Characterizing a serendipitous pathway in *E. coli* that can overcome a block in pyridoxal-5'-phosphate synthesis

By Hugh Haddox
MCDB, University of Colorado

Defense date:
April 9, 2012

Thesis advisor:
Dr. Shelley Copley, MCDB

Defense committee:
Dr. Shelley Copley, MCDB
Dr. Johannes Rudolph, CHEM
Dr. Christy Fillman, MCDB

Table of Contents

Abstract

Physiologically irrelevant metabolic pathways patched together from the reactions of multiple promiscuous enzymes can serve as a starting points for the evolution of novel pathways that degrade or generate new or pre-existing metabolites. It has been shown that at least three such 'serendipitous' pathways in *E. coli* can be evolved to over-come a block in the synthesis of pyridoxal-5'-phosphate (PLP). These pathways restore PLP synthesis by converting metabolites from core metabolism to intermediates in PLP synthesis that are downstream of the block, and can be elevated to a physiologically relevant level by the over-production of certain enzymes of the pathway.

Over-production of HisB, Php, or YjbQ complements a strain of *E. coli* with a deletion in *pdxB*, which encodes an enzyme that is essential in PLP biosynthesis, by elevating flux through one or more serendipitous pathways that restore PLP synthesis by producing a novel source of 2-oxo-3-hydroxy-4-phosphobutanoate—an intermediate in PLP biosynthesis downstream of the step catalyzed by *pdxB*. The aim of this honors thesis was to further characterize the serendipitous pathway(s) involved in this process.

First, this thesis indicated that HisB, Php, and YjbQ operate in the same serendipitous pathway. Although HisB, Php, and YjbQ are not epistatic, since over-production of each enzyme still complements a deletion of *pdxB* when genes encoding the other two enzymes are deleted, complementation by each enzyme is inhibited by a single metabolite, indicating that HisB, Php, and YjbQ participate in a single serendipitous pathway that is down-regulated by this metabolite.

Second, I determined which catalytic domain of the bi-functional HisB enzyme catalyzes the promiscuous activity in the serendipitous pathway that restores PLP synthesis upon HisB over-production. When the two catalytic domains of HisB were individually over-produced, only the imidazoleglycerol-phosphate dehydratase domain complemented a deletion in *pdxB*. Moreover,

mutating active site residues from this domain either partially or completely interfered with complementation by over-production of the full length HisB protein.

Lastly, I described a project that I am currently completing, which uses genome-wide transposon mutagenesis to search for enzymes that participate in the same serendipitous pathway as HisB in rescuing PLP synthesis. I screened an insertion library for cells that could no longer synthesize PLP using this serendipitous pathway and am in the process of locating transposons in the genomic DNA of select insertion mutants.

Chapter I: Introduction

Promiscuous enzymes and serendipitous pathways

Enzymes have evolved to have high catalytic efficiency and in most cases, high substrate specificity. To be highly efficient catalysts, enzymes bind substrates and transition states in very precise orientations relative to reactive components in the active site. This precision occurs at the level of sub-angstroms, where exact positioning of molecular orbitals influences catalytic efficiency. Catalysis is also often enhanced by binding pockets that make these precise interactions along large portions of the substrate/transition state. It follows that enzymes are usually highly substrate specific, since even the way that very similar compounds bind the active site could be sub-optimal for catalysis.

Still, high substrate specificity does not wholly prevent enzymes from catalyzing secondary reactions with alternative substrates. The adventitious binding of an alternative substrate in a favorable orientation can result in catalysis. In fact, most, if not all, enzymes are catalytically promiscuous, meaning that they can catalyze reactions other than the one(s) they have evolved to catalyze in cell physiology.

Promiscuous activities are defined as physiologically irrelevant (Copley, 2011). In contrast, physiologically relevant activities are under selective pressure to be maintained in the cell and operate at some specific level. Enzymes that catalyze multiple physiologically relevant reactions are classified as 'broadly specific' instead of 'promiscuous.'

