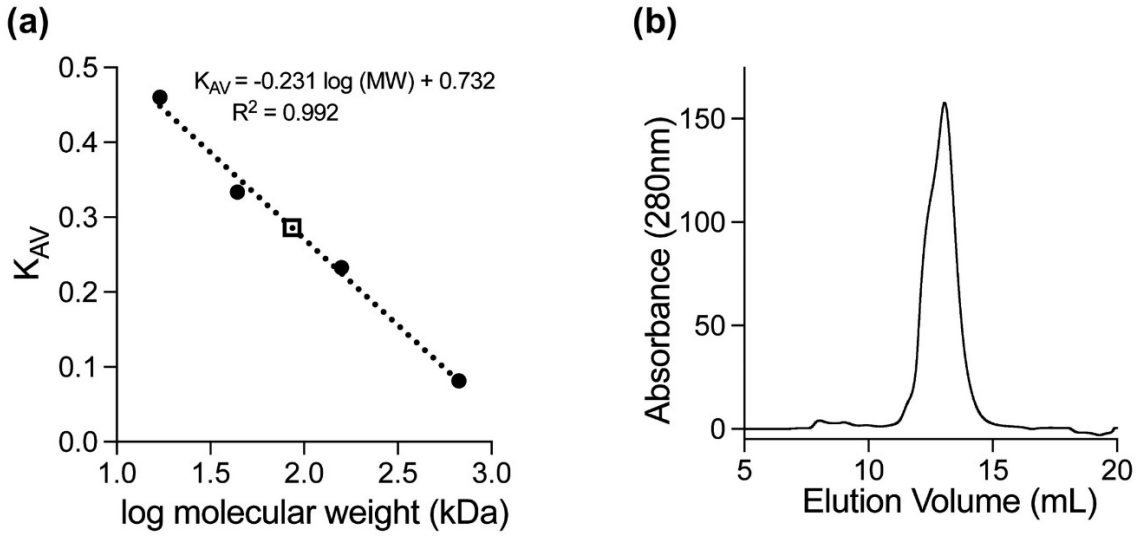


Figure 3.5 – *Lp*Thi5 is dimeric in solution The oligomeric state of *Lp*Thi5 in solution was determined by size exclusion chromatography. An SEC 650 column was equilibrated with 50 mM HEPES + 1 mM TCEP, pH 7.5. Molecular weight standards or *Lp*Thi5 (50 µL) were injected and eluted at a flow rate of 1 mL/min. A) Aliquots (50 µL) of thyroglobuline (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa), and myoglobin (17 kDa) were used as size standards to generate the standard curve and interpolate the size of *Lp*Thi5. The position of *Lp*Thi5 is represented by a square. B) Absorbance (280 nm) was followed over 20 mL of elution buffer after *Lp*Thi5 was injected on the column. The lone peak eluted at 13 mL.

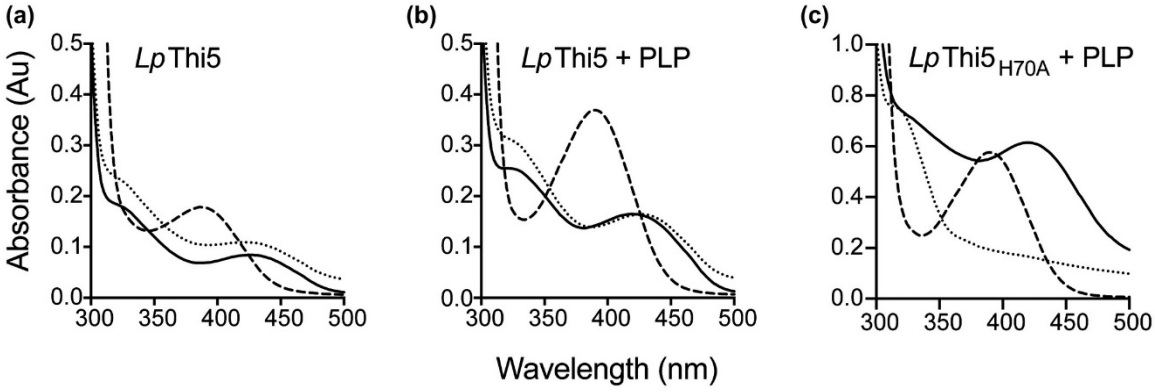


Figure 3.6 – *Lp*Thi5 purifies with pyridoxal-5'-phosphate. Recombinant *Lp*Thi5-His_6 was purified, a portion of the protein preparation was dialyzed against HEPES (50 mM, pH 7.5) containing PLP (200 µM) and TCEP (1 mM) at 4 °C overnight. Unbound PLP was removed using a PD-10 desalting column, and both preparations were characterized by UV-visible spectroscopy. Three spectra were taken of 100 µM *Lp*Thi5 (A), *Lp*Thi5 + PLP (B) and *Lp*Thi5$_{H70A}$ + PLP (C), untreated (solid line), treated with 25 mM $NaBH_4$ (dotted line), treated with 0.1 M NaOH (dashed line).

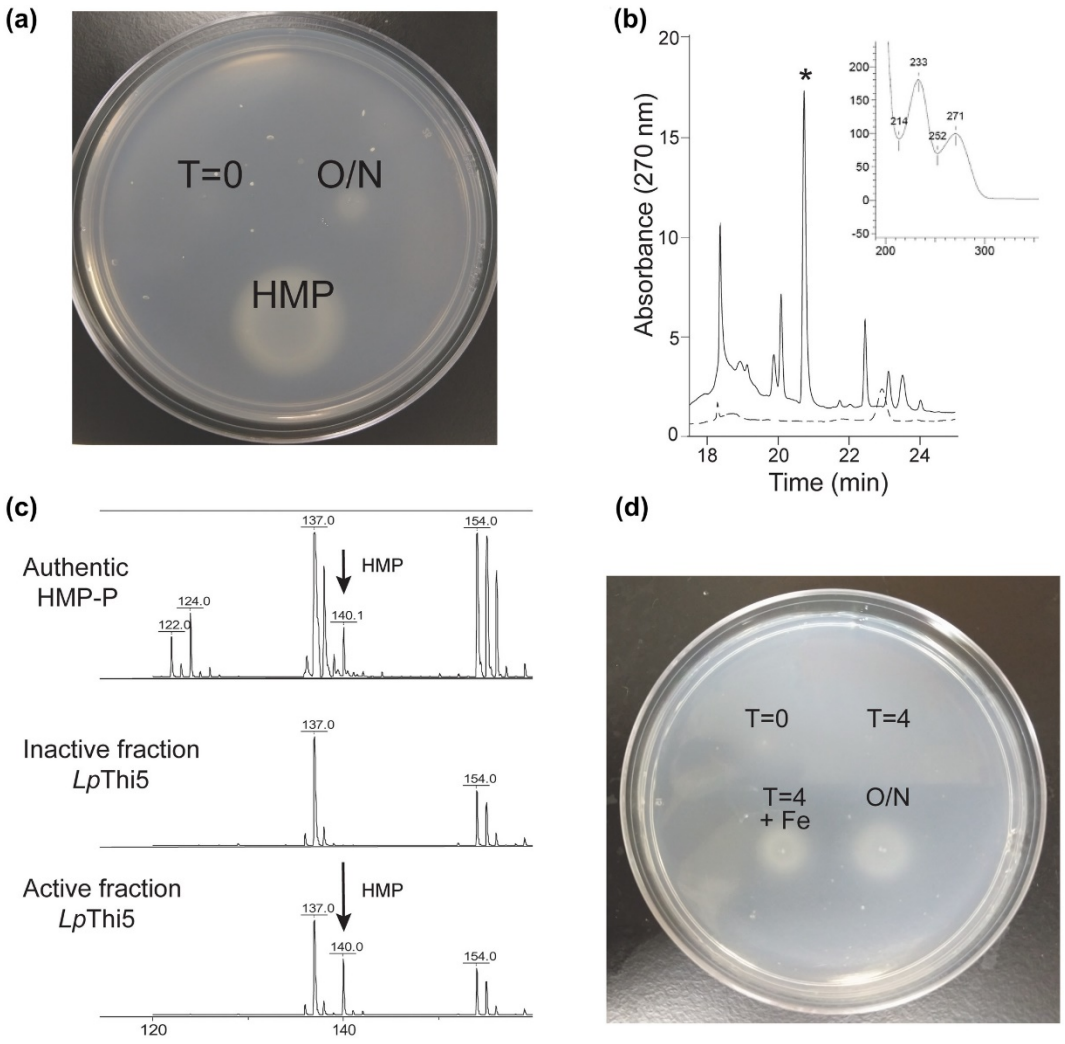


Figure 3.7 – Purified *Lp*Thi5 releases HMP (**A**) *Lp*Thi5 protein in buffer was denatured with heat before incubation (T=0), and after overnight incubation (O/N). In each case the supernatent was spotted on a lawn of a *thiC* mutant of *S. enterica* embedded in soft agar is shown. HMP was spotted as a control. Growth was detected as turbidity after 16 hr incubation at 37 °C. (**B**) *Lp*Thi5 (0.75 mg) was incubated overnight in buffer (black line) or not (hashed line) before denaturing protein at 95 °C for 5 minutes. The supernatents were injected onto an HPLC and absorbance was monitored over time at 270 nm. The star indicates the peak that contained activity in the overnight sample, and the inset shows the UV-Vis spectrum of this peak. (**C**) MALDI-TOF of authentic HMP-P, the biologically active and the inactive fractions from *Lp*Thi5 samples described are shown. (**D**) Impact of iron in the buffer bioassay for the production of HMP(-P) after a four hour incubation with iron.

(a)

```
ScThi5p     1 ----MSTDKITFLLNWQPTPYHIPIFLAQTKGYFKEQGLDMAILEPTNPSDVTELIGSGK
LpThi5      1 MAMSSLKSRVTLLLNWYTNPYHTPILVAQQLGFYAEEGIKLAILEPADPSDVTEIVGLGT

ScThi5p    57 VDMGLKAMIHTLAAKARGFPVTSVASLLDEPFTGVLYLKGSGITEDFQSLKGKKIGYVGE
LpThi5     61 VDFGVKAMIHTVAAKAKGYPVTSIGTLLDEPPTGLIALKSSGIN-SFQDIVGKRVGYIGE

ScThi5p   117 FGKIQIDELTKHYGMKPEDYTAVRCGMNVAKYIIEGKIDAGIGIECMQQVELEEYLAKQG
LpThi5    120 FGKKIIDDLASLAGIDPTSYKTVRIGMNVTDAIYRDVIDTGIGFINFQKVELEHLC----

ScThi5p   177 RPASDAKMLRIDKLACLGCCCFCTVLYICNDEFLKKNPEKVRKFLKAIKKATDYVLADPV
LpThi5    176 ---GETVFLRIDQLAGLGCCCFCSIQFIVPETTL-KQPELVKGFLRATQRGAAYTTEKPE

ScThi5p   237 KAWKEYIDFKPQLNNDLSYKQYQRCYAYFSSSLYNVHRDWKKVTGYGKRLAILPPDYVS-
LpThi5    232 EAYELLCQAKPQLRTPLYQKIFTRTLPFFSRTLINVDRDWDKVGRYTKHLNIIDEHFDIS

ScThi5p   296 -NYTNEYLSWPEPEEVSDPLEAQRLMAIHQEKCRQEGTFKRLALPA
LpThi5    292 QCYTNRFLPDTPYSDL-KP----------IACCLEN---------
```

(b)

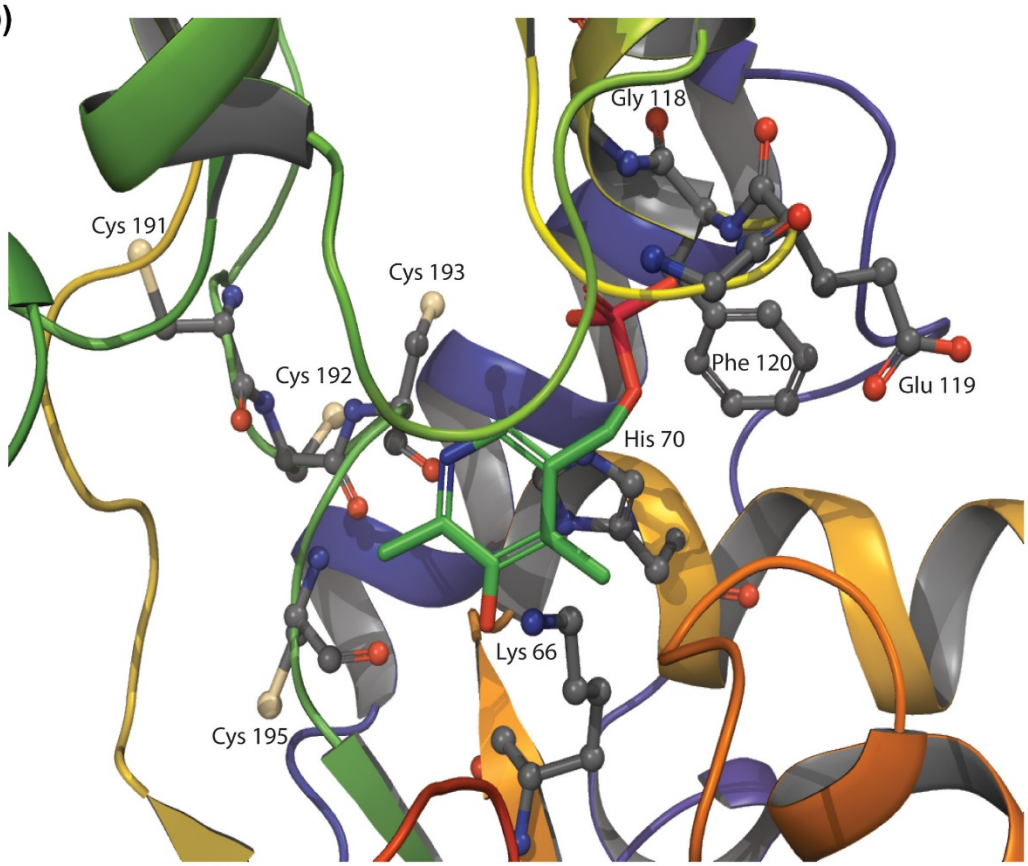


Figure 3.8 – *Lp*Thi5 and *Sc*Thi5p share structural motifs in the active site. (**A**) Amino acid sequences of *Lp*Thi5 and *Sc*Thi5p were aligned by CLUSTAL Omega (Sievers *et al.*, 2011), similar residues are highlighted in grey and identical residues are highlighted in black. Conserved residues interrogated in this study are boxed. (**B**) A structural homology model was generated with Phyre2 (Kelley *et al.*, 2015), and residues in the CCCXC motif or those predicted to coordinate PLP in the active site are modeled as sticks (K66, H70, G118, E119, F120) and visualized with PyMOL.

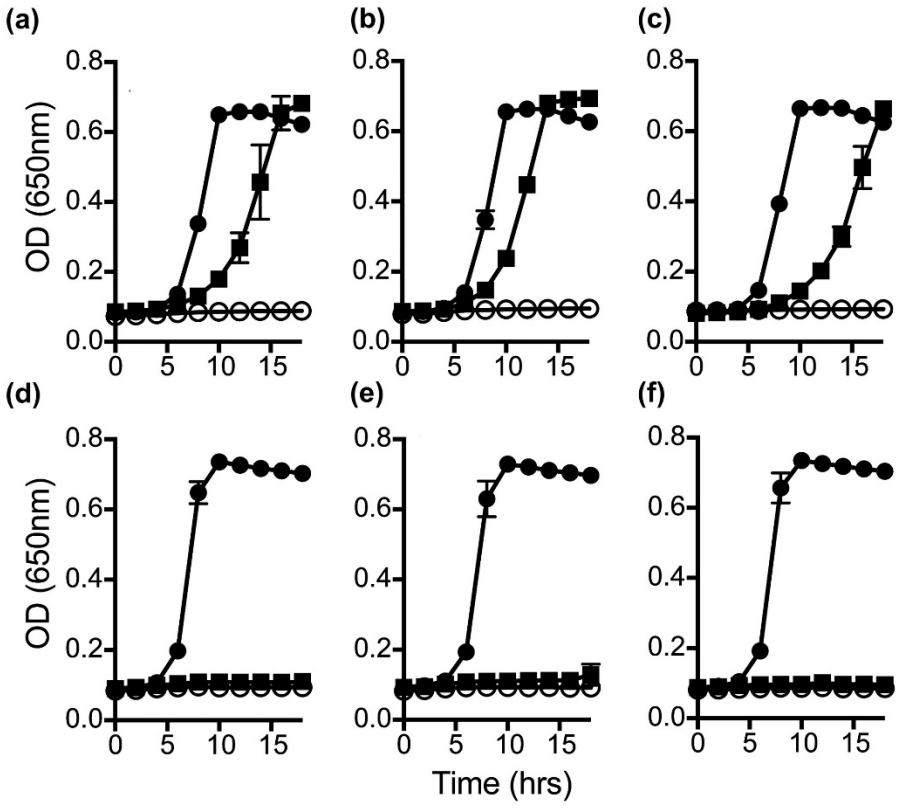


Figure 3.9 – *Lpthi5* variants have reduced ability to complement *thiC* mutant on glucose. Strains of *S. enterica* lacking *thiC* and containing a plasmid expressing *Lpthi5* alleles encoding *Lp*Thi5$_{G118A}$ (A,D), *Lp*Thi5$_{E119A}$ (B,E), or *Lp*Thi5$_{F120A}$ (C,F) under the regulation of P_{tac} were grown on a minimal medium. Panels A-C contained ribose (13.2 mM) as a carbon source, while panels D-F contained glucose (11 mM) as a carbon source. In each case the medium included no additions (open circles), 100 μM IPTG (solid squares) or 100 nM thiamine (solid circles). Growth was determined by following optical density at 650 nm over time. Error bars indicate the standard deviation of three independent biological replicates.

Supplemental Table 3.1 – HMP requirements with different carbon sources

Carbon Source and Additions	Final Cell Yield (OD_{650})
Glucose	0.08 ± 0.01
Glucose + 100 nM THZ	0.09 ± 0.01
Glucose + 10 nM HMP	0.28 ± 0.01
Glucose + 20 nM HMP	0.47 ± 0.03
Glucose + 100 nM THZ + 10 nM HMP	0.53 ± 0.01
Ribose	0.08 ± 0.01
Ribose + 10 nM HMP	0.42 ± 0.01

Final cell yield (OD_{650}) was recorded after 15 hours of growth with shaking at 37 °C. Values are averages ± standard deviations of three biological replicates.

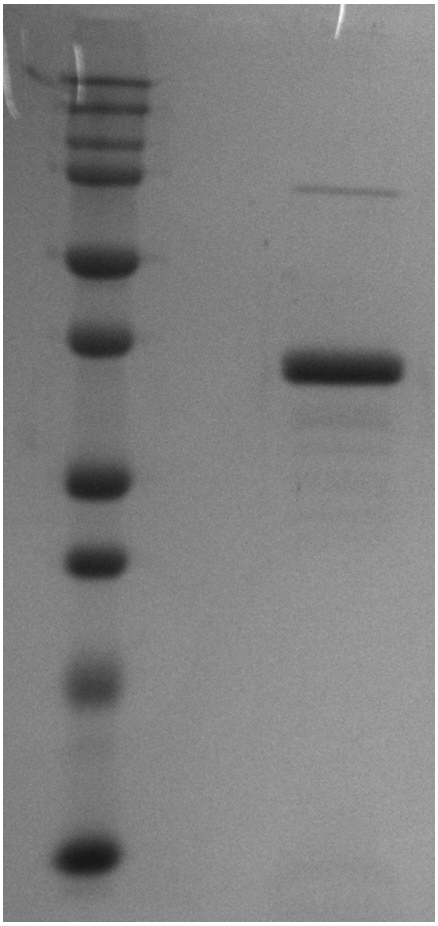


Figure 3.S1 – Representative *Lp*Thi5-His$_6$ protein purification

After purification by Ni-NTA chromatography, 1.5 µg of purified protein was denatured in SDS-PAGE loading dye (60 mM Tris pH 6.8, 0.1 M DTT, 2 % SDS, 10 % Glycerol) by incubating at 95 °C for 10 minutes, separated by SDS-PAGE using a 12 % acrylamide gel, stained with Coomassie Brilliant Blue, and purity was determined by TotalLab Quant v11 densitometry software. *Lp*Thi5 was enriched to > 85 % purity.

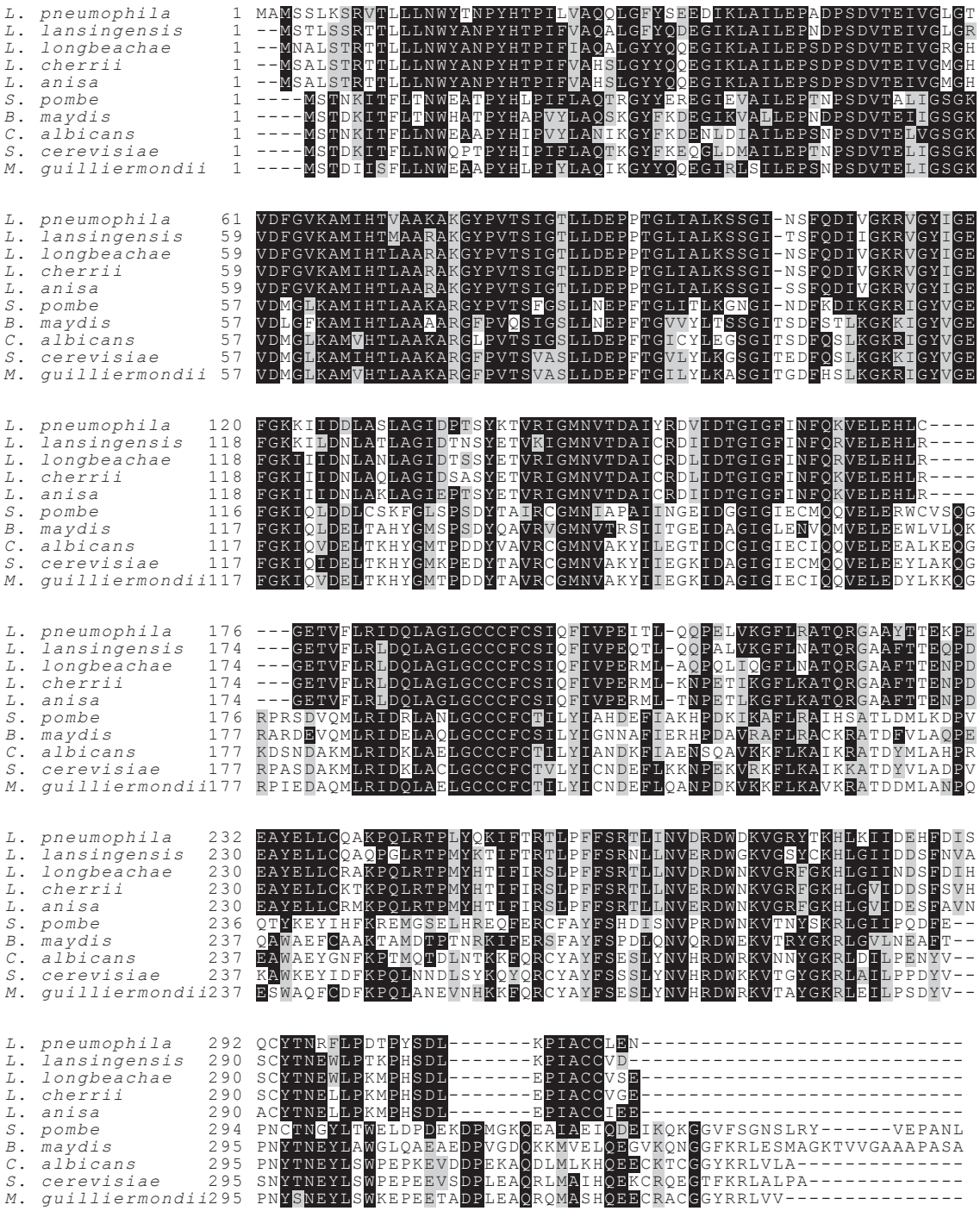


Figure 3.S2 – CLUSTAL Omega alignment of diverse *Legionella* and Eukaryotic Thi5

Diverse Thi5 orthologs from the Eukaryotic clade and the *Legionella* clade of Bacteria were aligned using the CLUSTAL Omega algorithm [Sievers, 2011] to highlight conserved, similar, and non-conserved structural features across these two domains. Strictly conserved residues are highlighted in black, similar residues are highlighted in gray, and non-conserved residues are annotated with a white background.

CHAPTER 4

EVIDENCE FOR A ROLE OF α-KETOGLUTARATE LEVELS IN THE FUNCTION OF THE HETEROLOGOUS Thi5 PATHWAY IN *SALMONELLA ENTERICA*[3]

[3] Paxhia MD and DM Downs. To be submitted to *Mol Microbiol*

4.1 ABSTRACT

Metabolism is structured as modules, however work in metabolic engineering has shown that errors can occur when transferring pathways between organisms. Previously it was shown that the Thi5-dependent pathway(s) from *Saccharomyces cerevisiae* and *Legionella pneumophila* had functional differences when incorporated into the metabolic network of *Salmonella enterica*. Function of *Sc*Thi5 required modification of the underlying metabolic network through restricting glycolysis, but how this led to *Sc*Thi5 function was unclear. Here we report a connection between alpha-ketoglutarate (αKG) and Thi5 function under multiple growth conditions. Two classes of genetic suppressors link metabolic flux or metabolite pool changes with Thi5 function. Directly modulating nitrogen assimilation through nutritional or genetic modification implicates αKG levels in Thi5 function, and addition of exogenous pyridoxal improves *Sc*Thi5 function. Finally, increasing αKG and PLP directly through supplementation improves both *Sc*Thi5 and *Lp*Thi5^{F120A} function. This suggests structural differences between *Sc*Thi5 and *Lp*Thi5 that are important for function and implicates αKG in supporting function of the Thi5 pathway in a heterologous context.

Importance

Thiamine biosynthesis is a model metabolic node that has been used to extend our understanding of metabolic network structure. The requirements for *in vivo* function of the Thi5-dependent pathway found in *Legionella* and yeast are poorly characterized. Here we demonstrate that αKG modulates function of the Thi5 pathway in *S. enterica* and provide

evidence that structural differences between *Sc*Thi5 and *Lp*Thi5 contribute to function in a heterologous context.

4.2 INTRODUCTION

The function, or failure to function, of metabolic modules across different organisms is an understudied area of metabolism, and one with implications for defining the acquisition of new metabolic capabilities by horizontal gene transfer or genetic engineering. A current view is that modules evolve in a specific metabolic network and are therefore optimized to function within the framework of the native organism. When moved, by either horizontal gene transfer or metabolic engineering, incompatibilities can occur and prevent optional function (Cardinale and Arkin, 2012; Kittleson *et al.*, 2012). These incompatibilities can reflect a depletion of metabolites, accumulation of toxic metabolites, inhibition of enzymes or pathways, lack of substrates, etc. Some of these incongruities can be solved by altering either the metabolic network of the host, or the module itself to restore a functional system (Close *et al.*, 2019). Identifying the changes needed to generate a functioning system can result in insights about existing metabolic integration and are facilitated by focusing on a defined node of metabolism. Thiamine biosynthesis is a well-defined metabolic network that has been used to investigate the plasticity of metabolism in *Salmonella enterica* (Downs, 2006; Koenigsknecht and Downs, 2010)].

Thiamine pyrophosphate (TPP), the biologically active form of Vitamin B_1, is an essential cofactor for nearly all organisms (Zhang *et al.*, 2016). It is required throughout central carbon metabolism and secondary metabolite biosynthesis primarily to facilitate decarboxylation and ligase reactions (Pohl *et al.*, 2004). TPP is produced by bacteria,

archaea, yeast and plants, where the two moieties 4-methyl-5-(2-hydroxyethyl)-thiazole-phosphate (THZ-P) and 4-amino-5-hydroxymethyl-2-methylpyrimidine-pyrophosphate (HMP-PP) are independently synthesized and combined to form thiamine monophosphate (TMP) (Jurgenson *et al.*, 2009). In a final step, TMP is phosphorylated to generate TPP (Jurgenson *et al.*, 2009). There have been two pathways described for the synthesis of the HMP-P moiety. The majority of plants, archaea and bacteria use a ThiC-dependent mechanism that converts the purine biosynthetic intermediate, 5-aminoimidazole ribotide, into HMP-P, formate and carbon monoxide (Chatterjee *et al.*, 2008; Martinez-Gomez and Downs, 2008; Chatterjee *et al.*, 2010) A second pathway was first described in yeast and subsequently found in the Legionellaceae clade of bacteria (Tazuya *et al.*, 1989; Tazuya *et al.*, 1995; Ishida *et al.*, 2008; Paxhia *et al.*, 2020). Labeling studies in *Saccharomyces cerevisiae* found that the atoms in the pyrimidine moiety of thiamine originated from histidine and a B_6 vitamer, presumably PLP (Tazuya *et al.*, 1989; Tazuya *et al.*, 1995; Ishida *et al.*, 2008). Genetic analyses in *S. cerevisiae* determined that genes of the *THI5* family (*THI5/11/12/13*) supported HMP synthesis, with a single isozyme being sufficient for growth (Wightman and Meacock, 2003).

The response of a metabolic network to a module exchange was probed in a study in which *ScTHI5* was introduced into *Salmonella enterica* as a substitution for the native ThiC. The study showed that *ScTHI5* alone was not broadly able to substitute for ThiC in the synthesis of HMP-P. Rather, Thi5p expressed *in trans* was sufficient to compliment the *thiC* strain of *S. enterica* only under specific limited growth conditions (Palmer *et al.*, 2015a)otably, Thi5p was unable to support thiamine synthesis when glucose was the sole carbon source, a phenotype unaltered by cAMP, histidine or PL addition. These data

indicated that an undefined metabolic change was required to integrate *Sc*Thi5 into the metabolic network of *S. enterica*. Supporting this notion, alleles of *sgrR* and *ptsI* were identified that restored *Sc*Thi5 function on glucose (Palmer *et al.*, 2015a). Genetic analysis supported the conclusion that the isolated allele of *sgrR* led to constitutive expression of *sgrS*, likely leading to reduced transport of glucose by PtsG and reduced glycolytic flux. Consistently, *Sc*Thi5 function in a *ptsG* background was dependent on *sgrR*, but the mechanism of activating or increasing *Sc*Thi5 was not determined. Somewhat surprisingly, *Lp*Thi5 from *Legeionella pneumophila* restored growth to a *thiC* mutant in multiple growth conditions, including glucose (Paxhia *et al.*, 2020). Variants of *Lp*Thi5 were identified that generally phenocopied the behavior of *Sc*Thi5, leading to the hypothesis that differences between the orthologs may contribute to function in a heterologous context (Paxhia *et al.*, 2020).

Knowledge of thiamine metabolism, and ease of genetic manipulation, makes *S. enterica* a model to probe the metabolic consequences of incorporating a heterologous module (Thi5) into a metabolic network. This study was initiated to further probe the mechanism that allows *Sc*Thi5 to function in *S. enterica.* Data herein identified two metabolites, pyridoxal and alpha-ketoglutarate (αKG), that impact the function of ScThi5 *in vivo.* Results presented are consistent with a model in which glucose-specific suppressors act by decreasing flux through glycolysis resulting in increased alpha-ketoglutarate (αKG) via the TCA cycle and increased flux through the pentose-phosphate pathway leading to increased PLP synthesis. Additional growth conditions and mutations identified herein that allow *Sc*Thi5 function increase αKG and/or PLP by other mechanisms. Excitingly, these metabolites appear to also facilitate function of a *Lp*Thi5

variant with reduced *in vivo* activity, suggesting structural differences in orthologs can reflect their evolution as part of a metabolic network with specific features.

4.3 RESULTS AND DISCUSSION

The finding that a bacterial Thi5 had the capacity to complement a *S. enterica thiC* mutant, while the metabolic network had to be modified to allow *Sc*Thi5 to function raised questions about i) differing metabolic networks in bacteria vs yeast, and ii) features of the Thi5 orthologs responsible for their differential function *in vivo*. Two mutations previously isolated that allow *Sc*Thi5-dependent growth of a *thiC* mutant with glucose as a carbon source were further characterized to address the former.

A variant of PtsI allows *Sc*Thi5 to function in *S. enterica*

A mutant allele of *ptsI* encoding $PtsI^{R361H}$ supported *Sc*Thi5 function on glucose (Figure 4.1) independent of *sgrR* (Palmer *et al.*, 2015a). PtsI, or Enzyme I (EI) of the phosphotransferase system (E.C. 2.7.3.9), is a regulatory node in metabolism that coordinates flux between glycolysis and the TCA cycle in response to alpha-ketoglutarate (αKG), phosphoenolpyruvate (PEP), and pyruvate levels (Huergo and Dixon, 2015). The substitution in $PtsI^{R361H}$ is located outside the active site and in a region of the enzyme that has a large $^{1}H_{N}/^{15}N$ chemical shift when PEP or αKG binds, suggesting the variant could have altered PEP turnover or regulation by αKG (Venditti *et al.*, 2013).

Both His_6-PtsI and His_6-$PtsI^{R361H}$ were purified and characterized *in vitro*. His_6-PtsI hydrolyzed PEP with a K_m of 0.26 ± 0.02 mM and a V_{max} of 235 ± 8.1 pmol/min (Figure 4.2; Table 4.2). His_6-$PtsI^{R361H}$ had a K_m of 0.011 ± 0.001 mM for PEP, and a V_{max} of 67 ±

0.7 pmol/min, indicating the variant had a higher affinity of PEP and a significant decrease in maximal turnover. αKG modulates the activity of PtsI, and the two proteins were assayed in the presence of different concentrations of αKG to determine the inhibitory constant, K_i. The K_i for αKG was slightly increased in the variant over wildtype (6.8 ± 0.3 mM αKG, 4.16 ± 0.34 mM, respectively) (Figure 4.2; Table 4.2). Based on these data, αKG was modeled as an inhibitor of His_6-PtsI and His_6-$PtsI^{R361H}$ (Supplemental Figure 4.2). Non-competitive inhibition appropriately modeled the effect of αKG on both His_6-PtsI and His_6-$PtsI^{R361H}$, consistent with a previous report (Doucette *et al.*, 2011). In total these data identified a decrease in maximal turnover of the $PtsI^{R361H}$ variant that could restrict flux through glycolysis, suggesting suppression occurred by a mechanism similar to that defined for the *sgrR* mutant (Palmer *et al.*, 2015a). Disruption of glycolytic flux has been shown to increase flux through the TCA cycle, and the pentose-phosphate pathway, suggesting potential changes responsible for the *Sc*Thi5 function in *S. enterica* (Hollywood and Doelle, 1976; Long *et al.*, 2018; Long and Antoniewicz, 2019).

Loss of *yggS* allows ScThi5 function

A previously identified mutant with restored *Sc*Thi5-dependent growth of a *S enterica thiC* mutant on glucose (Palmer *et al.*, 2015a) was further characterized. The strain did not have a mutation linked to the previously identified suppressors in *sgrR* or *ptsI*. This suppressor was of particular interest because it allowed *Sc*Thi5-dependent growth of a *S. enterica thiC* mutant on multiple carbon sources (e.g., gluconate, galactose and glycerol) in addition to glucose. This feature distinguished it from the *ptsI* allele above and other characterized

suppressors that had an effect only on glucose medium (Palmer *et al.*, 2015a), and putatively acted by decreasing flux through glycolysis. Whole genome sequencing of this suppressor identified multiple SNPs, but focus was drawn to an allele of *yggS* that was predicted to encode $YggS^{Q51Stop}$. Standard genetic mapping analyses using an insertion linked to the *yggSTU* operon (*nupG*) confirmed the causative mutation was linked to this locus.

The gene of unknown function, *yggS*, is a member of the COG0325 family and is conserved across all domains of life (Prunetti *et al.*, 2016). The YggS protein purifies with PLP, and mutations in the human homolog, PLPBP, are associated with Vitamin B_6-dependent epilepsy (Ito *et al.*, 2013; Shiraku *et al.*, 2018). An isogenic pair of strains (with and without the *yggS* lesion) were constructed in a *ΔthiC1225 ΔaraCBAD* background. Only the strain with a mutant allele of *yggS* allowed *Sc*Thi5-dependent growth on glucose, galactose and gluconate (Figure 4.3). Insertion-deletions of *yggS, yggT* and *yggU* were constructed and transduced into the relevant *thiC*/pScThi5 strain (DM16449). Of the three mutations, only the deletion of *yggS* allowed *Sc*Thi5-dependent growth, indicating the suppression phenotype was not due to a polar effect on a downstream gene(s) (data not shown).

The data showed that lack of a functional YggS was sufficient to allow *Sc*Thi5 to provide sufficient HMP for growth on all carbon sources tested. Subsequent selection for suppressors on other carbon sources (i.e., pyruvate, galactose) that allowed *Sc*Thi5-dependent growth yielded multiple alleles of *yggS*. Thus, the *yggS* locus appeared to be the predominant site of mutations that restored *Sc*Thi5-dependent growth in *S. enterica* on non-

glucose carbon sources. Previous work in our lab and others has identified pleotropic effects caused by the elimination of *yggS* in *S. enterica* and *E. coli* (Ito *et al.*, 2013; Prunetti *et al.*, 2016; Ito *et al.*, 2019; Ito *et al.*, 2020; Vu *et al.*, 2020). Relevant to the work herein, a lesion in *yggS* disrupts PLP homeostasis and is reported to alter a variety of metabolite pools (i.e., amino acids and keto acids).

Increased αKG allows *Sc*Thi5 function in *S. enterica*

Together the data above suggested that altering the metabolic network, via flux changes and/or metabolite pool sizes influenced the function of *Sc*Thi5 in *S. enterica.* Additional means to alter the metabolic network were explored. The standard minimal medium in the laboratory contains ammonia (16 mM) as the sole source of nitrogen. When glutamine (1 mM) was provided as the sole nitrogen source, *Sc*Thi5 function became sufficient for full growth of a *thiC* mutant on glucose, galactose or gluconate (Figure 4.4).

Changing the nitrogen source impacts flux in multiple pathways and changes the level of numerous metabolites. For a number of reasons αKG was identified as a metabolite of interest. αKG levels in the cell are dynamic and increase when *E. coli* is grown on nitrogen sources other than ammonia, or when nitrogen assimilation by either glutamate dehydrogenase or glutamate synthase are perturbed (Yan *et al.*, 2011; Huergo and Dixon, 2015). In a series of experiments, αKG levels were modulated by nutritional shifts or genetic perturbations and the resulting impact on *Sc*Thi5 function assessed. αKG could not be added exogenously to raise internal levels since *S. enterica* does not have a functional *kgtP,* the gene encoding the αKG transporter which has been manipulated in *E. coli.* (Yan *et al.*, 2011; You *et al.*, 2013). Dimethyl-ketoglutarate (dmKG) is a membrane-diffusible

metabolite that is converted into αKG by esterases in the cell (Doucette *et al.*, 2011). Titration of dmKG on glucose medium (0 – 20 mM) determined that 10 mM was the optimal concentration of this compound to support growth of the *thiC*/p*Sc*Thi5 strain on minimal glucose medium (Figure 4.5), while higher concentrations were inhibitory. Addition of dmKG (10 mM) also improved growth when galactose or gluconate were carbon sources (data not shown). The addition of ketobutryate, a ketoacid close in structure reported to be elevated in a *yggS* strain (Ito *et al.*, 2013), did not support *Sc*Thi5 function on glucose suggesting that *Sc*Thi5 function requires a subset of the metabolic changes caused by a *yggS* lesion.

A mutation in *gdhA* (glutamate dehydrogenase) was introduced into the *thiC*/p*Sc*Thi5 strain as an independent means to increase the cellular pool of αKG (Yan *et al.*, 2011). Consistently, the Δ*gdhA thiC*/p*Sc*Thi5 strain (DM17115) showed significant growth on minimal medium with ammonia as a nitrogen source when either glucose, galactose or gluconate were sole carbon sources (Figure 4.6). In total, results of these three experiments were consistent with the hypothesis that increasing αKG in *S. enterica* resulted in increased activity of *Sc*Thi5, as judged by the level of growth the protein supported.

Exogenous pyridoxal enhances *Sc*Thi5 function

Data above showed that the metabolic changes caused by using glutamine as sole nitrogen source, or a lesion in *yggS*, restored full growth to the *thiC* /pScThi5 (Figure 4.4) while a Δ*gdhA* mutation or dmKG supplementation only partially restored growth (Figure 4.6; Figure 4.5, respectively). These data suggested that utilizing glutamine as a nitrogen source, or eliminating *yggS,* not only increased αKG pools but altered another metabolic

feature necessary to allow *Sc*Thi5 activity sufficient for full growth. PLP is a substrate of Thi5 and past models suggested function of *Sc*Thi5 in *S. enterica* was limited by this metabolite or its delivery (Paxhia and Downs, 2019). When 100 nM pyridoxal (PL) was provided to the medium, limited growth was allowed on glucose. However, when both dmKG and PL were provided, near complete growth was achieved (Figure 4.5). Further, addition of PL dramatically enhanced the growth of the Δ*gdhA thiC* /pScThi5 strain on glucose (Figure 4.6). Thus, these data extended the model to suggest the metabolic network of *S. enterica* required both αKG and PLP be increased to support sufficient activity of *Sc*Thi5 to allow full growth. Further, the additive effect of the two metabolites suggested they had independent effects, each of which contributed to increased activity. When growing on glucose, constricted flux through glycolysis would increase αKG and PLP levels by resulting in increased flux through the TCA cycle (αKG) and pentose-phosphate pathway (generating erythrose-4-phosphate, a precursor to PLP) (Hollywood and Doelle, 1976; Long *et al.*, 2018; Long and Antoniewicz, 2019). While a stimulatory role for PLP might be anticipated by its role as a substrate, it is unclear if levels of αKG influence *Sc*Thi5 function in a direct or indirect manner. An important consideration in this working model is the fact that *S. enterica* strains lacking *yggS* have increased PLP and have increased keto acid levels (Ito *et al.*, 2013; Vu *et al.*, 2020), providing an explanation for the full growth restoration of *yggS* lesion on all carbon sources.