Promiscuous activities are usually orders of magnitude slower than well-evolved activities; however, these activities can be orders of magnitude faster than uncatalyzed reactions, making promiscuous enzymes good starting points for the evolution of new, proficient enzymes. The first step in the evolution of a promiscuous enzyme is for the promiscuous activity to be promoted to a physiologically relevant level by either a mutation or a change in environment. Promiscuous activities can be irrelevant based on a variety of factors that influence the rate of enzyme-catalyzed

reactions, including: k_{cat}, K_M, and the levels of substrate, enzyme, and inhibitors. A mutation that increases the reaction rate by changing any one of these factors could raise a promiscuous activity to a physiologically relevant level. Alternatively, if a promiscuous enzyme is never exposed to the promiscuous substrate, a mere change in environment that exposes the cell to this compound for the first time could make the promiscuous activity relevant.

Once a promiscuous activity becomes physiologically relevant, it is subject to selective pressure, and if the activity is beneficial to the organism's fitness, it will be under selective pressure to improve. There is functional constraint in evolving a single copy of a gene that is under selective pressure to improve a promiscuous activity and still retain the 'normal' activity since the active site must mediate both activities. Ultimately, this constraint can be relieved by gene duplication, which allows the two copies to each specialize in a single function.

In the intermediate stages of evolving a promiscuous enzyme, improving the promiscuous activity through the accumulation of beneficial amino acid changes can be accelerated if its gene is amplified in copy number (Copley, 2011). Initial amplification of the gene is directly advantageous since increasing gene dosage is one way to raise the total corresponding promiscuous activity within the cell. This process is also *indirectly* beneficial since a high copy number gene is able to explore more sequence space than a single-copy gene. Not only can multiple alleles explore sequence space at a faster rate than a single allele, but mutations from different alleles can be combined from homologous recombination, which can serve as mechanism to consolidate beneficial mutations into a single copy. Finally, when a single allele evolves that encodes an enzyme that is proficient in the promiscuous activity, the cell is under selective pressure to pare down the copy number to reduce the metabolic burden from high copy number.

The above considered, when evolving an enzyme with a new function there are two notable reasons why a promiscuous enzyme is a better starting point than an enzyme that entirely lacks the desired activity. If there is selective pressure for a new enzymatic activity, the corresponding

activity of an evolving enzyme must reach a minimum threshold to become physiologically relevant and thus under selection to further improve. The first advantage of promiscuous enzymes is that their activities are orders of magnitude faster than uncatalyzed reactions, and so orders of magnitude closer to this threshold. Considering that evolution is a stepwise process, it would be much more likely for a single amino acid change to improve a promiscuous function by a relatively small amount than for this change to improve a non-existent function by a relatively larger amount (orders of magnitude larger). The second advantage of promiscuous enzymes is that gene duplication can allow promiscuous enzymes, but not non-promiscuous enzymes (increasing the concentration of an enzyme with zero activity will not increase that activity), to cross the activity threshold. As described above, duplication and further amplification accelerates the process of evolution. So, not only do promiscuous enzymes that are evolving have more ways to cross the activity threshold than non-promiscuous enzymes, but evolution is more likely to occur using these paths.

Promiscuity is also relevant in the evolution of new metabolic pathways. Two promiscuous enzymes could form a low-flux pathway if the product of one promiscuous activity is the substrate of another. Pathways that are patched together from multiple promiscuous enzymes are called 'serendipitous' pathways. Like promiscuous activities, serendipitous pathways are, by definition, physiologically irrelevant and are not an evolved part of the cell's metabolic network. Also analogous to promiscuous activities, these pathways can serve as starting points for the evolution of new, physiologically relevant metabolic pathways. Given that microbes contain hundreds of enzymes, it seems likely that the proteome has many underlying serendipitous pathways available for recruitment.

Specific Context and Aims of Thesis

The motivation of this thesis, and the larger project in the Copley lab that this thesis relates to, is to identify promiscuous enzymes and serendipitous pathways within *E. coli* that can be used

to adapt to a block in an essential metabolic pathway. This block was made by deleting *pdxB*, which encodes an enzyme involved in the *de novo* biosynthesis of pyridoxal phosphate (PLP) (Figure 1.1), an enzyme cofactor essential to all life (Mukherjee *et al*, 2011). PLP participates in a diversity of processes including, transamination, racemization, decarboxylation, substitution, and elimination reactions, and is associated with over 60 enzymes in *E. coli* (Kim *et al*, 2010). Although *E. coli* can salvage PLP precursors on LB medium, the pathway for *de novo* biosynthesis of PLP is essential for growth on M9/glucose.