A single substitution in *Lp*Thi5 changes the metabolic network requirement

A bacterial Thi5 (*Legionella pneumophila*; *Lp*Thi5) functions to allow full growth of a *S. enterica thiC* mutant without host modification (Paxhia *et al.*, 2020). While *Sc*Thi5 and

*Lp*Thi5 share 48% amino acid identity, a homology model identified two putative structural differences between the proteins that could lead to differences in active site structure. The amino acid sequence of *Lp*Thi5 was modeled using the Phyre2 pipeline in intensive mode (Kelly *et al.*, 2015). The resultant model was aligned to a crystal structure of *Sc*Thi5 (PDB 4H67 chain D) to determine potential structural differences using Maestro (Schrodinger Inc.). *Sc*Thi5 has an extension of 20 amino acids at the c-terminus compared to *Lp*Thi5. This extension comprises alpha-helix 12 in the crystal structure and could influence the orientation of the two domains of the protein based on a proposed role of this helix in stabilizing Domain 2 (Coquille *et al.*, 2012). Additionally, based on the amino acid alignment and resulting structural alignment, *Lp*Thi5 has an eight amino acid truncation involving residues in alpha helix 6. Lack of these residues may constrict the active site, potentially increasing the affinity of *Lp*Thi5 for PLP which would be consistent with the working model herein.

While mutations in conserved active site residues of *Lp*Thi5 eliminated function, other variants retained the ability to complement an *S. enterica thiC* mutant on ribose, while losing the ability to allow growth of the same strain on glucose (Paxhia *et al.*, 2020). This pattern of growth was notable because it was similar to that allowed by the *Sc*Thi5 protein described herein. Analysis of the *Lp*Thi5^{F120A} variant on gluconate indicated a more significant shift in *in vivo* function. While the wildtype *Lp*Thi5 allowed full growth a *thiC* mutant with gluconate as a carbon source, *Sc*Thi5 only partially restored growth when induced [Paxhia, 2020]. The *Lp*Thi5^{F120A} variant allowed growth of the *thiC* mutant that was similar to that with the *Sc*Thi5 (Figure 4.7). Significantly, the strains, with p*Sc*Thi5 and p*Lp*Thi5 ^{F120A}, responded similarly to dmKG and PL (Figure 4.7). The substitution in

this variant has altered the GEFG motif that is involved in hydrogen bonding with the phosphate group of PLP in the active site. We hypothesize that *Sc*Thi5 has a lower affinity for PLP, which may be compensated for with the PLP delivery system in yeast (*SNZ3*). The additive effect of αKG and PLP suggest this result might implicate a role of αKG in directly or indirectly activating Thi5 enzyme(s) or variants with compromised PLP binding.

4.4 CONCLUSIONS

Previously it was found that incorporation of the Thi5 pathway for HMP-P biosynthesis from *S. cerevisiae* into *S. enterica* required modification of the metabolic network through reducing flux through glycolysis (Palmer *et al.*, 2015). A variety of suppressor mutations accomplished this and included a constitutive variant of SgrR, null mutations in *pgi, pfkA*, or the primary glucose transporter *ptsG* (Palmer *et al.*, 2015). The specific consequences of reducing flux through glycolysis, how this led to Thi5 function, and whether other mechanisms of suppression existed, were not addressed. Here we present evidence that supports a hypothesis that Thi5-dependent thiamine synthesis on multiple carbon sources in *S. enterica* requires increased pools of αKG and PLP. Mutations in *ptsI* reduce flux through glycolysis by reducing maximal turnover and maintaining high affinity for PEP, likely diverting flux through the pentose phosphate pathway and the TCA cycle. A null allele of *yggS* is associated with higher PLP and ketoacid pools, and this lesion leads to Thi5 function on multiple carbon sources.

Three different conditions that increased αKG increased *Sc*Thi5-dependent growth of a *S. enterica* strain. First, by changing nitrogen source from ammonia to glutamine,

*Sc*Thi5 function was restored, establishing a correlation between elevated αKG levels and Thi5 function. Second, nitrogen assimilation on ammonia was changed by eliminating *gdhA.* Both modifications of nitrogen assimilation led to improved *Sc*Thi5 function on multiple carbon sources. The addition of pyridoxal to a *gdhA* strain led to improved complementation by *ScTHI5*, establishing a link between elevated PLP levels and improved Thi5 function *in vivo*. Third, the additive effects of dmKG and pyridoxal supplementation led to *Sc*Thi5 function on glucose. These supplementations also restored function of a defective allele of *Lpthi5* encoding a variant that is assumed to be compromised in PLP binding. The fact that PL alone was not sufficient for *Lp*Thi5^{F120A} function suggests that increased αKG may be directly or indirectly improving flux through the Thi5 pathway. This also suggests that structural differences between *Sc*Thi5 and *Lp*Thi5 could explain complementation differences in *S. enterica*, and that *Sc*Thi5 may have a reduced affinity for PLP compared to *Lp*Thi5. The data herein provides insights into the remodeling of the metabolic network that is needed to allow growth of the heterologous *Sc*Thi5 in *S. enterica.* Going forward this information will be valuable in efforts to define the biochemical mechanism of Thi5 action *in vivo.*

A direct model would implicate αKG in the reaction mechanism of Thi5. The proposed mechanism for Thi5 included an oxidation of the PLP and histidine substrates (Lai *et al.*, 2012). Several oxygenases exist that facilitate oxidation of their substrates by using keto acids like αKG as co-substrates. These mechanisms are found in diverse biochemical pathways but thus far no examples that incorporate the co-substrate into the final product have been described (Herr and Hausinger, 2018). Importantly, several of these enzymes are oxygen sensitive, and require enzymes purified under aerobic conditions to

be reconstituted anaerobically with the addition of strong reducing agents like dithionite. For example, when purified aerobically, enzymes like TauD spontaneously auto-oxidize and damage residues within the active site (Ryle *et al.*, 2003; Hausinger, 2004). This characteristic is consistent with the previous report of an oxygen-sensitive reaction mechanism for Thi5 *in vitro* (Lai *et al.*, 2012).

Moving pathways between organisms can lead to incompatibilities caused by accumulation of toxic intermediates or depletion of molecules necessary for either the pathway of interest or a competing reaction (Cardinale and Arkin, 2012; Kittleson *et al.*, 2012). Fortunately, categorizing errors and determining their cause and potential solutions not only increases knowledge about a particular node of metabolism, but also improves knowledge on the construction of metabolic networks as a whole. Further work toward understanding how αKG impacts Thi5 function may lead to a deeper knowledge on the differences in network construction between *L. pneumophila*, *S. enterica*, and *S. cerevisiae*, and may start to determine why certain organisms use Thi5 over ThiC in the biosynthesis of thiamine-pyrophosphate.

4.5 MATERIALS AND METHODS

Strains, media and chemicals

All strains of *S. enterica* are derived from strain LT2 and their relevant genotypes are described in Table 4.1. Rich media used in this study included Nutrient Broth (NB; 8 g/L Difco Nutrient broth and 5 g/L NaCl), or super broth (SB; 32 g/L tryptone (Fisher Scientific), 20 g/L yeast extract (Fisher Scientific), 5 g/L NaCl with 0.05 N NaOH). Solid media contained 15 gm/L agar. Antibiotic concentrations in rich media are as follows:

Ampicillin (Amp) - 100 mg/L; Chloramphenicol – 20 mg/L; Kanamycin – 50 mg/L; Tetracycline - 20 mg/L. Minimal medium was no-carbon E medium (NCE) (Vogel and Bonner, 1956) with 1 mM $MgSO_4$, 0.1 x trace minerals (Balch *et al.*, 1979), with glucose, galactose or gluconate as a sole carbon source (11 mM). No carbon or nitrogen (NCN) (RW *et al.*, 1980) salts with glutamine (1 mM) as a nitrogen source and indicated carbon source at 11 mM was used as low nitrogen medium. Thiamine and pyridoxal were added as indicated to final concentration of 100 nM. IPTG was purchased from Gold Biotechnology, (St. Louis, MO), other chemicals were purchased from Sigma-Aldrich, (St. Louis, MO).

Genetic methods

Suppressor mutants (*ptsI611, yggS661*) were isolated in a *thiC /pSc*Thi5 strain of *S. enterica* on minimal glucose medium after DES mutagenesis as reported (Palmer *et al.*, 2015a). The causative mutation in each strain was identified by Illumina whole-genome sequencing, linked to transposon Tn10d(Tc) using standard genetic techniques and reconstructed into the DM13337 (*ΔthiC1225 ΔaraCBAD*) background using the high-frequency generalized transducing mutant of bacteriophage P22 (HT105/1, *int*-201) (Schmieger, 1972), as previously described (Ernst *et al.*, 2014; Palmer *et al.*, 2015). Insertion-deletion mutations of *yggS*, *yggT*, and *yggU* were constructed by Lambda Red recombineering using the primers in Table 4.1 (Datsenko and Wanner, 2000). Other mutations were moved by transduction into the DM13337 background using standard genetic approaches as needed.

Growth Analysis

Growth of *S. enterica* strains were monitored at OD_{650} in 96 well plates with a BioTek ELx808 plate reader. Strains were grown for 16-20 hours in NB Amp (2 mL) at 37 °C before pelleting and resuspension in an equal volume of 0.85 % NaCl. The cell suspension was used to inoculate (1 %) the appropriate media (200 µL). Plates were incubated at 37 °C with medium shaking and data were plotted using Prism 8 (GraphPad).

Molecular techniques

Plasmids were constructed using standard molecular techniques. Plasmid DNA was isolated using the PureYield Plasmid MiniPrep System (Promega, Madison, WI). Q5 DNA polymerase (New England Biolabs, Ipswich, MA) was used to amplify *ptsI* and *ptsI611* from genomic DNA of the appropriate strain with primers synthesized by Integrated DNA Technologies (Coralville, IA) listed in Table 4.1. The resulting PCR products were purified using a PCR purification kit (Qiagen, Venlo, Limburg, The Netherlands). *Nhe*I and *Xho*I restriction endonucleases (New England Biolabs, Ipswich, MA) were used to digest the PCR product. The digested product was then purified using the PCR purification kit and ligated into the pTEV5 plasmid, generating pDM1671and pDM1672. The identity of each plasmid insert was confirmed by sequencing by Eton Bioscience, Inc, Research Triangle Park, NC.

Expression and Purification of PtsI and PtsI^{R361H} proteins

E. coli BL21-AI strains containing pDM1671 or pDM1672 were inoculated into each of two flasks containing NB Amp (50 mL) and grown overnight at 30 °C with shaking. Each

of two Fernbach flasks containing SB Amp (1.5 L) were inoculated from an overnight culture (1 %), and the flasks were incubated at 37 °C with shaking (180 rpm) until OD_{650} reached 0.15. At that point, the temperature was lowered to 30 °C and when the OD_{650} reached 0.6, expression was induced with IPTG (1 mM) and arabinose (0.02 % w/v). Incubation at 30 °C continued for 20 hours prior to harvesting by centrifugation. The cell pellet was stored at -80 °C until use.

Twenty-five grams of the cell pellet was resuspended to a total volume of 50 mL in Buffer A [50 mM HEPES (Fisher Scientific), 150 mM NaCl, 20 mM Imidazole (Fisher Scientific) pH 7.5 at 4 °C] with DNAse (0.025 mg/mL), lysozyme (1 mg/mL) and phenylmethylsulfonyl fluoride (0.1 mg/mL) and incubated on ice for one hour. The cell suspension was lysed at 1.45 kbar using a Constant Systems Limited One Shot (United Kingdom), and the resultant cell lysate was cleared at 48,000 x g (50 min, 4 °C). Particulates from the cell-free extract were removed using a 0.45 µM PVDF filter (Millipore) and the filtered extract was injected onto two pre-equilibrated 5 mL HisTrap HP Ni-Sepharose columns connected in sequence. The protein was washed with 5 column volumes of Buffer A, 5 column volumes of 4 % Buffer B (Buffer A + 480 mM Imidazole (Fisher Scientific), pH 7.5 at 4 °C) and eluted from the column with a gradient of Buffer B from 4 % to 100 % over 10 column volumes. Fractions containing His_6-PtsI (or His6-$PtsI^{R361H}$) as determined by SDS-PAGE were combined and concentrated by centrifugation using a 50 kDa filter (Millipore), exchanged into 50 mM HEPES buffer with 10 % glycerol, pH 7.5 at 4 °C, using a PD10 column following the manufacturer's instructions (GE Healthcare), flash-frozen in liquid nitrogen and stored at -80 °C until use. Protein concentration was determined by extinction coefficient using the theoretical molecular

weight and A_{280} extinction coefficient of His_6-PtsI as determined by the ExPASy Protparam database (ε_{280} = 30620 M^{-1} cm^{-1}) (Gasteiger *et al.*, 2003). Protein purity was evaluated by SDS-PAGE coupled with densitometry, and both enzymes were purified to > 95% purity (Supplemental Figure 4.1).

PEP hydrolysis by PtsI

The enzymatic activity of His_6-PtsI and His_6-$PtsI^{R361H}$ was measured by coupling the formation of inorganic phosphate from the hydrolysis of PEP with the conversion of P_i and 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) to ribose-1-phosphate and 2-amino-6-mercapto-7-methylpurine by purine nucleoside phosphorylase (PNP) (Webb, 1992; Venditti and Clore, 2012; Venditti *et al.*, 2013). Formation of 2-amino-6-mercapto-7-methylpurine was monitored at 360 nm using the EnzChek Phosphate Assay kit (ThermoFisher Scientific). Briefly, 20 µM His_6-PtsI or His_6-$PtsI^{R361H}$ were incubated with 1 mM $MgSO_4$, 200 µM MESG and 1 U/mL PNP, in 50 mM HEPES, pH 7.5 at 37 °C for 10 minutes. The reaction was initiated with the addition of PEP and the change in absorbance at 360 nm was monitored continuously for an hour. To monitor inhibition, alpha-ketoglutarate was added to a final concentration between 1 and 3 mM during the 10 minute incubation at 37 °C. Rates of P_i formation were calculated from a standard curve of A_{360} from 0 – 10000 pmol KH_2PO_4 in triplicate ($\Delta Abs = 2.54x10^{-5}$pmol P_i, $R^2 = 0.9994$). The effect of αKG on PtsI and $PtsI^{R361H}$ activity was modeled using the non-competitive inhibition model. The following equation was used in Prism 8 and the fit was evaluated.

$$v_{app} = \frac{V_{max}[PEP]}{\left(1+\frac{[\alpha KG]}{K_i}\right)(K_m+[PEP])} \quad \text{(Equation 1)}$$

4.6 REFERENCES

Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., and Wolfe, R.S. (1979) Methanogens: reevaluation of a unique biological group. *Microbiol Rev* **43**: 260–96.

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

Chatterjee, A., Hazra, A.B., Abdelwahed, S., Hilmey, D.G., and Begley, T.P. (2010) A Radical dance in thiamin biosynthesis: mechanistic analysis of the bacterial hydroxymethylpyrimidine phosphate synthase. *Angew Chemie - Int Ed* **49**: 8653–8656.

Chatterjee, A., Li, Y., Zhang, Y., Grove, T.L., Lee, M., Krebs, C., *et al.* (2008) Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. *Nat Chem Biol* **4**: 758–765.

Close, D.M., Cooper, C.J., Wang, X., Chirania, P., Gupta, M., Ossyra, J.R., *et al.* (2019) Horizontal transfer of a pathway for coumarate catabolism unexpectedly inhibits purine nucleotide biosynthesis. *Mol Microbiol* **112**: 1784–1797.

Coquille, S., Roux, C., Fitzpatrick, T.B., and Thore, S. (2012) The last piece in the vitamin B1 biosynthesis puzzle: Structural and functional insight into yeast 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase. *J Biol Chem* **287**: 42333–42343.

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 U S A* **97**: 6640–5.

Doucette, C.D., Schwab, D.J., Wingreen, N.S., and Rabinowitz, J.D. (2011) A-Ketoglutarate Coordinates Carbon and Nitrogen Utilization Via Enzyme I Inhibition. *Nat Chem Biol* **7**: 894–901.

Downs, D.M. (2006) Understanding microbial metabolism. *Annu Rev Microbiol* **60**: 533–559.

Ernst, D.C., Lambrecht, J.A., Schomer, R.A., and Downs, D.M. (2014) Endogenous

synthesis of 2-aminoacrylate contributes to cysteine sensitivity in Salmonella enterica. *J Bacteriol* **196**: 3335–3342.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**: 3784–3788.

Hausinger, R.P. (2004) Fe (II)/ α -Ketoglutarate-Dependent Hydroxylases and Related Enzymes. *Crit Rev Biochem Mol Biol* **39**: 21–68.

Herr, C.Q., and Hausinger, R.P. (2018) Amazing Diversity in Biochemical Roles of Fe(II)/2-Oxoglutarate Oxygenases. *Trends Biochem Sci* **43**: 517–532.

Hollywood, N., and Doelle, H.W. (1976) Effect of specific growth rate and glucose concentration on growth and glucose metabolism of Escherichia coli K-12. *Microbios* **17**: 23–33.

Huergo, L.F., and Dixon, R. (2015) The Emergence of 2-Oxoglutarate as a Master Regulator Metabolite. *Microbiol Mol Biol Rev* **79**: 419–435.

Ishida, S., Tazuya-Murayama, K., Kijima, Y., and Yamada, K. (2008) The direct precursor of the pyrimidine moiety of thiamin is not urocanic acid but histidine in Saccharomyces cerevisiae. *J Nutr Sci Vitaminol (Tokyo)* **54**: 7–10.

Ito, T., Hori, R., Hemmi, H., Downs, D.M., and Yoshimura, T. (2020) Inhibition of glycine cleavage system by pyridoxine 5'-phosphate causes synthetic lethality in glyA yggS and serA yggS in Escherichia coli. *Mol Microbiol* **113**: 270–284.

Ito, T., Iimori, J., Takayama, S., Moriyama, A., Yamauchi, A., Hemmi, H., and Yoshimura, T. (2013) Conserved pyridoxal protein that regulates ile and val metabolism. *J Bacteriol* **195**: 5439–5449.

Ito, T., Yamamoto, K., Hori, R., Yamauchi, A., Downs, D.M., Hemmi, H., and Yoshimura, T. (2019) Conserved Pyridoxal 5'-Phosphate-Binding Protein YggS Impacts Amino Acid Metabolism through Pyridoxine 5'-Phosphate in Escherichia coli. *Appl Environ Microbiol* **85**.

Jurgenson, C.T., Begley, T.P., and Ealick, S.E. (2009) The structural and biochemical foundations of thiamin biosynthesis. *Annu Rev Biochem* **78**: 569–603.

Kelly, L.A., Mezulis, S., Yates, C., Wass, M., and Sternberg, M. (2015) The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat Protoc* **10**: 845–858.

Kittleson, J.T., Wu, G.C., and Anderson, J.C. (2012) Successes and failures in modular genetic engineering. *Curr Opin Chem Biol* **16**: 329–336.

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

Lai, R.Y., Huang, S., Fenwick, M.K., Hazra, A., Zhang, Y., Rajashankar, K., *et al.* (2012) Thiamin pyrimidine biosynthesis in candida albicans: A remarkable reaction between histidine and pyridoxal phosphate. *J Am Chem Soc* **134**: 9157–9159.

Long, C.P., and Antoniewicz, M.R. (2019) Metabolic flux responses to deletion of 20 core enzymes reveal flexibility and limits of E. coli metabolism. *Metab Eng* **55**: 249–257.

Long, C.P., Gonzalez, J.E., Feist, A.M., Palsson, B.O., and Antoniewicz, M.R. (2018) Dissecting the genetic and metabolic mechanisms of adaptation to the knockout of a major metabolic enzyme in Escherichia coli. *Proc Natl Acad Sci U S A* **115**: 222–227.

Martinez-Gomez, N.C., and Downs, D.M. (2008) ThiC is an [Fe-S] cluster protein that requires AdoMet to generate the 4-amino-5-hydroxymethyl-2-methylpyrimidine moiety in thiamin synthesis. *Biochemistry* **47**: 9054–9056.

Palmer, L.D., Paxhia, M.D., and Downs, D.M. (2015) Induction of the sugar-phosphate stress response allows Saccharomyces cerevisiae 2-methyl-4-amino-5-hydroxymethylpyrimidine phosphate synthase to function in Salmonella enterica. *J Bacteriol* **197**.

Paxhia, M.D., and Downs, D.M. (2019) SNZ3 Encodes a PLP Synthase Involved in Thiamine Synthesis in Saccharomyces cerevisiae. **9**: 335–344.