A publication from the Copley lab, Kim *et al*, 2010, describes at least three serendipitous pathways in *E. coli*, patched together from multiple promiscuous enzymes, that are able to complement the *pdxB* deletion. In this study, a library of *E. coli* ORFs cloned into expression vectors (ASKA clone library, Baba *et al*, 2006) was transformed into *E. coli* Δ*pdxB* in a multi-copy suppression experiment that screened for growth on M9/glucose. The over-production of enzymes in this screening elevates their promiscuous activities, allowing potentially restorative activities to reach physiologically relevant levels. This screen revealed seven enzymes that complemented the Δ*pdxB* strain: PdxA, AroB, ThrB, YeaB, HisB, YjbQ, and Php. Interestingly, none of these enzymes restore PLP biosynthesis by catalyzing the *pdxB* reaction. Instead, these enzymes are distributed between at least three serendipitous pathways that shunt material from *E. coli* metabolism into PLP biosynthesis, downstream of the *pdxB* step.

Pathway #1 consists of three enzyme catalyzed steps and one non-enzymatic step that transforms 3PHP, an intermediate in serine synthesis, to 4PHT, an intermediate in PLP synthesis (Figure 1.2) (Kim *et al*, 2010). Kim *et al*, 2010 uses a variety of experiments to demonstrate that the enzymes and intermediates in Figure 1.2 constitute a *bona fide* serendipitous pathway *in vivo* that is able to complement the Δ*pdxB* strain on M9/glucose upon the over-production of select enzymes or the addition of pathway intermediates to the growth medium.

The aim of my thesis is to better characterize pathway #2. The over-expression of *hisB*, *php*, and *yjbQ* each lead to the production of 2-oxo-3-hydroxy-4-phosphobutanoate (OHPB), but little more is known about this pathway. Indeed, it is possible that the enzymes encoded in these genes participate in different serendipitous pathways that each generate OHPB. It is also possible that the serendititous pathway(s) involving *hisB*, *php*, and *yjbQ* include additional promiscuous enzymes. HisB is also a bi-functional enzyme with two structurally discrete catalytic domains, posing the question: which catalytic domain mediates the promiscuous activity involved in complementation of the Δ*pdxB* strain?

In chapter II, my aim was to determine if HisB, Php, and YjbQ are in the same serendipitous pathway. I tested for epistasis between *hisB*, *php*, and *yjbQ* by determining if these genes complement strains of *E. coli* with deletions in *pdxB* and another of the three genes. There were no signs of epistasis, indicating that *hisB*, *php*, and *yjbQ* do not form a simple, linear serendipitous pathway. In this chapter, I also showed that complementation by *hisB*, *php*, and *yjbQ* is sensitive to glycine in each case, indicating that these genes participate in the same serendipitous pathway.

Chapter III describes a set of experiments that demonstrate the IGPD domain of HisB catalyzes a promiscuous activity involved in complementation. Finally, chapter IV describes how genome-wide transposon mutagenesis was used to probe for insertions that disrupt the serendipitous pathway involving *hisB*, in an effort to identify additional enzymes in this serendipitous pathway.

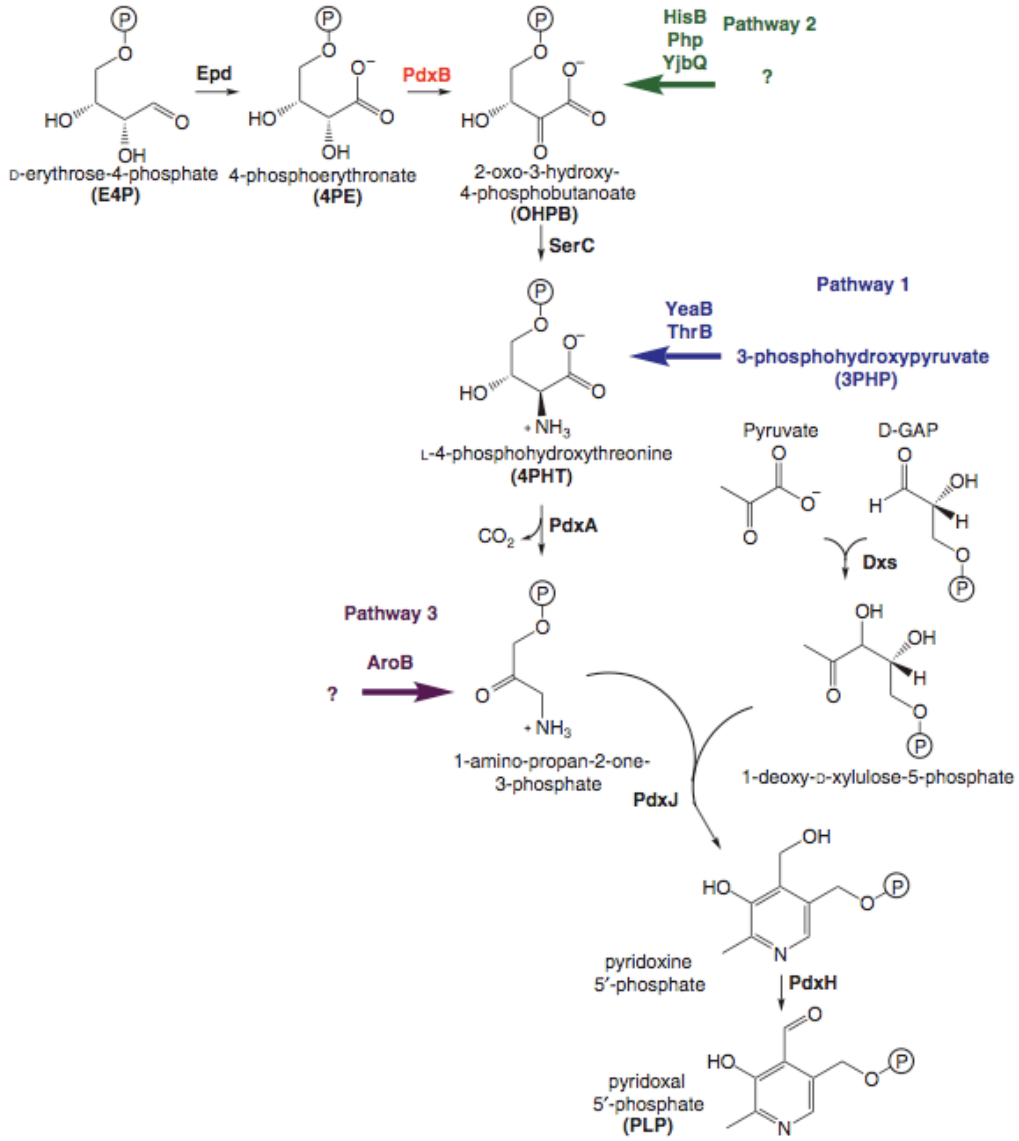


Figure 1.1. From Kim *et al*, 2010, showing the *de novo* biosynthetic pathway for pyridoxal-5'-phosphate (PLP) in *E. coli*.