Paxhia, M.D., Swanson, M.S., and Downs, D.M. (2020) Functional characterization of the HMP-P synthase of Legionella pneumophila (Lpg1565). *Mol Microbiol* **5**: 1–15.

Pohl, M., Sprenger, G. a., and Müller, M. (2004) A new perspective on thiamine catalysis. *Curr Opin Biotechnol* **15**: 335–342.

Prunetti, L., Yacoubi, B. El, Schiavon, C.R., Kirkpatrick, E., Huang, L., Bailly, M., *et al.* (2016) Evidence that COG0325 proteins are involved in PLP homeostasis. *Microbiol (United Kingdom)* .

RW, D., D, B., and Roth, J. (1980) Advanced bacterial genetics. *Cold Spring Harb Lab* .

Ryle, M.J., Koehntop, K.D., Liu, A., Que, L., and Hausinger, R.P. (2003) Interconversion of two oxidized forms of taurine/α-ketoglutarate dioxygenase, a non-heme iron hydroxylase: Evidence for bicarbonate binding. *Proc Natl Acad Sci U S A* **100**: 3790–3795.

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

Shiraku, H., Nakashima, M., Takeshita, S., Khoo, C.S., Haniffa, M., Ch'ng, G.S., *et al.* (2018) PLPBP mutations cause variable phenotypes of developmental and epileptic encephalopathy. *Epilepsia Open* **3**: 495–502.

Tazuya, K., Azumi, C., Yamada, K., and Kumaoka, H. (1995) Pyrimidine moiety of thiamin is biosynthesized from pyridoxine and histidine in Saccharomyces cerevisiae. *Biochem Mol Biol Int* **36**: 883–888.

Tazuya, K., Yamada, K., and Kumaoka, H. (1989) Incorporation of histidine into the pyrimidine moiety of thiamin in Saccharomyces cerevisiae. *Biochim Biophys Acta - Gen Subj* **990**: 73–79.

Venditti, V., and Clore, G.M. (2012) Conformational selection and substrate binding regulate the monomer/dimer equilibrium of the c-terminal domain of Escherichia coli

enzyme I. *J Biol Chem* **287**: 26989–26998.

Venditti, V., Ghirlando, R., and Clore, G.M. (2013) Structural basis for enzyme i inhibition by α-ketoglutarate. *ACS Chem Biol* **8**: 1232–1240.

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

Vu, H.N., Ito, T., and Downs, D.M. (2020) The Role of YggS in Vitamin B6 Homeostasis in Salmonella enterica Is Informed by Heterologous Expression of Yeast SNZ3. *J Bacteriol* **202**: e00383-20.

Webb, M.R. (1992) A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems. *Proc Natl Acad Sci U S A* **89**: 4884–4887.

Wightman, R., and Meacock, P. a. (2003) The THI5 gene family of Saccharomyces cerevisiae: Distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine. *Microbiology* **149**: 1447–1460.

Yan, D., Lenz, P., and Hwa, T. (2011) Overcoming fluctuation and leakage problems in the quantification of intracellular 2-oxoglutarate levels in Escherichia coli. *Appl Environ Microbiol* **77**: 6763–6771.

You, C., Okano, H., Hui, S., Zhang, Z., Kim, M., Gunderson, C.W., *et al.* (2013) Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature* **500**: 301–306.

Zhang, K., Bian, J., Deng, Y., Smith, A., Nunez, R.E., Li, M.B., *et al.* (2016) Lyme disease spirochaete Borrelia burgdorferi does not require thiamin. **16213**: 1–6.

Table 4.1 – Strain, plasmids and primers.

Strain	**Relevant Genotype**	**Source**
DM16449	*ΔthiC1225 ΔaraCBAD* / pDM1625	(Paxhia *et al.*, 2020)
DM17114	*ΔthiC1225 ΔaraCBAD ΔyggS638*::*Km* / pDM1625	This study
DM17115	*ΔthiC1225 ΔaraCBAD ΔgdhA631*::*Cm* / pDM1625	This study
DM17147	*ΔthiC1225 ΔaraCBAD nupG*::*Tn10d(Tc) yggS661* / pDM1625	This study
DM17148	*ΔthiC1225 ΔaraCBAD nupG*::*Tn10d(Tc)* / pDM1625	This study
DM17305	*ΔthiC1225 ΔaraCBAD yfeA85*::*Tn10d(Tc)* / pDM1625	This study
DM17306	*ΔthiC1225 ΔaraCBAD yfeA85*::*Tn10d(Tc) ptsI611* / pDM1625	This study
Plasmid		
pDM1625	pTac85-*ScTHI5* (codon-optimized)	(Paxhia *et al.*, 2020)
pDM1671	pTEV5-*ptsI*	This study
pDM1672	pTEV5-*ptsI611*	This study
Primers		
*Nhe*I *ptsI* F	TAGGGCTAGCATGATTTCAGGCATTTTAGCATCC	This study
*Xho*I *ptsI* R	TAGGCTCGAGTTAGCAGATTGTTTTTTCTTCAATGA	This study
yggS F	GTCACAATTTCTTCCATATTCATATAAACATCCTCGGAAAGTGTAGGCTGGAGCTGCTTC	This study
yggS R	ACAGCTCAATTACCGTTGAGAGCAGGAAGGTCAACGTATTCATATGAATATCCTCCTTAG	This study
yggT 5’	TCGTGATTACACAAAAAATtaaGGAAAACTGAGGAACGCCGTGTAGGCTGGAGCTGCTTC	This study
yggT 3’	GCCGTAACACCAGACCGTCTTCGCAGCGAGTCACAGCACTCATATGAATATCCTCCTTAG	This study
yggU 5’	CGACAGGCAACATGTTGCTGCCGGGGCTGTGGATGGCGTTGTGTAGGCTGGAGCTGCTTC	This study
yggU 3’	GCACTTTACCGGCGTTGCCGGTTGCGAGAACAACTTTTTGCATATGAATATCCTCCTTAG	This study

Table 4.2 – Kinetic parameters of PtsI and PtsI^{R361H}

	K_m (PEP; mM)	V_{max} (PEP; pmol Pi/min)	K_i (αKG; mM)
PtsI	0.26 ± 0.02	235 ± 8.1	4.16 ± 0.34
PtsI^{R361H}	0.011 ± 0.001	67 ± 0.7	6.8 ± 0.3

Rate of Pi formation from the hydrolysis of PEP was measured using a coupled assay with purine nucleoside phosphorylase and 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG). αKG was titrated in the assay from 0 – 3 mM to determine the inhibitory constant. The effect of αKG was modeled as a non-competitive inhibitor, and the K_m, V_{max} and K_i were calculated using Equation 1.

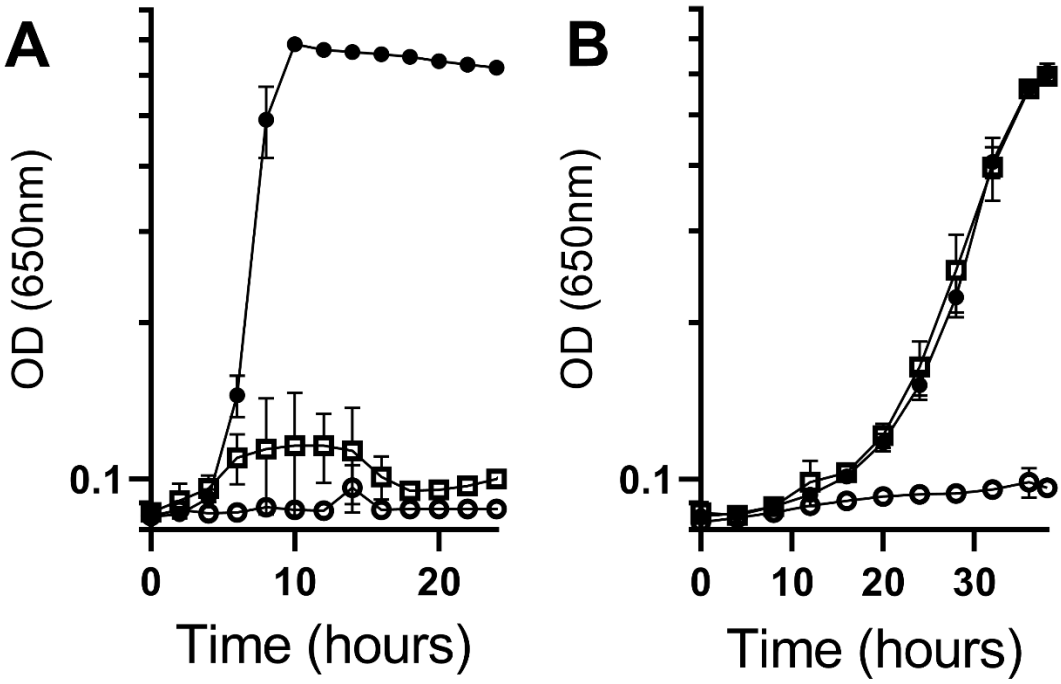


Figure 4.1 – *ptsI611* allows *Sc*Thi5 function on glucose medium
An isogenic pair of *thiC* strains with pTac85-*ScTHI5* (pDM1625) that either lack (A) or contain (B) the *ptsI611* allele were grown on; NCE minimal medium with glucose as a sole carbon source (open circles), plus added thiamine (100 nM) (closed circles), plus IPTG (100 μM) (open squares). Growth was monitored as a function of optical density at 650 nm with shaking at 37 °C. Error bars indicate the standard deviation of three independent biological replicates.

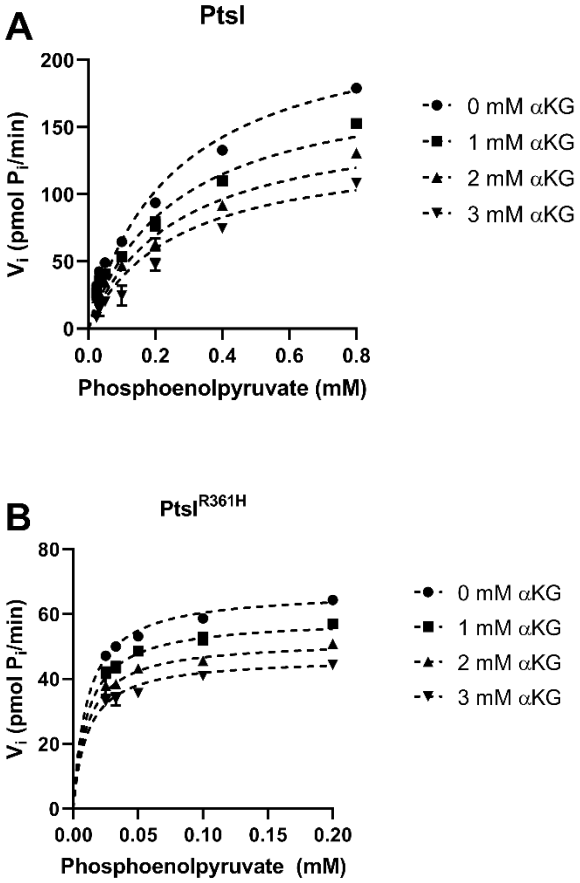


Figure 4.2 – PtsI^{R361H} has a higher affinity for PEP but has a lower V_{max}
The initial velocity of PEP hydrolysis by A) PtsI and B) PtsI^{R361H} was measured using a coupled assay to detect P_i formation with purine nucleoside phosphorylase and 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG). Reactions were performed in 50 mM HEPES pH 7.5, 1 mM $MgSO_4$, 200 µM MESG, 1 U/mL purine nucleoside phosphorylase, with 20 µM PtsI or the variant at 37 °C. αKG was titrated from 0 mM to 3 mM as indicated to determine the K_i for each variant. Error bars indicate the standard deviation of three technical replicates.

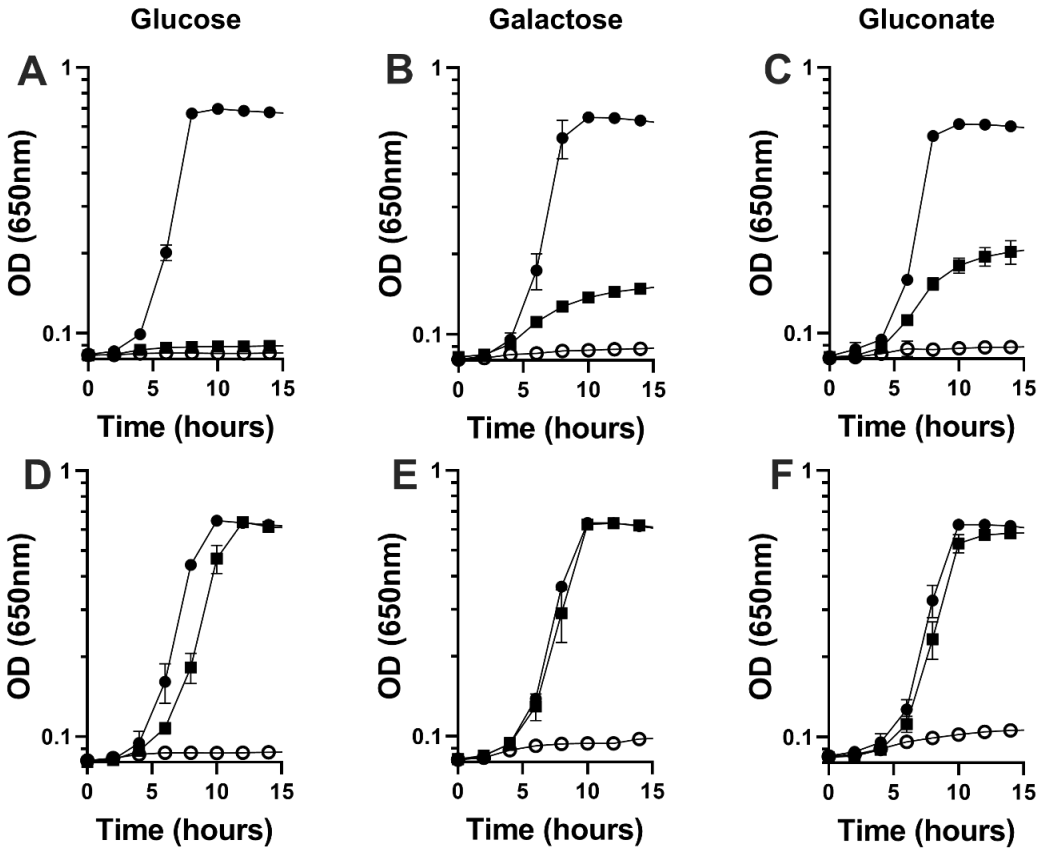


Figure 4.3 – A null allele of *yggS* allows *Sc*Thi5 dependent growth on multiple carbon sources

An isogenic pair of *thiC* strains that contained pTac85-*ScTHI5* (pDM1625) and either lack (DM17148; A-C) or contain (DM17147; D-F) the *yggS661* allele were grown on NCE minimal medium with glucose, gluconate or galactose as a sole carbon source. Filled and empty circles signify the presence and absence of 100 nM Thi, respectively. Filled squares signify the presence of 100 µM IPTG in the medium. The strains and growth conditions shown are A) *thiC* p*ScTHI5* on glucose; B)– *thiC* p*ScTHI5* on galactose; C) – *thiC* p*ScTHI5* on gluconate; D) – *thiC yggS661* p*ScTHI5* on glucose; E) – *thiC yggS661* p*ScTHI5* on galactose; F) – *thiC yggS661* p*ScTHI5* on gluconate. Growth was monitored as a function of optical density at 650 nm with shaking at 37 °C. Error bars indicate the standard deviation of three independent biological replicates.

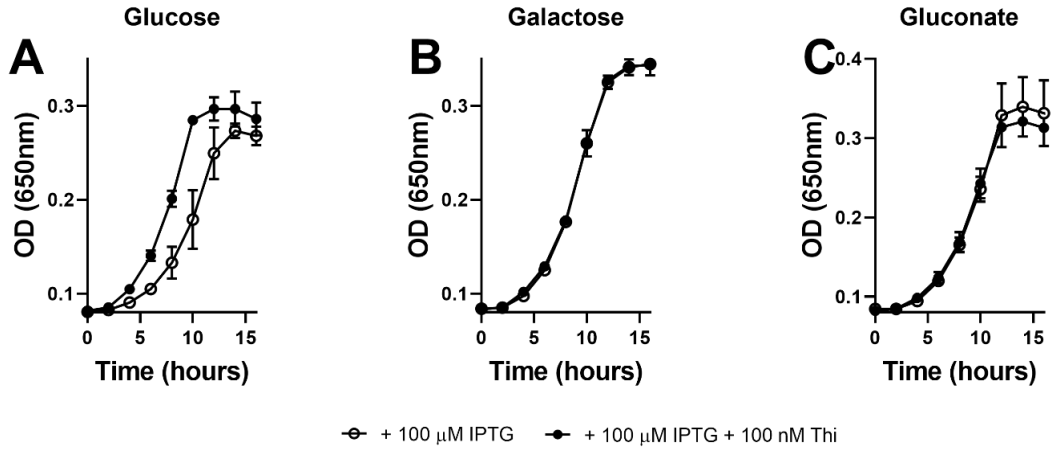


Figure 4.4 – Modifying nitrogen source restores ScThi5 growth
A *thiC* strain with *ScTHI5* on pTac85 was grown on minimal NCN medium with 100 μM IPTG, glutamine (1 mM) as a nitrogen source and glucose (A), gluconate (B) or galactose (C) as a sole carbon source. Filled and empty circles signify the presence and absence of 100 nM Thi, respectively. Growth was monitored as a function of optical density at 650 nm with shaking at 37 °C. Error bars indicate the standard deviation of three independent biological replicates.

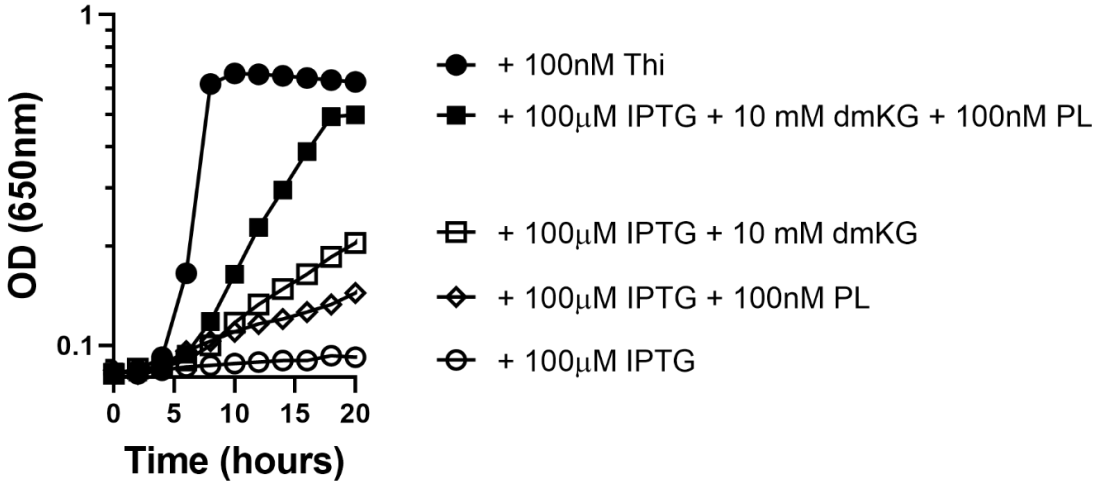


Figure 4.5 – Dimethyl-ketoglutarate and pyridoxal supplementation are additive in supporting *Sc*Thi5 function on glucose
A *thiC* strain with *ScTHI5* on pTac85 was grown on minimal medium with glucose as a sole carbon source. Additions of 100 μM IPTG, 10 mM dimethyl-ketoglutarate (dmKG), 100 nM pyridoxal (PL), or 100 nM thiamine (Thi) are indicated. Growth was monitored as a function of optical density at 650 nm with shaking at 37 °C. A representative of three independent experiments performed in triplicate is shown.

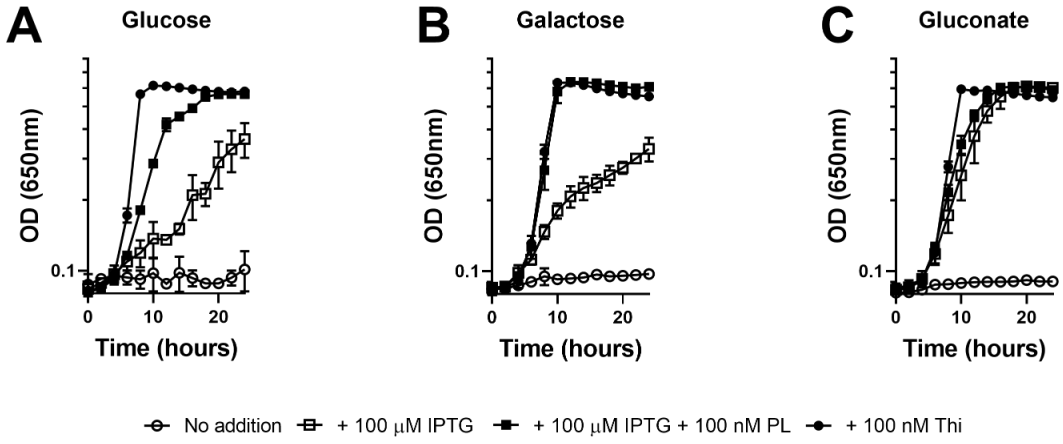


Figure 4.6 – Modifying nitrogen assimilation restores *Sc*Thi5 function
A *thiC gdhA* strain with *ScTHI5* on pTac85 was grown on minimal medium with under high ammonia conditions (NCE) with glucose, galactose or gluconate as a sole carbon source. Empty and filled circles indicate the absence and presence of 100 nM Thi, respectively. Empty squares indicate the presence of 100 μM IPTG, and filled squares indicate 100 μM IPTG and 100 nM PL in the medium. Growth was monitored as a function of optical density at 650 nm with shaking at 37 °C. Error bars indicate the standard deviation of three independent biological replicates.