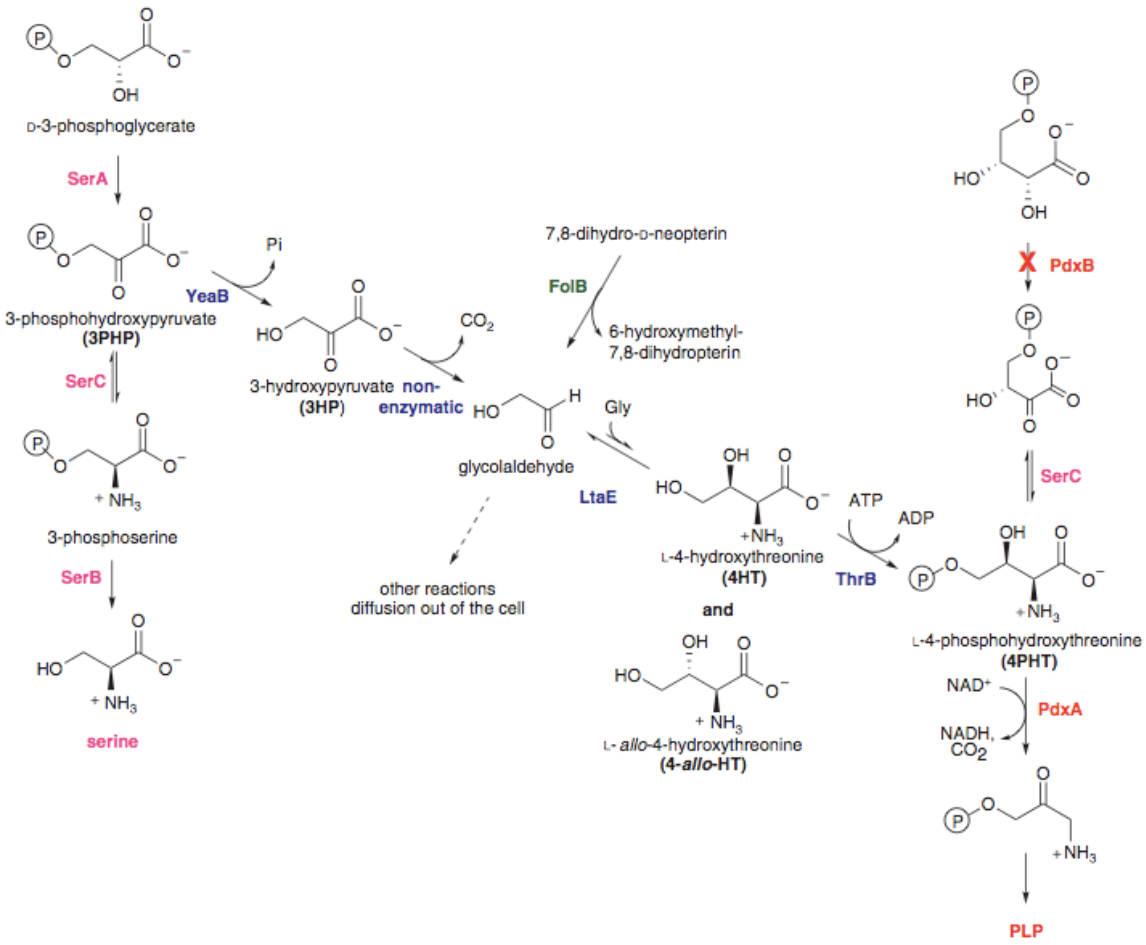


Figure 1.2. From Kim *et al*, 2010, showing how serendipitous pathway #1 converts 3PHP, an intermediate in serine biosynthesis, to 4PHT, an intermediate in PLP biosynthesis.

Chapter II: HisB, Php, and YjbQ appear to reside in the same serendipitous pathway despite showing no signs of epistasis

Introduction

Based on the results from Kim *et al*, 2010, there are two lines of evidence suggesting that *hisB*, *php*, and *yjbQ* take part in the same serendipitous pathway. First, neither gene is able to complement *E. coli* strains lacking *serC*, *pdxA*, *pdxJ*, or *pdxH*, indicating that *hisB*, *php*, and *yjbQ* are all involved in the production of 2-oxo-3-hydroxy-4-phosphobutanoate (OHPB) (Figure 1.1). If one of these genes was involved in the production of 1-amino-propan-2-one-3 phosphate, for example, it would be expected to complement strains deficient in *serC* and *pdxA*, but fail to complement deletions of *pdxJ* or *pdxH*. Second, out of the seven genes that were identified in the multi-copy suppression experiment that complement the Δ*pdxB* strain, only the over-expression of *hisB*, *php*, and *yjbQ* support growth on both liquid and solid medium. This physiological similarity suggests a mechanistic relatedness. Ultimately, however, it is unclear whether *hisB*, *php*, and *yjbQ* participate in the same serendipitous pathway.