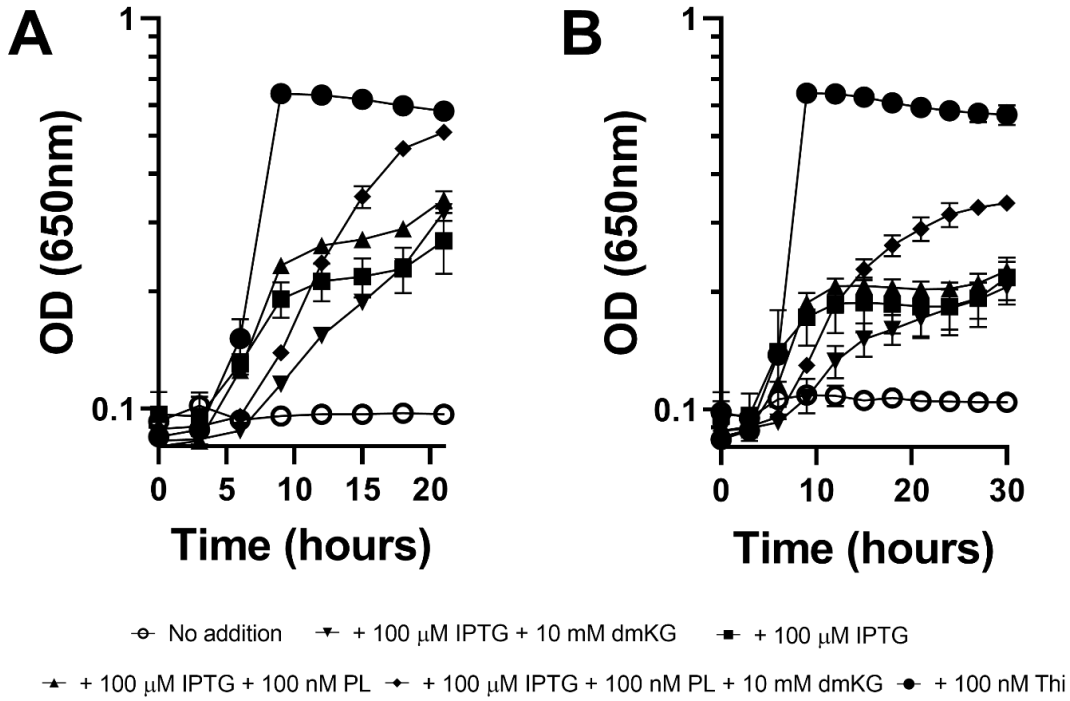
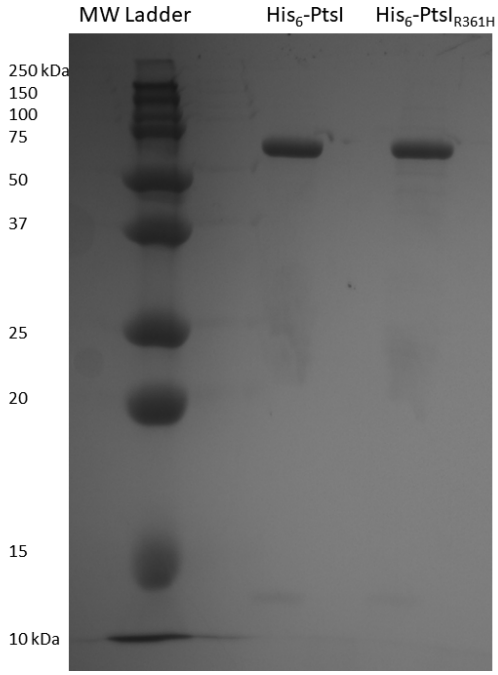
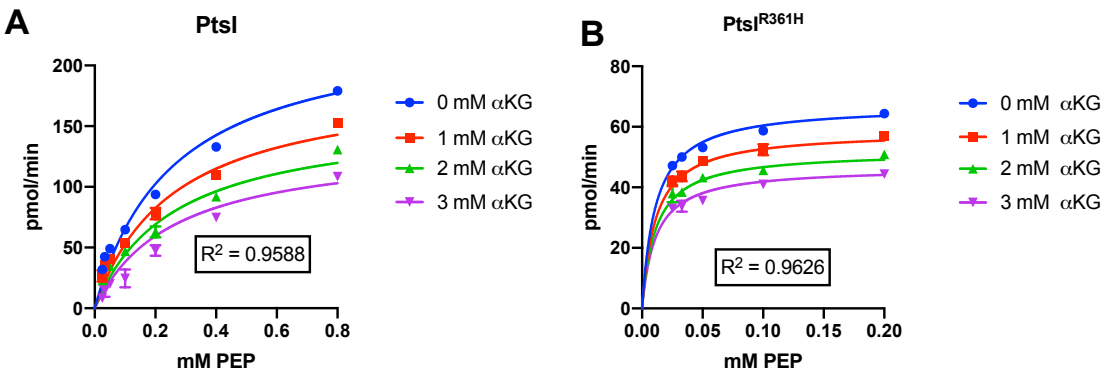


Figure 4.7 – *Lp*Thi5^{F120A} phenocopies *Sc*Thi5 with the addition of dmKG and PL during growth on gluconate

A *thiC* strain with either **A**) *ScTHI5* or **A**) *Lpthi5*^{F120A} on pTac85 was grown on minimal medium with gluconate as a sole carbon source. Additions of 100 μM IPTG, 10 mM dimethyl-ketoglutarate (dmKG), 100 nM pyridoxal (PL), or 100 nM thiamine (Thi) are indicated. Growth was monitored as a function of optical density at 650 nm with shaking at 37 °C. Error bars indicate the standard deviation of three independent biological replicates.



Supplemental Figure 4.1 – Representative purification of His_6-PtsI and His_6-$PtsI_{R361H}$ Two µg of purified His_6-PtsI and His_6-$PtsI_{R361H}$ were denatured at 95 °C for 15 minutes in a loading buffer (60 mM Tris Cl pH 6.8, 2 % SDS, 0.1 M DTT, 10% glycerol) and separated by SDS-PAGE using a 12 % polyacrylamide gel. The gel was stained with Coomassie Blue and visualized using PC Image software and purity was assessed via densitometry using Total Lab TL100 software. Both His_6-PtsI and His_6-$PtsI_{R361H}$ were enriched to > 95 % purity.



Supplemental Figure 4.2 – Modeling the effect of αKG on PtsI and PtsI^{R361H} by non-competitive inhibition

The initial velocity of PEP hydrolysis by PtsI (A) and PtsI$_{R361H}$ (B) was measured using a coupled assay to detect P_i formation with purine nucleoside phosphorylase and 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG). Reactions were performed in 50 mM HEPES pH 7.5, 1 mM $MgSO_4$, 200 µM MESG, 1 U/mL purine nucleoside phosphorylase, 20 µM PtsI or the variant at 37 °C. αKG was titrated from 0 mM to 3 mM as indicated to determine the K_i for each variant. Error bars indicate the standard deviation of three technical replicates. The data were modeled as non-competitive inhibition using Prism 8, and R-squared was quantified to measure goodness of fit.

CHAPTER 5

CONCLUSIONS

S. cerevisiae encodes additional PLP synthases for thiamine synthesis

Chapter 2 adds to our understanding of the structure of the metabolic network around PLP and HMP-P synthesis in *Saccharomyces cerevisiae*. Previously it was thought that the PLP synthase encoded by *SNZ1* was the primary PLP synthase in *S. cerevisiae*. Using a defined media with minimal vitamin additions, the additional paralogs *SNZ2* and *SNZ3* were found to be required for thiamine biosynthesis during fermentative growth. Importantly, *SNZ2* or *SNZ3* alone were sufficient for both PLP and thiamine synthesis during growth on glucose, suggesting that *SNZ1* plays a minor role under these growth conditions.

The Thi5 pathway is found in the Legionellaceae clade of Bacteria

Chapter 3 extends the presence of the Thi5 pathway for HMP biosynthesis into *Legionella pneumophila* and related microorganisms. A *thi5* strain of *L. pneumophila* is a thiamine auxotroph that specifically requires the pyrimidine moiety (HMP) for growth. The high requirement of HMP for growth may be due to lack of a *thiD* homolog in this organism. Instead, thiamine synthesis in this organism is predicted to use a *thiD2-thiE* fusion that likely lacks HMP kinase activity. In contrast to *ScTHI5*, *Lpthi5* is sufficient to complement a *thiC* strain of *S. enterica*. Several conserved residues (K66, H70, C191, C192, C193, and C195) are required for function of this homolog. Additionally, alanine substitutions within

a conserved GEF motif implicated in PLP binding modulates function of this homolog *in vivo*. This extends our knowledge of *in vivo* function of this metabolic pathway.

Changes in both αKG and PLP abundance are required for Thi5 function in *S. enterica*

Chapter 4 identifies two discrete requirements for Thi5 function in a heterologous context. Nutritional and suppressor analysis led to a connection between αKG and Thi5 function on multiple carbon sources. Additionally, exogenous PL is required for full function during growth on glucose. This also extends to a variant of *Lp*Thi5 that is likely compromised in PLP binding, *Lp*Thi5^{F120A}. This suggests that αKG may activate Thi5 enzymes with compromised PLP binding and suggests that *Lp*Thi5 has a higher affinity than *Sc*Thi5 for PLP.

5.2 FUTURE DIRECTIONS

Extending the model of Thi5 function *in vivo*

The incorporation of Thi5 into the *S. enterica* network gives us a new tool to understand connections between central metabolic pathways. Two well-defined growth phenotypes of a *thiC* p*ScTHI5* strain include no growth on glucose without the addition of thiamine and full growth on 50 mM acetate. Previously it was found that a constitutive variant of SgrR restored Thi5 function on glucose (1). Consistent with the working model, increasing glycolytic flux by removing the repressor *mlc* eliminates Thi5 function. Elimination of this repressor would increase expression of genes encoding components of the phosphotransferase system (*ptsI*, *ptsH*, *crr*, *ptsG*, *manXYZ*), leading to greater

glycolytic flux. A suppressor that restores *Sc*Thi5 function in a *sgrR1 mlc* background was isolated in *csrA*. The predicted role of this suppressor is either reducing glycolytic flux by downregulating glycolytic enzymes or upregulating gluconeogenic enzymes to compete with glycolytic flux due to its reciprocal role in regulating both pathways (2). How change in *csrA* regulation leads to lack of growth on acetate is unclear. Characterization of this *csrA* allele may lead to further knowledge about the coupling of glycolysis, gluconeogenesis and the TCA cycle.

Additionally, two genetic suppressors are missing from the current working model of Thi5 function in the heterologous context of *S. enterica*. Incorporating them will likely extend our understanding of central carbon metabolism and how flux is diverted between the three major modules of glycolysis, the pentose-phosphate pathway, and the TCA cycle in *S. enterica*. Based on our working model, reducing glycolytic flux would divert carbon flux through the pentose-phosphate pathway (PPP) and the TCA cycle. Two suppressors that restore *Sc*Thi5 function that seem to conflict with this model are lesions of *zwf* and *gnd* within the oxidative branch of the pentose phosphate pathway. Do these mutations lead to increased flux to the PPP through glycolysis instead? And does the loss of NADPH generation through the oxidative-pentose phosphate pathway lead to increased flux through the TCA cycle to compensate for a lack of reducing power? If not, do these lesions lead to a change in the reduction potential of the cell that impacts Thi5 function?

Finally, Thi5 could potentially be used as a sensitive tool to expand our understanding about the subtle differences in the reduction state of the cell. Previously it was found that lesions in *gshA* and *gshB* eliminate Thi5 function. This is surprising, because glutathione and glutamyl-cysteine have been shown to be redundant for function

of many pathways sensitive to oxidation. Previously there were no identified phenotypes for a *gshB* strain, however strains lacking *gshB* have 50 % of the ribonucleotide reductase activity as a wild-type strain (3). If Thi5 is sensitive to the reduction state of the cell, isolating and characterizing suppressors of a *thiC gshB* /p*ScTHI5* strain may illuminate differences between glutathione and glutamyl-cysteine, which would contribute to our understanding of their apparent redundancy.

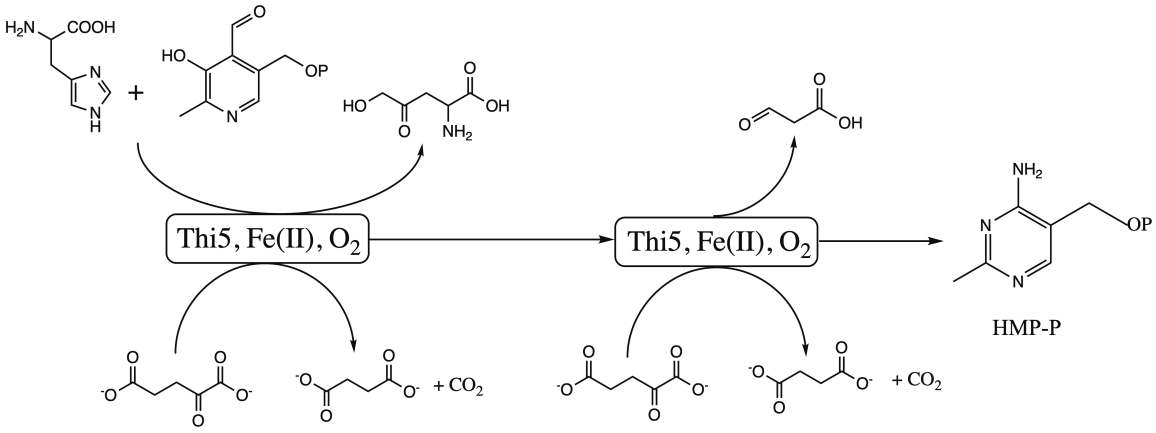


Figure 5.1 – Proposal for a balanced catalytic Thi5 reaction
A balanced reaction proposal for the Thi5 enzyme is shown. HMP-P is formed from histidine and PLP, corresponding to the labeling studies in *S. cerevisiae*, while two molecules of αKG are consumed as co-substrates. Additional putative reaction products include 5-hydroxy-4-oxo-L-norvaline, 3-oxopropanoic acid, two molecules of CO_2 and two molecules of succinate.

Is Thi5 an Fe(II)-ketoglutarate dependent oxygenase?

The connection between αKG and Thi5 function *in vivo* is an exciting addition to our understanding of the requirements within the metabolic network for the enzyme. This may have implications for its function *in vitro*, and lead to a deeper understanding of the structural differences between the pathways found in *S. cerevisiae* and *L. pneumophila.*

Media supplementation and growth conditions that likely lead to increased αKG and PLP levels allow function of both *Sc*Thi5 and *Lp*Thi5^{F120A} in the cell. A simple model is that αKG activates Thi5 homologs with compromised PLP binding. If the mechanism is direct, the oxidation of PLP and histidine in the formation of HMP-P by labeling studies (4–6) might require αKG as a co-substrate (Figure 5.1). Within several metabolic pathways, Fe(II)-ketoglutarate oxygenases have been shown to become inactive due to oxidation of Fe(II) or a residue lining the active site. In the case of Fe(II) oxidation, this inactivation is reversible with the addition of ascorbate (7). However, as the examples of TauD, TfdA and AlkB demonstrate, initiating turnover in the absence of all substrates can irreversibly damage some enzymes through the formation of oxidized residues within the active site (8–11). If the trend holds for Thi5, an anaerobic purification/reconstitution and the addition of αKG and Fe(II) with ascorbate may be required for catalytic turnover with histidine and PLP. Alternatively, other ketoacids like pyruvate or oxaloacetate may be used in the Thi5 reaction, as some Fe(II)-ketoglutarate oxygenases can use alternative ketoacids (12).

The functional differences between *Sc*Thi5 and *Lp*Thi5 *in vivo* may reflect functional differences *in vitro*. Quantifying and comparing the Kd for PLP and/or histidine *in vitro* may distinguish these enzymes from each other and will add to our understanding of the differences in function *in vivo*. Alternatively, if αKG or another keto acid is required for activity of each enzyme, differences in binding may distinguish the two homologs.

In organisms containing both Thi5 and ThiC, how is HMP-P biosynthesis regulated?

A major question that has yet to be addressed is why some organisms use Thi5 over ThiC for HMP-P biosynthesis. Genome sequences for two organisms reveal the presence of both ThiC and Thi5 homologs: *Endozoicomonas elysicola* and *Endozoicomonas atrinae*. The ThiC homologs from these organisms share 67 % amino acid identity to *E. coli* ThiC, 63 % amino acid identity to *Arabidopsis thaliana* ThiC, and only 40 % amino acid identity to BzaF from *Desulfuromonas acetoxidans*. The high identity to true ThiC enzymes, as opposed to than other related radical SAM enzymes suggest that the homolog may encode a genuine ThiC. Initially, a heterologous approach would be useful to test if both homologs can support HMP-P synthesis in *S. enterica*. Beyond this initial experiment, it may be possible to detect expression differences across the thiamine biosynthetic genes within the *Endozoicomonas* species, which might indicate difference in use of each pathway in response to environmental conditions. Some *Endozoicomonas* species are facultative anaerobes (13), so it is plausible that ThiC may be used during anaerobic HMP-P biosynthesis while Thi5 is used under aerobic conditions. Furthermore, biochemical analysis of these homologs may reveal additional features that regulate use of each pathway.

The finding that structural differences between the Thi5 enzymes from *S. cerevisiae* and *L. pneumophila* contribute to function in a heterologous metabolic network is an exciting addition to our understanding about the integration of metabolic modules. It should be rigorously tested with multiple metabolic modules, but it may be that the structure-function relationship found within enzymes reflects the metabolic network in which they have

evolved. If so, theoretically it would be possible to model structure-function relationships within metabolism that better predict function of disparate modules in an organism of interest. One factor that should be tested in the predictive model is whether the local metabolic structure around the module in its native host is predictive of function in a heterologous metabolic network. If successful, this will be another level of optimization in creating synthetic metabolic networks to produce value-added products.

5.3 REFERENCES

1. Palmer LD, Paxhia MD, Downs DM. 2015. Induction of the sugar-phosphate stress response allows Saccharomyces cerevisiae 2-methyl-4-amino-5-hydroxymethylpyrimidine phosphate synthase to function in Salmonella enterica. J Bacteriol 197.

2. Sabnis NA, Yang H, Romeo T. 1995. Pleiotropic regulation of central carbohydrate metabolism in Escherichia coli via the gene csrA. J Biol Chem 270:29096–29104.

3. Fuchs JA, Warner HR. 1975. Isolation of an Escherichia coli Mutant Deficient in Glutathione Synthesis. J Bacteriol 124:140–148.

4. Tazuya K, Yamada K, Kumaoka H. 1989. Incorporation of histidine into the pyrimidine moiety of thiamin in Saccharomyces cerevisiae. Biochim Biophys Acta - Gen Subj 990:73–79.

5. Tazuya K, Azumi C, Yamada K, Kumaoka H. 1995. Pyrimidine moiety of thiamin is biosynthesized from pyridoxine and histidine in Saccharomyces cerevisiae. Biochem Mol Biol Int 36:883–888.

6. Ishida S, Tazuya-Murayama K, Kijima Y, Yamada K. 2008. The direct precursor of the pyrimidine moiety of thiamin is not urocanic acid but histidine in Saccharomyces cerevisiae. J Nutr Sci Vitaminol (Tokyo) 54:7–10.

7. Myllyla R, Majamaa K, Gunzler V. 1984. Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. J Biol Chem 259:5403–5405.

8. Liu A, Ho RYN, Que L, Ryle MJ, Phinney BS, Hausinger RP. 2001. Alternative reactivity of an α-ketoglutarate-dependent iron(II) oxygenase: Enzyme self-hydroxylation. J Am Chem Soc 123:5126–5127.

9. Ryle MJ, Koehntop KD, Liu A, Que L, Hausinger RP. 2003. Interconversion of two oxidized forms of taurine/α-ketoglutarate dioxygenase, a non-heme iron hydroxylase: Evidence for bicarbonate binding. Proc Natl Acad Sci U S A 100:3790–3795.