This chapter describes two experiments directed towards determining whether *hisB*, *php*, and *yjbQ* operate in the same serendipitous pathway. One experiment tests for epistasis between *hisB*, *php*, and *yjbQ* by seeing if a deletion in one of these genes interferes with the ability of the other two, when over-expressed, to complement the Δ*pdxB* strain. Epistatic genes would be predicted to participate in the same serendipitous pathway. In comparison, nothing definitive can be said about non-epistatic genes. An obvious explanation is that these genes participate in separate pathways. However, these genes may act in the same pathway if they are inessential to the pathway because they are functionally redundant by other, undiscovered promiscuous enzymes.

I also tested whether *hisB*, *php*, and *yjbQ* are in the same serendipitous pathway by seeing if complementation by these genes is inhibited by the same metabolites. The flux through well-evolved pathways is usually tightly regulated by other proteins and metabolites. Although serendipitous pathways are not subject to this kind of coordinate regulation (serendipitous

pathways are not an integrated part of cell physiology), their flux will be responsive to metabolites or regulatory proteins that regulate the pathway's enzymes or source of material. The addition of a combination of amino acids, nucleotide bases, and vitamins inhibits complementation of the Δ*pdxB* strain by over-expression of each gene on M9/glucose. I used drop-out experiments to determine which nutrients are inhibitory. If the same nutrients are inhibitory in each case, *hisB*, *php*, and *yjbQ* most likely operate in the same serendipitous pathway; if inhibitory nutrients are different, the reverse is more likely.

Results

Creation of the double knockout strains: Δpdxb ΔhisB, Δpdxb Δphp, and Δpdxb ΔyjbQ

I constructed three double knockout strains (Δ*pdxB* Δ*hisB*, Δ*pdxB* Δ*php*, Δ*pdxB* Δ*yjbQ*) by replacing *hisB*, *php*, or *yjbQ* with a chloramphenicol resistance gene (cm^r) in the Δ*pdxB* strain. Figure 2.1 diagrams the strategy I used for the replacement of target genes with an antibiotic resistance cassette, but illustrates this process using a kanamycin resistance gene. In this figure, the antibiotic resistance cassette is removed from the genomic DNA in the final step using homologous recombination; however, I left the chloramphenicol resistance cassette intact in each double knockout strain.

To begin, I used PCR to amplify cm^r from the pACYC184 vector with primers with unique flanking sequences that were tailored to direct homologous recombination with the target gene. I visualized the PCR generated chloramphenicol resistance cassettes after electrophoresis on a 1% agarose gel, which appeared as single bands that approximately matched the expected size of 922 base pairs. Next, I transformed the Δ*pdxB* strain, first with a temperature sensitive vector encoding lambda-red recombinase and ampicillin resistance, the former of which was used to catalyze homologous recombination of the chloramphenicol resistance cassettes with genomic DNA, and then with one of the chloramphenicol resistance cassettes constructed to target either *hisB*, *php*, or *yjbQ*. Knockout strains were selected on LB plates with chloramphenicol. Finally, lambda-red

recombinase, was removed by growing cells at a high temperature, which rids cells of the temperature sensitive vector that was used to express this protein. Removal of this vector was confirmed by showing the loss of ampicillin resistance by the knockout strains.