10. Henshaw TF, Feig M, Hausinger RP. 2004. Aberrant activity of the DNA repair enzyme AlkB. J Inorg Biochem 98:856–861.

11. Grzyska PK, Müller TA, Campbell MG, Hausinger RP. 2007. Metal ligand substitution and evidence for quinone formation in taurine/α-ketoglutarate dioxygenase. J Inorg Biochem 101:797–808.

12. Eichhorn E, Van Der Ploeg JR, Kertesz MA, Leisinger T. 1997. Characterization of α-ketoglutarate-dependent taurine dioxygenase from Escherichia coli. J Biol Chem 272:23031–23036.

13. Chen WM, Lin KR, Sheu SY. 2019. Endozoicomonas coralli sp. nov., isolated from the coral Acropora sp. Arch Microbiol 201:531–538.

APPENDIX A

INDUCTION OF THE SUGAR-PHOSPHATE STRESS RESPONSE ALLOWS SACCHAROMYCES CEREVISIAE 2-METHYL-4-AMINO-5-HYDROXYMETHYLPYRIMIDINE PHOSPHATE SYNTHASE TO FUNCTION IN *SALMONELLA ENTERICA*[4]

Acknowledgements: My contribution to this work consisted of determining which suppressors required presence of *sgrR* for function, growth curves in the presence of alpha-methylglucoside, Western blot for *Sc*Thi5 and beta-galactosidase assays detailed in Figure A.5.

[4] Palmer LD, Paxhia MD, and DM Downs. 2015. *J Bacteriol.* Nov;197(22):3554-62.
Reprinted here with the permission of the publisher

A.1 ABSTRACT

Thiamine pyrophosphate is a required cofactor for all forms of life. The pyrimidine moiety of thiamine, 2-methyl-4-amino-5-hydroxymethylpyrimidine phosphate (HMP-P), is synthesized by different mechanisms in bacteria and plants compared to fungi. In this study, *Salmonella enterica* was used as a host to probe requirements for activity of the yeast HMP-P synthase, Thi5p. Thi5p synthesizes HMP-P from histidine and pyridoxal-5-phosphate and was reported to use a backbone histidine as substrate, requiring that it was a single turnover enzyme. Heterologous expression of Thi5p did not complement a *S. enterica* HMP-P auxotroph during growth with glucose as the sole carbon source. Genetic analyses described here showed that Thi5p was activated in *S. enterica* by alleles of *sgrR* that induced the sugar-phosphate stress response. Deletion of *ptsG* (encodes EII^{BC} of the PTS system) also allowed function of Thi5p and required *sgrR*, but not *sgrS,* for this effect. In total, the data herein supported the hypothesis that one mechanism to activate Thi5p in *S. enterica* grown on minimal glucose medium required decreased PtsG activity and an unidentified gene regulated by SgrR.

Importance

This work describes a metabolic link between the sugar-phosphate stress response and the yeast thiamine biosynthetic enzyme Thi5p when heterologously expressed in *S. enterica* during growth on minimal glucose medium. Suppressor analysis (i) identified a mutant class of the regulator SgrR that activate sugar-phosphate stress response constitutively, and (ii) determined that Thi5p is conditionally active in *Salmonella enterica*.

These results emphasized the power of genetic systems in model organisms to uncover enzyme function and underlying metabolic network structure.

A.2 INTRODUCTION

Thiamine pyrophosphate (TPP) is a cofactor for many central metabolic enzymes, and is required at low levels by all organisms. Humans require dietary intake of thiamine, which is biosynthesized by a variety of plants, bacteria, and fungi. TPP is composed of two independently synthesized moieties, 5-(2-Hydroxyethyl)-4-methylthiazole phosphate (THZ-P) and 2-methyl-4-amino-5-hydroxymethylpyrimidine phosphate (HMP-P). In bacteria and plants, the first steps of the HMP-P biosynthesis pathway are shared with purine biosynthesis (Figure A.1) (1-3). In these organisms, the radical *S*-adenosylmethionine enzyme ThiC catalyzes an intra-molecular rearrangement of the purine intermediate 5-aminoimidazole ribotide (AIR) to the pyrimidine HMP-P (4, 5).

In contrast, fungi do not contain a ThiC homolog; under aerobic conditions HMP-P is synthesized by the Thi5p enzyme family. *In vivo* labeling in yeast implicated histidine and pyridoxine as precursors to HMP-P biosynthesis under aerobic conditions (Figure A.1) (6-8). There are four members of the *THI5* gene family in *Saccharomyces cerevisiae* (*THI5* (*YFL058w*), *THI11* (*YJR156c*), *THI12* (*YNL332w*) and *THI13* (*YDL244w*)), and other species have between zero and five copies of this gene (9). Genetic analysis in *S. cerevisiae* demonstrated that the four enzymes are functionally redundant, as only the quadruple mutant displayed thiamine auxotrophy under aerobic conditions (9). A less efficient HMP-P biosynthetic pathway operates under anaerobic conditions and does not rely on pyridoxine or Thi5p family enzymes (9, 10).

The *in vivo* labeling pattern for aerobic HMP-P synthesis in *S. cerevisiae* suggested unique chemistry was involved. The *Candida albicans* Thi5p activity has been reconstituted (11). These authors reported that the Thi5p protein was oxygen sensitive, but it required oxygen, iron, and PLP for HMP-P synthesis. The Lai study further reported that addition of histidine was not required for activity. Rather, their data suggested that residue His66 of the Thi5p protein served as the histidine substrate, implying Thi5p was a single-turnover enzyme (11). The importance of His66 for Thi5p activity was corroborated by Coquille, et al., who presented the structure of Thi5p and reported His66 was required for Thi5 activity *in vivo* in *S. cerevisiae* (12). The report that Thi5p is a single-turnover enzyme raised questions about potential differences in the physiology between fungi and bacteria/plants that might select for the maintenance of single turnover enzyme when a catalytic mechanism is available (i.e., ThiC). Maintaining such an enzyme mechanism would appear to have significant energy cost implications for a dynamic metabolic network.

Here we describe characterization of *S. cerevisiae* Thi5p activity *in vivo* in the model organism *S. enterica*. This study was initiated with two goals: to begin to describe metabolic network differences between two well-studied organisms of different domains; and to improve future studies of Thi5p enzymatic activity. Metabolic networks are thought to be composed of conserved metabolic modules (18), and general network organizing principles are broadly conserved (19). However, it remains unclear the degree to which metabolic network structure is impacted by component substitution. Thiamine biosynthesis in *S. enterica* is an established system for metabolic dissection (reviewed in (3)), suggesting analysis of Thi5p activity in *S. enterica* could uncover metabolic differences between *S.*

enterica and *S. cerevisiae* in addition to informing mechanistic studies of the Thi5p enzyme.

Although this study was initiated before *in vitro* Thi5p activity was reported, the proposed Thi5p mechanism suggested that further study *in vivo* could uncover enzymatic properties important for Thi5p activity. Historically, physiological studies have lead to successful reconstitution or improvement of *in vitro* activity assays for enzymes that have been difficult to study. For example, the ThiC bacterial HMP-P synthase activity *in vivo* was linked to methionine and iron sulfur cluster metabolism in *Salmonella enterica* (13, 14), leading to its subsequent reconstitution of activity and identification as a Radical SAM enzyme (4, 5). Similarly, although many reports described biotin synthase BioB as a single-turnover enzyme *in vitro*, *in vivo* work in *E. coli* showed that BioB was capable of multiple turnovers and inhibited by its product 5'-deoxyadenosine (15, 16). These findings informed later studies demonstrating that BioB was capable of multiple turnovers *in vitro* when product inhibition was alleviated (17).

The analysis herein has uncovered integration between the sugar-phosphate stress response regulator and Thi5p activity in *S. enterica*. The sugar-phosphate stress response is coordinated by the transcription factor SgrR (formerly YabN) in enteric bacteria (20, 21). SgrR was defined for its activation of the small RNA *sgrS* (20, 22), which destabilizes transcripts of genes encoding PTS transporters (*ptsG* (20) and *manXYZ* (23)) and stabilizes transcripts of a sugar phosphatase (*yigL* (24)). The small RNA *sgrS* also encodes a small peptide, SgrT, which inhibits Enzyme $IIBC^{Glc}$ activity post-translationally (20, 22). Further, SgrR regulates additional genes in *E. coli* including those encoding a sugar efflux pump (*setA* (25)), a glutamic-pyruvic transaminase (*alaC*, formerly *yfdZ* (21)), and itself

(21). The results described herein suggest that one mechanism to allow Thi5p function in *S. enterica* is by remodeling the metabolic network associated with the sugar phosphate stress response regulator.

A.3 MATERIALS AND METHODS

Strains, media and chemicals.

Minimal medium was No-carbon E medium (NCE) (26) supplemented with $MgSO_4$ (1 mM), trace minerals (0.1X) (27) and carbon sources based on 66 mM available carbon units with the following exceptions: pyruvate (50 mM, prepared from powder as needed), acetate (as noted), malate (40 mM), succinate (20 mM) fumarate (50 mM). Thiamine was supplemented at 100 nM. Rich media was Difco nutrient broth (NB; 8 g/L) with NaCl (5 g/L), lysogeny broth (LB) or superbroth (SB; tryptone (32 g/L), yeast extract (20 g/L), NaCl (5 g/L) with NaOH (0.05 N)). Solid media contained 1.5% agar. Antibiotics were added at the following concentrations in rich media, unless otherwise noted: chloramphenicol (Cm), 20 mg/L; ampicillin (Ap), 150 mg/L; tetracycline (Tc), 20 mg/L; kanamycin (Kn), 50 mg/L. All chemicals were purchased from Sigma-Aldrich, St Louis, MO. The strains used in this study were derivatives of *S. enterica* strain LT2 and were generated for this study or were part of the laboratory collection, and their genotypes are listed in Table A.1.

Genetic methods.

The high-frequency generalized transducing mutant of bacteriophage P22 (HT105/1, *int*-201) (28) was used for all transductional crosses. Transduction and subsequent purification

was performed as previously described (29). Mutations in *sgrR, alaC,* and *sgrS* were generated by using lambda Red recombination using primers listed in Table A.2 [Datsenko 2000].

Molecular techniques.

Plasmids were constructed using standard molecular techniques. DNA was amplified using Herculase (Agilent, Santa Clara, CA) or Q5 (New England Biolabs, Ipswich, MA) DNA polymerase. Primers were purchased from Integrated DNA Technologies, Coralville, IA. Plasmids were isolated using the Wizard Plus SV Miniprep kit (Promega, Madison, WI), and PCR products were purified using the PCR purification kit (Qiagen, Venlo, Limburg). Restriction endonucleases were purchased from New England Biolabs, Ipswich, MA, and ligase was purchased from Thermo Scientific, Waltham, MA.

The plasmids and primers are listed in Table A.2. The *THI5* yeast GST-tagged plasmid from Thermo Open Biosystems (Huntsville, AL) was isolated using the modified Qiagen mini-prep protocol for yeast and used as template for amplification of *THI5*. Later analysis revealed the clone ID YDR155C plasmid included a mixed population of wild-type and variant *THI5* (cloned in pDM1336) with the following substitutions: M37I, A138V, G152D. We did not detect any functional differences between the variant *THI5* or wild-type *THI5*. pSU18 is a pACYC184 derivative compatible with pBR322 that uses the *lac* promoter and encodes Cm^R (30). To construct pDM1336 and pDM1381, *THI5* was amplified using primers *S. cerevisiae* Thi5 for 5' and *S. cerevisiae* Thi5 rev 3'; the resulting PCR products were purified and blunt-end ligated into pSU18 digested with SmaI.

pBAD vectors are pBR322 derivatives that use the *ara* promoter and encode Amp^R; pBAD18S includes the N-terminal sequence of AraB Met-Ala-Ile-Ala-Gly prior to the multiple cloning site (31). *THI5* was cloned into SacI/HindIII sites in pBAD18S to construct pDM1361. pFZY1 is a mini-F derivative (averages 1-2 copies per cell) with a multiple-cloning site upstream of a promoterless galK9-lacZYA reporter segment (32). The promoter region of *sgrS* was cloned into pFZY1 using the KpnI/BamHI sites to construct pDM1402 *sgrSp-lacZ*).

Western blot.

Cells containing pSU18-*THI5* or pSU18 were grown overnight in 10 mL of minimal glucose or ribose medium. Cultures were centrifuged (10 min at 17000 × *g*) and resuspended in 14.5 mM NaCl before sonication (1 second pulses for 1 minute) (Fisher Scientific). Lysate was clarified by centrifugation (10 min at 17000 × *g*) and total protein was quantified by the bicinchoninic acid **(**BCA) assay (Pierce) using bovine serum albumin (BSA) as a standard. Protein samples were denatured in a loading buffer (60 mM TrisCl pH 6.8, 2 % SDS, 0.1 M DTT, 10 % glycerol) at 85°C for 15 min before sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using a 12% gel. Proteins were then transferred to a polyvinylidene fluoride **(**PVDF) membrane (Bio-Rad Laboratories) and detected using a rabbit polyclonal antibody for Thi5p from Harlan Bioproducts (1/1000 dilution) as a primary antibody and a horseradish peroxidase conjugated polyclonal goat anti-rabbit IgG (1/50,000 dilution) as the secondary antibody (Thermo Scientific Pierce). The blot was then visualized using an ECL Plus Western Blotting Substrate (Pierce) on a Typhoon Trio + (GE Healthcare).

For antibody production, *THI5* was cloned into the pET14b vector (Novagen) using the NdeI/XhoI sites and transformed into *E. coli* BL21AI. Cells were grown, expression induced, and protein purified by laboratory protocols used routinely for purifying His tagged proteins using nickel-nitrilotriacetic acid (Ni-NTA) Superflow resin (Qiagen). Protein aliquots were stored at −80°C prior to use. Purified His_6 Thi5p was used as a control in the Western blot procedure.

Suppressor isolation.

Ten independent cultures of DM13623 in NB Cm were incubated overnight with shaking at 37°C. A 100-μL sample (~ 1 X 10^8 cells) was spread on a minimal glucose medium plate and 5 μL diethyl sulfate (DES) was spotted in the center of the plate as indicated. Suppressor frequency was analyzed after incubation at 37°C for 2 days. Six colonies per plate were streaked for individual colonies on selective medium (minimal glucose), then non-selective medium (NB Cm), before patching to NB Cm and printing to minimal glucose for confirmation of selected phenotype.

The resultant strains were separated into classes based on growth in minimal glucose medium. For each class, a transposon (Tn*10d*(Tc)) genetically linked to the causative mutation was isolated by standard genetic techniques and used to reconstruct the mutant for phenotypic confirmation. The chromosomal location of relevant insertions was determined by sequencing using a PCR-based protocol (33). A DNA product was amplified with degenerate primers and primers derived from the Tn*10d*(Tc) insertion sequence and sequenced. The genome of the reconstructed strains was sequenced to identify the causative mutation (see below). Alternatively, putative loci were PCR amplified and sequenced by

GeneWiz (South Plainfield, NJ). The strains were reconstructed in the DM14419 background.

Genome sequencing.

Whole-genome sequencing was used to identify the causative suppressor mutations in DM13631 and DM13718 (*ptsI611*). The protocol used for preparation of genomic DNA has been described (34).

Genomic DNA was submitted to the Georgia Genomics Facility (GGF) at the University of Georgia (Athens, GA) for paired-end (2 x 250 bp) sequencing using the Illumina MiSeq platform. DNA samples were fragmented and tagged with sequencing adapters using the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA). Processing and assembly of the sequencing data was performed by the Georgia Advanced Computing Resource Center (GACRC) at the University of Georgia. Briefly, the raw sequencing data was cleaned up using Trimmomatic (Usadel Lab, Max Planck Institute, Germany) with a read length cut-off of 100 bp, resulting in >300-fold coverage of the 4.95 Mb *S. enterica* LT2 genome (35). Trimmed reads were mapped to the published genome using Bowtie 2 (Source Forge). Variant calling was performed using the Genome Analysis Toolkit (Broad Institute, Cambridge, MA), and single nucleotide polymorphisms (SNPs) were identified using the Integrative Genomics Viewer (Broad Institute).

β-galactosidase assays. β-galactosidase assays were performed with modifications of previously described methods (36, 37). To determine β-galactosidase activity in NB medium, a 50-μL sample of overnight culture grown in NB medium with ampicillin (30

mg/L) was inoculated into 5 mL medium with ampicillin (30 mg/L). The cultures were incubated at 37°C with shaking until OD_{600} = 0.5-0.7, at which time samples were removed and incubated with α-MG (0.5%) or an equal volume ddH_2O in a 96-well plate. After cells were incubated +/- α-MG for 45 min at 37°C with shaking, the OD_{600} of 175-µL cell samples was determined by a SpectraMax 385 Plus microplate spectrophotometer (Molecular Devices, Sunnyvale, CA). Twenty-µL samples of cells were then added to 80 µL permeabilization solution that contained Na_2HPO_4 (100 mM), KCl (20 mM), $MgSO_4$ (2 mM), hexadecyltrimethylammonium bromide (1.6 mM), deoxycholic acid sodium salt (0.9 mM), and β-mercaptoethanol (77 mM). After at least 10 min incubation, 25 µL permabilized cell mixture was added to 150 µL substrate solution that contained Na_2HPO_4 (60 mM), NaH_2PO_4 (40 mM), *o*-nitrophenyl-β-D-galactoside (ONPG, 3.3 mM), and β-mercaptoethanol (39 mM). The reactions were incubated at 30°C and ONP product formation was monitored at A_{420} over time. Rates (ΔA_{420}/min) were determined by fitting the data to a linear equation with outlier elimination in GraphPad Prism 6.0d (La Jolla, CA). Specific activity A was calculated using the formula 1000 X rate (ΔA_{420}/min)/OD_{600} using OD_{600} as measured from 175-µL cell samples in a flat-bottom 96-well plate using a Spectramax 385.

Alternatively, to determine β-galactosidase activity in minimal glucose medium, a 50-µL sample of overnight culture grown in NB medium with ampicillin (30 mg/L) was inoculated into 5 mL minimal glucose medium with ampicillin (15 mg/L) and with α-MG (1%) as indicated. The cultures were incubated at 37°C with shaking until mid-logarithmic phase growth, and OD_{600} was 0.5-0.7 in 18 X 150 mm borosilicate tubes as measured by a Spectronic 20+ (Thermo Scientific, Waltham, MA). Then, 20-µL samples were added to

80 μL permeabilization solution and β-galactosidase activity was measured as above. Specific activity B was calculated using the formula 1000 X rate (ΔA_{420}/min)/OD_{600} using OD_{600} as measured from 5-mL cell samples in 18 X 150 mm borosilicate tubes using a Spectronic 20+.

A.4 RESULTS AND DISCUSSION

Thi5p functions conditionally in *S. enterica*. The *THI5* gene was cloned into the pSU18 vector (pDM1336 and pDM1381) with expression from the *lac* promoter. *S. enterica* lacks the lactose utilization operon, and therefore expression in the pSU18 vector was constitutive. *THI5* expression did not allow growth of a *S. enterica thiC* mutant strain on minimal glucose medium (Table A.3, Figure A.2A). Addition of 2 or 5 mM cAMP did not alter growth, indicating the glucose-specific effect was not due to catabolite repression (data not shown). Further analysis assessed Thi5p-dependent growth of strain DM14419 (Δ*thiC*/p*THI5*) on a suite of carbon sources (Table A.3). Thiamine allowed full growth of this strain on all carbon sources, as expected. Significant Thi5p-dependent thiamine synthesis was observed on a limited number of permissive carbon sources, including ribose, xylose and mannose. Even on these permissive carbon sources, the growth rates in the absence of thiamine were low; however, a high final density was reached. In contrast, Thi5p did not support significant growth in the absence of thiamine on glucose and other carbon sources (e.g., pyruvate). Expression of *THI5* from pDM1361 (pBAD18S-*THI5*) induced with arabinose (0.02% or 0.2%) did not change the growth pattern (data not shown). These results indicated that the differential Thi5p-dependent growth was not the result of poor gene expression on different carbon sources. This conclusion was further

demonstrated by western blot analyses that showed the accumulation of Thi5p on a non-permissive (glucose) and permissive (ribose) carbon source was not significantly different (Figure A.3). Taken together, these data showed that Thi5p function was impacted by components of the metabolic network that differed based on carbon source.

A genetic approach identified metabolic factors that allowed Thi5p activity in *S. enterica*. Suppressor mutations that allowed growth of a Δ*thiC1225* Δ*araCBAD* p*THI5* strain on minimal glucose medium were isolated. Colonies arose at a rate of ~1 X 10^{-7} when DM13623 was plated on glucose medium, but none retained this phenotype after purification. Addition of diethyl sulfate (DES) to the selection plate increased the number of colonies and resulted in stable revertants. Four independent, stable derivatives of DM13623 that grew on minimal glucose plates were further analyzed. Genetic mapping, along with targeted and whole-genome sequencing identified one causative mutation in the *ptsI* locus and two in the *sgrR* locus. A fourth mutation (*zxx10175*) was not identified, and was genetically unlinked to either the *ptsI* or the *sgrR* locus. The *sgrR* suppressor strains were chosen for further investigation and their growth pattern characterized in minimal glucose medium with and without thiamine. Representative data are shown in Figure A.2. The data showed that, as expected, the growth rate of the parent strain DM13623 was increased by the addition of thiamine (0.04 ± 0.03 h^{-1}, 0.60 ± 0.03 h^{-1}, without and with thiamine, respectively). In contrast, the strain carrying mutant allele *sgrR1* (Figure A.2B) grew approximately as well with or without thiamine supplementation. It was noted that the growth rate of the *sgrR1* containing strain with thiamine was less than the parental strain (0.33 ± 0.01 h^{-1}). This result suggested that the suppressing mutations carried a fitness cost to the strains on glucose medium, a result not atypical when the metabolic

network is remodeled to compensate for a perturbation or to generate a new function (38-42).