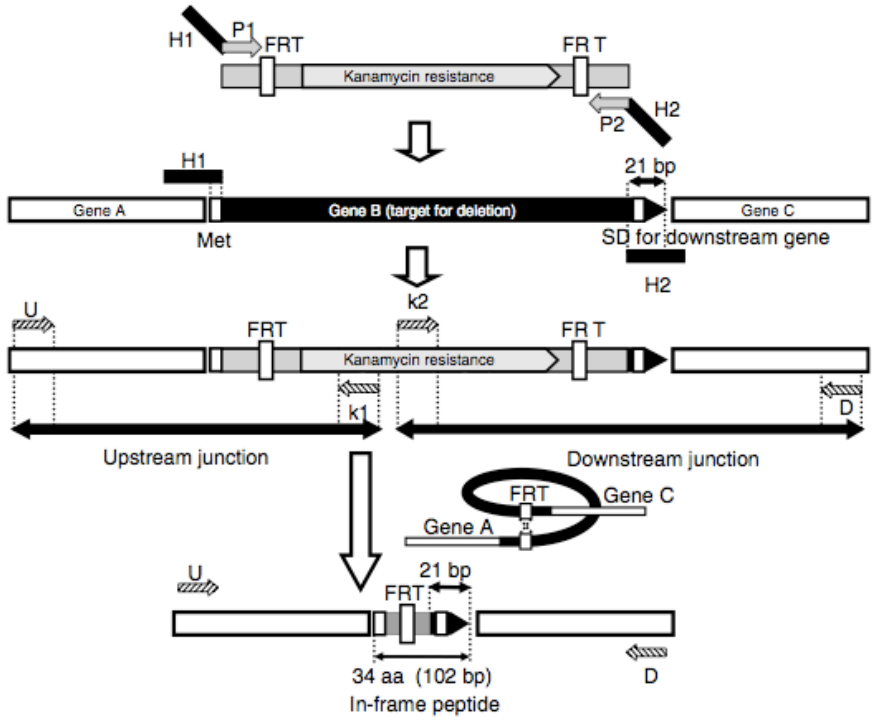


Figure 2.1. From Baba *et al*, 2006, diagramming a strategy for targeted replacement of genomic DNA with an antibiotic resistance cassette, in this case encoding a gene for kanamycin resistance. This method was used to replace *hisB*, *php*, and *yjbQ* in the Δ*pdxB* strain with a chloramphenicol resistance gene. Confirmation of gene deletions was confirmed using PCR, with pairs of primers similar to U/k1 and k2/D.

A series of PCR reactions were used to verify replacement of target genes with cm^r in each double knockout strain. For example, in the case of the Δ*pdxB* Δ*hisB* strain, a forward primer that anneals upstream of the *hisB* locus was paired with one of two reverse primers that anneal to sequences in either *hisB* or cm^r. Likewise a reverse primer that anneals downstream of the *hisB* locus was paired with one of two forward primers that anneal to *hisB* or cm^r. Using the Δ*pdxB* Δ*hisB* strain as a template in colony PCR, if *hisB* was truly replaced with cm^r, PCR should only generate a product if the reaction involved a primer that anneals with cm^r, since primers that anneal with *hisB* should not have a complementary sequence in the genomic DNA. As a control, each primer pair was

used in colony PCR of the Δ*pdxB* strain, where the opposite pattern is expected. I conducted this procedure with each double knockout strain and the PCR products were visualized after electrophoresis on a 1% agarose gel (Figures 2.3-2.5).

All PCR reactions that were expected to generate a product had, as their most prominent product, a band of DNA that approximately matched an anticipated length (Tables 2.4-2.6). These bands were always absent in PCR with the same pair of primers but with the other strain of *E. coli*, except in the PRC reaction 2y, which produced a faint band that corresponded to the prominent one generated in 6y. It is possible that this band was generated by the adventitious annealing of primers to unintended sequences. Alternatively, this band could result from colony PCR with a mixed population. However, considering that the other reactions that were used to confirm the replacement of *yjbQ* with cm^r followed expectations, it seems likely that this replacement was successful and that the population was not mixed.

Histidine is required for growth of the ΔpdxB ΔhisB strain on M9/glucose, but does not interfere with complementation of the pdxB deletion

Of *hisB*, *php*, and *yjbQ*, only *hisB* is essential for growth on M9/glucose. To account for this essentiality, histidine was added to M9/glucose in complementation experiments that involved the Δ*pdxB* Δ*hisB* strain. As a control, to ensure that the addition of histidine does not complement the *pdxB* deletion, I tested for growth of the Δ*pdxB* Δ*hisB* strain on M9/glucose in the presence of this amino acid. Visible colonies of this strain failed to appear within seven days. In another control experiment, I tested whether adding histidine to M9/glucose affected complementation of the Δ*pdxB* strain by *pdxB*, *hisB*, *php*, *yjbQ*, and pTrcHisB. In each case, complementation was indistinguishable in the presence and absence of histidine with respect to the number of days it took visible colonies to appear (data not shown).