Alleles of *sgrR* allow Thi5p function in *S. enterica*.

Two alleles of *sgrR* (*sgrR1, sgrR2*) were isolated as suppressor mutations that allowed growth of a Δ*thiC1225* Δ*araCBAD* p*THI5* strain on glucose medium. The same strain lacking p*THI5* failed to grow, indicating that Thi5p was satisfying the thiamine requirement of the cell. SgrR mediates the sugar-phosphate stress response that has been characterized in *E. coli*. SgrR belongs to a novel class of transcription regulators (COG4533) and has a predicted N-terminal DNA-binding domain and C-terminal solute-binding domain (20, 21). Analysis of *S. enterica* SgrR in the Interpro database (43) identified the N-terminal binding domain as amino acids 5-118, with the solute binding domain beginning at amino acid 163. The *sgrR1* allele encodes a variant with a substitution in the putative solute-binding domain ($SgrR^{G525R}$), while the *sgrR2* allele encodes a variant with a substitution immediately following the predicted DNA-binding domain ($SgrR^{R119W}$). A null allele of *sgrR* did not allow Thi5p-dependent growth on minimal glucose medium, indicating the suppressing alleles encoded variants with altered function (data not shown).

Expression from the *sgrS* promoter (*sgrSp*) has been used as a reporter of SgrR activity (20) and was used to probe the function of the suppressing variants of SgrR. A vector (pDM1402) in which the *sgrS* promoter drives expression of the *lacZ* gene was constructed. This plasmid allowed β-galactosidase activity to be assayed as a proxy for expression from the *sgrS* promoter. The non-metabolizable glucose analog α-

methylglucoside (α-MG) is a gratuitous inducer of sugar-phosphate stress and SgrR activity (20). The effect of SgrR variants on expression of *lacZ* from pDM1402 with and without induction by α-MG was measured in nutrient medium. The data in Figure A.4 showed that both $SgrR^{G525R}$ and $SgrR^{R119W}$ increased transcription in the absence of inducer. These data suggested the SgrR variants were constitutively active and resulted in the continual induction of the sugar-phosphate stress response. In the presence of the α-MG inducer, expression of *lacZ* did not increase to the same level allowed by the wild-type protein. The above results were consistent with the hypothesis that the variants of SgrR that allowed function of Thi5p had an altered ability to sense the appropriate signal.

In the above assays the cells were grown in nutrient medium, potentially masking a response that was relevant for the ability of the *sgrR* alleles to allow Thi5p to function *in vivo.* Strains with representative *sgrR* alleles and the *sgrS*p-*lacZ* reporter construct were grown in glucose medium with thiamine. The β-galactosidase activity was measured during mid-logarithmic growth to assess transcription mediated by $SgrR^{G525R}$ encoded by *sgrR1*. The data in Figure A.5 showed that *sgrR1* increased expression from the *sgrS* promoter on minimal thiamine medium when glucose was the carbon source. Therefore, a variant SgrR that allowed Thi5p-dependent thiamine synthesis activated the sugar-phosphate stress response constitutively during growth on glucose.

Disruption of glycolysis allowed Thi5p-dependent growth on glucose.

Glycolytic flux has been tied to the induction of the sugar-phosphate stress response (44). Mutations that eliminated phosphoglucose isomerase activity (encoded by *pgi*) or phosphofructokinase A (encoded by *pfkA*, and responsible for ~95% of

phosphofructokinase activity in *E. coli* (45)) were introduced into a Thi5p-containing strain and thiamine synthesis was assessed by growth on minimal medium. Eliminating either *pgi* or *pfkA* allowed Thi5p-dependent growth on glucose (Table A.4). Previous reports found that *E. coli* strains disrupted in *pgi* and *pfkA* had destabilized *ptsG* mRNA (46, 47), presumably through the sugar-phosphate stress response (20, 21). Therefore, the strains carrying *pgi* or *pfkA* mutations were assessed for their sugar-phosphate stress response induction by monitoring the expression of *sgrSp-lacZ* during growth on glucose medium. The data in Figure A.5 showed that elimination of either locus induced the expression of *sgrS*, again consistent with the conclusion that induction of the sugar-phosphate stress response was associated with the activation of Thi5p in *S. enterica.*

Induction of the sugar phosphate stress response is required to decrease PtsG activity. The data above supported a general model in which the constitutive induction of the sugar-phosphate stress response allowed Thi5p function in *S. enterica* on glucose medium. α-MG was used as a gratuitous inducer of the sugar-phosphate stress response to further test this conclusion. Titration of α-MG determined that when 1% α-MG was added to the medium, growth of strain DM14419 (Δ*thiC* pTHI5) occurred on glucose in the absence of thiamine (growth rate 0.139 ±0.01 with α-MG, <0.01 ± 0.01 without α-MG). Expression from the *sgrSp-lacZ* reporter plasmid in strain DM14531 (Δ*thiC* p*sgrS*p) during mid-log growth in minimal glucose medium containing 1% α-MG was at a low level of expression, similar to that induced by *sgrR1* on glucose medium (Figure A.5).

The primary target of the sugar phosphate stress response, PtsG, was tested for an effect on Thi5p function in minimal glucose medium. An insertion mutation in *ptsG*

allowed Thi5p-dependent thiamine synthesis in a Δ*thiC* strain, as shown in Figure A.6A. The effect of the *ptsG* mutation was independent of *sgrS,* consistent with a role for *sgrS* in decreasing PtsG activity (Fig. 6C). In contrast, *sgrR* was required for Thi5p-dependent thiamine synthesis allowed by the *ptsG* mutation (Fig. 6B). Together these results indicated that the suppression observed in the *sgrR1* and *sgrR2* mutant strains had two components: the constitutive expression of *sgrS* resulted in decreased PtsG activity and an additional, undefined role of SgrR. Further, the role of SgrR was not to express *alaC*, which is the only known SgrR target in *S. enterica* that is not regulated via *sgrS* (Fig. 6D). Finally, these data showed that the secondary role of SgrR did not require its "activation" since the allele present was wild-type and no known inducing conditions were needed. Consistent with this interpretation, the presence of a *ptsG* mutation did not affect the basal level of expression from the *sgrSp-lacZ* reporter, or its ability to be induced by α-MG (data not shown). It was noted that the mutations in *sgrR* or *alaC* did affect the growth pattern of the strains, in each case increasing the lag before growth. The specific cause for this is not obvious but could indicate a subtle role of the relevant genes in the metabolic network required to support Thi5p-dependent thiamine synthesis.

Induction of the sugar phosphate stress response is not necessary for Thi5p activity.

Together, the results above suggested that induction of the sugar-phosphate stress response was involved in at least one mechanism that allowed Thi5p function in *S. enterica* during growth on glucose. The data showed that one relevant component of this response was the induction of *sgrS* that led to decreased PtsG. Other observations indicated that the stress response, and in fact SgrR, was not required for all mechanisms that allowed Thi5p

function in *S. enterica.* For instance, when the parental strain (DM14419) was grown on a permissive carbon-source (ribose), expression of the *sgrSp-lacZ* reporter was at a low basal level and deletion of *sgrR* did not prevent Thi5p-dependent growth (data not shown). This result indicated that induction of the sugar-phosphate response was not the only mechanism that allowed function of Thi5p in *S. enterica* and implicates a role for the different carbon source-specific metabolic networks in Thi5p activity in *S. enterica*. In addition, the *ptsI* allele that restored Thi5p-dependent growth on minimal glucose medium did not require *sgrR* or *sgrS* for this effect (data not shown).

A.5 CONCLUSIONS

S. cerevisiae HMP-P synthase Thi5p is capable of supporting thiamine-independent growth in a *ΔthiC* mutant strain of *S. enterica* under certain conditions. The conditionality of Thi5p function in *S. enterica* provided an opportunity to dissect the metabolic components involved. Two suppressor mutations that allowed Thi5p function on glucose medium were alleles of *sgrR,* which encodes a transcriptional regulator. The suppressing alleles resulted in constitutive expression of *sgrS*, suggesting a link between Thi5p and the sugar-phosphate stress response. This link was extended when null mutations in *pgi* (encodes phosphoglucose isomerase) or *pfkA* (encodes phosphofructokinase A) also, (i) induced *sgrS* expression and (ii) allowed Thi5p-dependent thiamine synthesis on glucose medium. Further, inducing the sugar phosphate stress response with the gratuitous inducer α-MG allowed Thi5p-dependent thiamine synthesis in glucose medium.

The finding that disrupting *ptsG* independently restored thiamine synthesis revealed one mechanism connecting Thi5p function and induction of the sugar-phosphate

stress response. The results associated with strains lacking *ptsG* supported the working model in which one of the mechanisms of Thi5p activation requires decreased PtsG function and an undefined role of SgrR. The expectation is that the role of SgrR is to regulate a gene whose product is required to remodel the metabolic network to allow Thi5p function. Interestingly, the data suggest the relevant role of SgrR does not depend on conditions that are known to activate the regulator. Thus this work has the potential to expand our understanding of SgrR and how it contributes to metabolism beyond its role in the sugar phosphate stress response.

When considered in total, the data suggest Thi5p must be activated in order to function in *S. enterica*, possibly by phosphorylation or another post-translational modification,. This scenario would demand a Thi5p-activating enzyme in *S. cerevisiae* with no *S. enterica* ortholog. Instead, robustness in the *S. enterica* metabolic network would support Thi5p activation under certain conditions, including those generated by growing on a permissive carbon source or by remodeling the metabolic network by inducing the sugar-phosphate stress response during growth on glucose. Alternative hypotheses include that Thi5p activity is inhibited during growth on glucose, or that a relevant cofactor is not available. In either case, the limitation can be alleviated through activation of the sugar-phosphate stress response, specifically the decrease in PtsG activity. Regardless of which scenario is correct, further definition of this system will have implications for *in vitro* studies on Thi5p. Production of Thi5p in minimal glucose medium (as reported by Lai *et al.*, 2012) could deleteriously affect its activity in reconstituted assays. In total, the findings here emphasize that metabolism is characterized by complex metabolic networks that are not always amenable to “plug-and-play” approaches, even if all required substrates are

present. The data further presented a number of provocative questions about the enzymology of Thi5p and the sugar-phosphate stress response regulator. The results reported here identify a new area of investigation that has the potential to yield unexpected knowledge about central metabolic processes and the integration of heterologous components to a network.

Acknowledgments

We acknowledge Man Him (Sammy) Leung for construction of pDM1336 and isolation and mapping of the *sgrR1* suppressor mutation. The National Institutes of Health grant GM47296 to DMD supported this work. LDP was supported by NSF through Graduate Research Fellowship grant DGE-0718123.

A.6 REFERENCES

1. **Jurgenson CT, Begley TP, Ealick SE.** 2009. The structural and biochemical foundations of thiamin biosynthesis. Annual review of biochemistry **78:**569-603.

2. **Begley TP, Chatterjee A, Hanes JW, Hazra A, Ealick SE.** 2008. Cofactor biosynthesis--still yielding fascinating new biological chemistry. Curr Opin Chem Biol **12:**118-125.

3. **Koenigsknecht MJ, Downs DM.** 2010. Thiamine biosynthesis can be used to dissect metabolic integration. Trends Microbiol **18:**240-247.

4. **Martinez-Gomez NC, Downs DM.** 2008. ThiC is an [Fe-S] cluster protein that requires AdoMet to generate the 4-amino-5-hydroxymethyl-2-methylpyrimidine moiety in thiamin synthesis. Biochemistry **47:**9054-9056.

5. **Chatterjee A, Li Y, Zhang Y, Grove TL, Lee M, Krebs C, Booker SJ, Begley TP, Ealick SE.** 2008. Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. Nat Chem Biol **4:**758-765.

6. **Tazuya K, Yamada K, Kumaoka H.** 1993. Pyridoxine is a precursor of the pyrimidine moiety of thiamin in Saccharomyces cerevisiae. Biochem Mol Biol Int **30:**893-899.

7. **Tazuya K, Yamada K, Kumaoka H.** 1989. Incorporation of histidine into the pyrimidine moiety of thiamin in *Saccharomyces cerevisiae*. Biochim Biophys Acta **990:**73-79.

8. **Ishida S, Tazuya-Murayama K, Kijima Y, Yamada K.** 2008. The direct precursor of the pyrimidine moiety of thiamin is not urocanic acid but histidine in *Saccharomyces cerevisiae*. Journal of nutritional science and vitaminology **54:**7-10.

9. **Wightman R, Meacock PA.** 2003. The THI5 gene family of *Saccharomyces cerevisiae*: distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine. Microbiology **149:**1447-1460.

10. **Tanaka K, Tazuya K, Yamada K, Kumaoka H.** 2000. Biosynthesis of thiamin under anaerobic conditions in *Saccharomyces cerevisiae.* Biological & pharmaceutical bulletin **23:**108-111.

11. **Lai RY, Huang S, Fenwick MK, Hazra A, Zhang Y, Rajashankar K, Philmus B, Kinsland C, Sanders JM, Ealick SE, Begley TP.** 2012. Thiamin Pyrimidine Biosynthesis in *Candida albicans* : A Remarkable Reaction between Histidine and Pyridoxal Phosphate. J Am Chem Soc.

12. **Coquille S, Roux C, Fitzpatrick TB, Thore S.** 2012. The last piece in the vitamin B1 biosynthesis puzzle: structural and functional insight into yeast 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase. J Biol Chem **287:**42333-42343.

13. **Newell PC, Tucker RG.** 1968. Biosynthesis of the pyrimidine moiety of thiamine. A new route of pyrimidine biosynthesis involving purine intermediates. Biochemistry Journal **106:**279-287.

14. **Dougherty MJ, Downs DM.** 2006. A connection between iron-sulfur cluster metabolism and the biosynthesis of 4-amino-5-hydroxymethyl-2-methylpyrimidine pyrophosphate in *Salmonella enterica*. Microbiol. **152:**2345-2353.

15. **Choi-Rhee E, Cronan JE.** 2005. Biotin synthase is catalytic in vivo, but catalysis engenders destruction of the protein. Chem Biol **12:**461-468.

16. **Choi-Rhee E, Cronan JE.** 2005. A nucleosidase required for in vivo function of the S-adenosyl-L-methionine radical enzyme, biotin synthase. Chem Biol **12:**589-593.

17. **Farrar CE, Siu KK, Howell PL, Jarrett JT.** 2010. Biotin synthase exhibits burst kinetics and multiple turnovers in the absence of inhibition by products and product-related biomolecules. Biochemistry **49:**9985-9996.

18. **von Mering C, Zdobnov EM, Tsoka S, Ciccarelli FD, Pereira-Leal JB, Ouzounis CA, Bork P.** 2003. Genome evolution reveals biochemical networks and functional modules. Proc Natl Acad Sci U S A **100:**15428-15433.

19. **Albert R, Jeong H, Barabasi AL.** 2000. Error and attack tolerance of complex networks. Nature **406:**378-382.

20. **Vanderpool CK, Gottesman S.** 2004. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. Mol Microbiol **54:**1076-1089.

21. **Vanderpool CK, Gottesman S.** 2007. The novel transcription factor SgrR coordinates the response to glucose-phosphate stress. J Bacteriol **189:**2238-2248.

22. **Wadler CS, Vanderpool CK.** 2007. A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. Proc Natl Acad Sci U S A **104:**20454-20459.

23. **Rice JB, Vanderpool CK.** 2011. The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. Nucleic Acids Res **39:**3806-3819.

24. **Papenfort K, Sun Y, Miyakoshi M, Vanderpool CK, Vogel J.** 2013. Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. Cell **153:**426-437.

25. **Sun Y, Vanderpool CK.** 2011. Regulation and function of *Escherichia coli* sugar efflux transporter A (SetA) during glucose-phosphate stress. J Bacteriol **193:**143-153.

26. **Vogel HJ, Bonner DM.** 1956. Acetylornithase of *Escherichia coli*: partial purification and some properties. Journal of Biological Chemistry **218:**97-106.

27. **Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS.** 1979. Methanogens: reevaluation of a unique biological group. Microbiol. Rev. **43:**260-296.

28. **Schmieger H.** 1972. Phage P22-mutants with increased or decreased transduction abilities. Mol Gen Genet **119:**75-88.

29. **Downs DM, Petersen L.** 1994. apbA, a new genetic locus involved in thiamine biosynthesis in *Salmonella typhimurium*. J Bacteriol **176:**4858-4864.

30. **Bartolomé B, Jubete Y, Martinez E, de la Cruz F.** 1991. Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. Gene **102:**75-78.

31. **Guzman LM, Belin D, Carson MJ, Beckwith J.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. Journal of Bacteriology **177:**4121-4130.

32. **Koop AH, Hartley ME, Bourgeois S.** 1987. A low-copy-number vector utilizing beta-galactosidase for the analysis of gene control elements. Gene **52:**245-256.

33. **Caetano-Anolles G.** 1993. Amplifying DNA with arbitrary oligonucleotide primers. PCR Methods Appl **3:**85-94.

34. **Ernst DC, Lambrecht JA, Schomer RA, Downs DM.** 2014. Endogenous synthesis of 2-aminoacrylate contributes to cysteine sensitivity in *Salmonella enterica*. J Bacteriol **196:**3335-3342.

35. **McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK.** 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature **413:**852-856.

36. **Griffith KL, Wolf RE, Jr.** 2002. Measuring beta-galactosidase activity in bacteria: cell growth, permeabilization, and enzyme assays in 96-well arrays. Biochem Biophys Res Commun **290:**397-402.

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

38. **Maisnier-Patin S, Andersson DI.** 2004. Adaptation to the deleterious effects of antimicrobial drug resistance mutations by compensatory evolution. Res Microbiol **155:**360-369.

39. **Portnoy VA, Bezdan D, Zengler K.** 2011. Adaptive laboratory evolution--harnessing the power of biology for metabolic engineering. Curr Opin Biotechnol **22:**590-594.

40. **Koenigsknecht MJ, Fenlon LA, Downs DM.** 2010. Phosphoribosylpyrophosphate synthetase (PrsA) variants alter cellular pools of ribose 5-phosphate and influence thiamine synthesis in *Salmonella enterica*. Microbiology **156:**950-959.

41. **Koenigsknecht MJ, Lambrecht JA, Fenlon LA, Downs DM.** 2012. Perturbations in Histidine Biosynthesis Uncover Robustness in the Metabolic Network of *Salmonella enterica*. PLoS One **7:**e48207.

42. **Schmitz GE, Downs DM.** 2006. An allele of *gyrA* prevents *Salmonella enterica* serovar Typhimurium from growing with succinate as a carbon source. Journal of Bacteriology **188:**3126-3129.

43. **Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi H, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJ, Scheremetjew M, Tate J, Thimmajanarthanan M, Thomas PD, Wu CH, Yeats C, Yong SY.** 2012. InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res **40:**D306-312.

44. **Richards GR, Patel MV, Lloyd CR, Vanderpool CK.** 2013. Depletion of glycolytic intermediates plays a key role in glucose-phosphate stress in *Escherichia coli*. J Bacteriol **195:**4816-4825.

45. **Vinopal RT, Clifton D, Fraenkel DG.** 1975. PfkA locus of *Escherichia coli.* J Bacteriol **122:**1162-1171.

46. **Kimata K, Tanaka Y, Inada T, Aiba H.** 2001. Expression of the glucose transporter gene, ptsG, is regulated at the mRNA degradation step in response to glycolytic flux in *Escherichia coli*. EMBO J **20:**3587-3595.

47. **Morita T, El-Kazzaz W, Tanaka Y, Inada T, Aiba H.** 2003. Accumulation of glucose 6-phosphate or fructose 6-phosphate is responsible for destabilization of glucose transporter mRNA in *Escherichia coli.* J Biol Chem **278:**15608-15614.

48. **Way JC, Davis MA, Morisato D, Roberts DE, Kleckner N.** 1984. New Tn10 derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene **32:**369-379.

Table A.1. Bacterial strains.

Strain Number	**Genotype**
DM13623	Δ*thiC1225* Δ*araCBAD* pDM1336
DM13631	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d(Tc)[a] *sgrR1* pDM1336
DM13632	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d(Tc) pDM1336
DM14419	Δ*thiC1225* Δ*araCBAD* pDM1381
DM14422	Δ*thiC1225* Δ*araCBAD* Δ*sgrR3*::Kn pDM1381
DM14426	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d(Tc) *sgrR1* pDM1381
DM14427	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d(Tc) pDM1381
DM14430	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d *sgrR2* pDM1381
DM14431	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d(Tc) pDM1381
DM14443	Δ*thiC1225* Δ*araCBAD pgi*::Tn*5*(Kn) pDM1381
DM14517	Δ*thiC1225* Δ*araCBAD ptsG4152::*Tn*10*d(Tc) pDM1381
DM14531	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d(Tc) pDM1402
DM14532	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d(Tc) *sgrR1* pDM1402
DM14533	Δ*thiC1225* Δ*araCBAD zxx10167*::Tn*10*d(Tc) *sgrR2* pDM1402
DM14534	Δ*thiC1225* Δ*araCBAD* Δ*sgrR3*::Kn pDM1402
DM14540	Δ*thiC1225* Δ*araCBAD* Δ*pfkA*::Kn pDM1381
DM14668	Δ*thiC1225* Δ*araCBAD ptsG4152::*Tn*10*d(Tc) Δ*sgrR3*::Kn pDM1381
DM14896	Δ*thiC1225* Δ*araCBAD ptsG4152::*Tn*10*d(Tc) Δ*sgrS6*::Kn pDM1381
DM14932	Δ*thiC1225* Δ*araCBAD ptsG4152::*Tn*10*d(Tc) Δ*alaC251*::Kn pDM1381
DM14980	Δ*thiC1225* Δ*araCBAD pgi*::Tn*5*(Kn) pDM1402
DM14985	Δ*thiC1225* Δ*araCBAD* Δ*pfkA*::Kn pDM1402

[a]Tn*10*d(Tc) refers to the transposition-defective mini-Tn*10*(Tn*10*Δ16Δ17 tet^R) (48)

Table A.2. Plasmid and primers.

Plasmid Name	**Description**
pDM1336	pSU18-*THI5* (encodes Thi5p$^{M37I, A138V, G152D}$)
pDM1381	pSU18-*THI5*
pDM1361	pBAD18S-*THI5*
pDM1373	pET14b-*THI5*
pDM1402	pFZY1-*sgrSp*
Primer Name	**Sequence**
Sc THI5 for SacI pBAD18	AAGAGCTCTGATGTCTACAGACAAGATC
Sc THI5 Rev HindIII	TTTAAGCTTTTAAGCTGGAAGAGCCAATC
sgrSp_KpnI_F	ATGGTACC CATAAAAGGGGAACTC
sgrSp_BamHI_R	ATAGGATCC CGAAAGATATTATTGGC
S. cerevisiae Thi5 for 5'	CATATGATGTCTACAGACAAGATCAC
S. cerevisiae Thi5 rev 3'	CTCGAGTTAAGCTGGAAGAGCCAATC
pfkA_wanner_for	CATTCCAAAGTTCAGAGGTAGTCATGATTAAG AAAATCGGGTGTAGGCTGGAGCTGCTTC
pfkA_wanner_rev	CCATCAGGCGCGCAAAAACAATCAGTACAGT TTTTTCGCGCATATGAATATCCTCCTTAG
yabN (sgrR) wanner for	TTTTCATCGGAGTTCCCCTTTTATGCCCTCAGG TCGCCTGGTGTAGGCTGGAGCTGCTTC
yabN (sgrR) wanner rev	CAGCAATCAAGAGCTGGCGTTAAGGATCTGG CGGCGCAAACATATGAATATCCTCCTTAG
yabN(sgrR)_NcoI_for	ATATCCATGGCAATGCCCTAGGTCGC
yabN(sgrR)_XbaI_rev	GCGATCTAGATTAAGGATCTGGCGGCGC
sgrS wanner for	GCAATTTTATTATCCCTATATTAGGCCAATAA TATCTTTCGTGTAGGCTGGAGCTGCTTC
sgrS wanner rev	ATTTCGGCTGTTTCTGGATGACGATGATGGGA CGGCGTTTCATATGAATATCCTCCTTAG
alaC(yfdZ) wanner for	ACGTTAATCTGAGGATATTATGGCTGACTTCC GCCCTGAAGTGTAGGCTGGAGCTGCTTC
alaC(yfdZ) wanner rev	GGGGCTCCTGTTTGCTTTTCTACTCCGTGCTCG CTTCAACCATATGAATATCCTCCTTAG

Table A.3. Thi5p-dependent growth on different carbon sources

Carbon Source	Growth Rate[a] (h^{-1})		Final Cell Yield[b]	
	- Thiamine	+ Thiamine	- Thiamine	+ Thiamine
Acetate (50 mM)	0.01 ± 0.01	0.16 ± 0.02	0.10 ± 0.01	0.60 ± 0.06
Fructose	0.07 ± 0.04	0.55 ± 0.02	0.23 ± 0.07	0.74 ± 0.01
Fructose-6-P	0.04 ± 0.01	0.64 ± 0.02	0.15 ± 0.01	0.65 ± 0.01
Fumarate(50 mM)	0.03 ± 0.03	0.07 ± 0.05	0.16 ± 0.01	0.21 ± 0.01
Galactose	0.03 ± 0.01	0.54 ± 0.04	0.13 ± 0.01	0.75 ± 0.01
Gluconate	0.02 ± 0.01	0.57 ± 0.05	0.11 ± 0.01	0.52 ± 0.01
Glucose	0.01 ± 0.01	0.59 ± 0.03	0.12 ± 0.02	0.76 ± 0.02
Glucose-6-P	0.06 ± 0.02	0.61 ± 0.08	0.12 ± 0.02	0.60 ± 0.01
Glycerol	0.12 ± 0.06	0.50 ± 0.03	0.25 ± 0.09	0.81 ± 0.01
Malate (40 mM)	0.01 ± 0.01	0.40 ± 0.02	0.12 ± 0.03	0.64 ± 0.01
Mannose	0.19 ± 0.05	0.53 ± 0.01	0.59 ± 0.09	0.74 ± 0.01
Pyruvate (50 mM)	0.02 ± 0.01	0.41 ± 0.03	0.12 ± 0.01	0.86 ± 0.01
Ribose	0.15 ± 0.01	0.43 ± 0.04	0.51 ± 0.03	0.67 ± 0.01
Succinate(20 mM)	0.05 ± 0.03	0.12 ± 0.04	0.19 ± 0.05	0.26 ± 0.08
Xylose	0.15 ± 0.01	0.36 ± 0.01	0.64 ± 0.01	0.73 ± 0.01

[a]Growth rate is reported as μ, which is equal to $\ln(X/X_0)/T$, where X is OD_{650}, X_0 is the initial OD_{650} value of the period analyzed during exponential growth, and T is the time (in h). DM14419 (Δ*thiC1225* Δ*araCBAD* p*THI5*) was grown at 37°C in minimal medium for 36.5 h with each carbon source at 66 mM available carbon units unless otherwise indicated. Data represents averages and standard deviations of three independent cultures.
[b]Maximum OD_{650}.

Table A.4. Thi5p-dependent growth on glucose in strains disrupted in glycolysis.[a]

Strain	Relevant genotype	Growth Rate[a] (h^{-1})		Final Cell Yield[b]	
		-Thi	+Thi	-Thi	+Thi
DM14419	*thiC* p*THI5*	0.01 ± 0.01	0.60 ± 0.01	0.11 ± 0.01	0.84 ± 0.01
DM14443	*pgi thiC* p*THI5*	0.09 ± 0.01	0.10 ± 0.01	0.66 ± 0.01	0.87 ± 0.01
DM14540	*pfkA thiC* p*THI5*	0.11 ± 0.03	0.14 ± 0.01	0.65 ± 0.07	0.90 ± 0.02

[a]Growth rate is reported as μ, which is equal to $\ln(X/X_0)/T$, where X is OD_{650}, X_0 is the initial OD_{650} value of the period analyzed during exponential growth, and T is the time (in h). [b]Maximum OD_{650}. Strains DM14419, DM14443 and DM14540 were grown with shaking at 37°C in minimal medium with 11 mM glucose. Data represents average and standard deviations of three independent cultures.

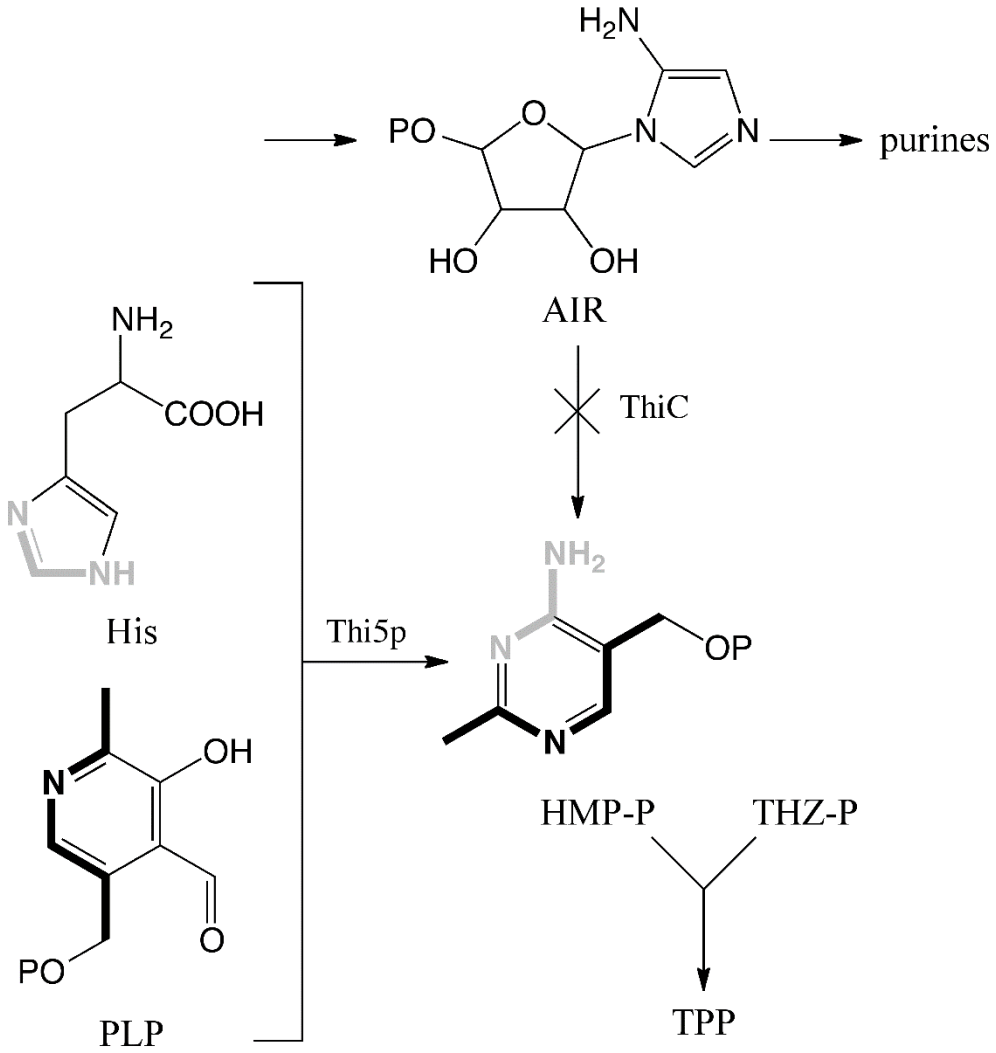


Figure A.1. Schematic of Thi5p-dependent thiamine biosynthesis in *S. enterica.* The genes encoding the relevant enzymes are shown. Atoms in HMP-P are shaded to indicate their origin, with bold indicating PLP, and grey indicating histidine. Abbreviations: AIR, 5-aminoimidazole ribotide; His, histidine; PLP, pyridoxal phosphate; HMP-P, 2-methyl-4-amino-5-hydroxymethylpyrimidine phosphate; THZ-P, 4-methyl-5-hydroxyethylthiazole phosphate.

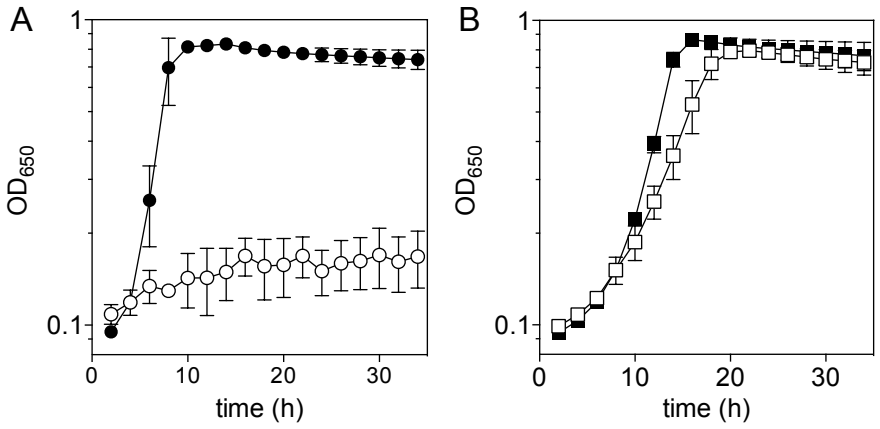


Figure A.2. Suppressor alleles allow Thi5p-dependent growth on minimal glucose medium. The *thiC* p*THI5* strain and derivative carrying *sgrR1* were grown in minimal glucose medium without (empty symbols) and with thiamine (filled symbols). The strains shown are (**A**) DM13623 (Δ*thiC*/p*THI5*) and (**B**) DM13631 (Δ*thiC sgrR1* /p*THI5*) Growth was monitored by optical density at 650 nm with shaking at 37°C. Data represent averages and standard deviations of three independent cultures.

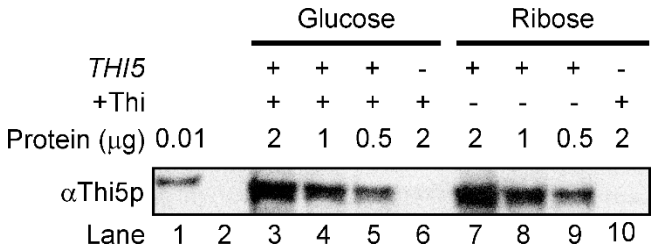


Figure A.3. Thi5p accumulates in minimal medium with glucose or ribose as carbon source. Strains containing pSU18 expressing *THI5* (DM14419; lanes 3-5, 7-9) or the empty vector (DM14148; lanes 6 and 10) were grown overnight in minimal medium with chloramphenicol (5 mg/L) and glucose or ribose, as indicated. Thiamine was included when necessary for growth, as indicated. Lane 1 contained 12.5 ng purified His_6Thi5p, and lane 2 was empty.

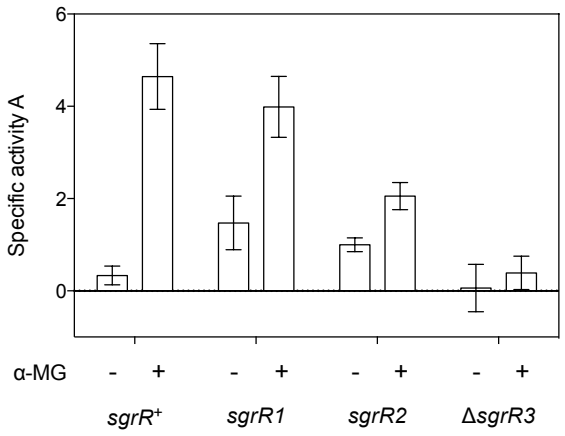


Figure A.4. Effects of *sgrR* alleles on transcription from the *sgrS* promoter. β-galactosidase assays were performed on cultures grown in NB medium containing ampicillin (30 mg/L) to mid-log phase and then incubated with and without α-MG (0.5%) for 45 minutes before permeabilization. The data are reported in Specific activity A (1000 X A_{420} min^{-1} OD_{600}^{-1}, where OD_{600} was determined from a 175-µL cell sample in a flat-bottom 96-well plate using a Spectramax 385) and represent averages and standard deviations from three independent cultures. These results are representative of four independent experiments.

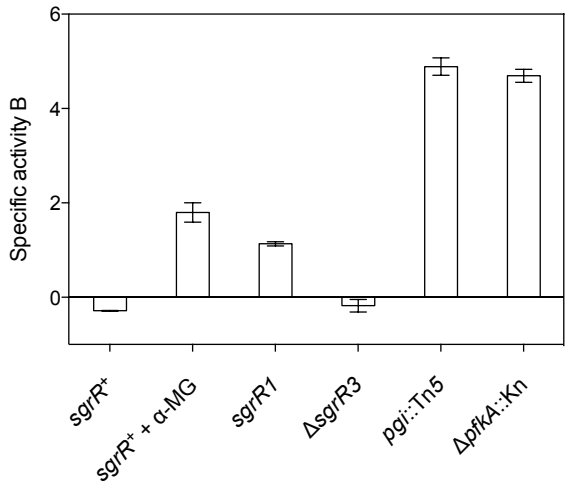


Figure A.5. Suppressor mutations affect transcription from the *sgrS* promoter. β-galactosidase assays were performed on cultures grown in minimal glucose medium containing ampicillin (15 mg/L), thiamine, and 1% α-MG as indicated to mid-log phase before permeabilization. The data are reported in Specific activity B (1000 X A_{420} min^{-1} OD_{600}^{-1}, where OD_{600} was determined from a 5-mL culture in 18 X 150 mm borosilicate tubes using a Spectronic 20+) and represent averages and standard deviations from three independent cultures.

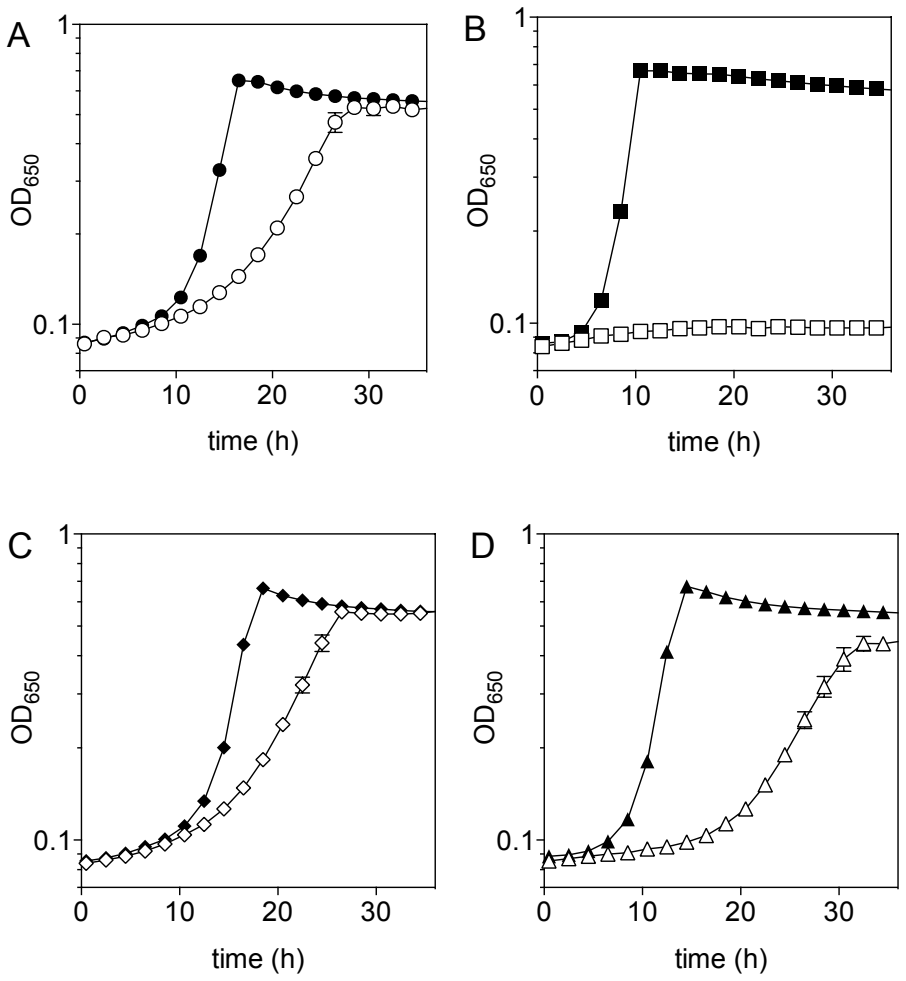


Figure A.6 – An insertion in *ptsG* allows Thi5p-dependent growth on glucose.
The *thiC ptsG* p*THI5* strain and derivatives with insertions or deletions of *sgrR*, *sgrS* and *alaC* were grown in minimal glucose medium without (empty symbols) and with thiamine (filled symbols). The strains shown are (**A**) DM14517 (*thiC ptsG* /p*THI5*), (**B**) DM14668 (*thiC ptsG sgrR* /p*THI5*), (**C**) DM14896 (*thiC ptsG sgrS* /p*THI5*) and (**D**) DM14932 (*thiC ptsG alaC* /p*THI5*). Growth was monitored by optical density at 650 nm with shaking at 37°C. Data represent averages and standard deviations of two independent cultures and are representative of three experiments.