Knocking out hisB, php, or yjbQ from the ΔpdxB strain does not interfere with complementation of the pdxB deletion by over-expression of the other two genes

I tested for epistasis between *hisB*, *php*, and *yjbQ* by seeing if a deletion in one of these genes interferes with complementation of the Δ*pdxB* strain by over-expression of the other two. Growth of double knockout strains during the over-expression of *pdxB*, *hisB*, *php*, or *yjbQ*, or when these strains were transformed with the blank pTrcHisB vector, was tested on M9/glucose, in the presence of histidine in cases involving the Δ*pdxB* Δ*hisB* strain (Table 2.1, Figure 2.6). All double knockout strains over-expressing *hisB*, *php*, or *yjbQ* grew within the same number of days as when these genes are over-expressed in the Δ*pdxB* strain.

Table 2.1. Growth of *E. coli* double knockout strains on plates containing M9/glucose at 37 ºC when the indicated genes were over-expressed from pTrcHisB.

Gene	Δ*pdxB* Δ*hisB*	Δ*pdxB* Δ*php*	Δ*pdxB* Δ*yjbQ*
pdxB	+2[a]	+2	+2
hisB	+2	+2	+2
php	+2	+2	+2
yjbQ	+5	+3	+3
no gene insert	-7	-7	-7

[a]+N or -N indicates growth or no growth, respectively, of visible colonies after N days.

Importantly, the deletions do not appear to disrupt native PLP synthesis, as the over-expression of *pdxB* complemented each double knockout strain. Neither does the act of deleting *hisB*, *php*, or *yjbQ* directly complement the *pdxB* deletion, since double knockout strains transformed with the blank pTrcHisB vector failed to produce visible colonies after seven days.

Dropping out glycine from M9/glucose supplemented with nutrients from EZ Rich Defined Medium restores complementation of the ΔpdxB ΔltaE strain by hisB, php, and yjbQ over-expression

Inhibition of complementation of a strain of *E. coli* with deletions in both *pdxB* and *ltaE* (hereafter referred to as the Δ*pdxB* Δ*ltaE* strain) by HisB, Php, and YjbQ was tested on M9/glucose with the nutrients from a drop-out medium called EZ Rich Defined Medium, which includes: the 20 L-amino acids; the nucleotide bases: adenine, guanine, cytosine, and uracil; and the vitamins: thiamine, pantothenate, *p*-aminobenzoic acid, *p*-hydroxybenzoic acid, and 2,3-dihydroxy benzoic acid (I refer to this complex medium as M9/glucose +EZ) (Neidhardt *et al*, 1974). I deleted *ltaE* to

protect against pseudo-reversion of the Δ*pdxB* strain using pathway #1, which occasionally occurs on M9/glucose (Kim *et al*, 2010). In growth inhibition experiments, growth was sometimes severely reduced and it was useful to know in these cases that no growth was due to pseudo-reversion.

I tested for the ability of the Δ*pdxB* Δ*ltaE* strain to grow on this medium when transformed with *pdxB*, *hisB*, *php*, *yjbQ*, or the pTrcHisB vector (Table 2.2; Figure 2.7). I also repeated these complementation experiments on the same medium, but with a ten-fold dilution in the amino acid concentration, which I refer to as M9/glucose +EZ (aa/10), as well as on M9/glucose, without the addition of any nutrients.

Table 2.2. Growth on plates containing M9/glucose and indicated supplements at 37 ºC of *E. coli* Δ*pdxB* Δ*ltaE* when cells are transformed with indicated gene inserts on pTrcHisB.

	Δ*pdxB* Δ*ltaE*				
Supplements	*pdxB*	*hisB*	*php*	*yjbQ*	no gene instert
M9/g	+[a]	+	+	+	-
+EZ (aa/10)	+	+/-	+/-	+/-	-
+EZ	+	-	-	-	-

[a]+ indicates growth, - indicates no growth, and +/- indicates abnormal, restricted growth

Both complex media inhibited complementation by *hisB*, *php*, and *yjbQ* over-expression. In the case of the M9/glucose +EZ complex medium, growth was severely inhibited. On M9/glucose +EZ (aa/10), growth was inhibited compared with growth on M9/glucose, but less severely than in the presence of the full strength amino acid supplement. Importantly, these nutrients did not directly complement the Δ*pdxB* Δ*ltaE* strain, as shown by the failure of this strain to grow when transformed with the blank pTrcHisB vector. Nor do these nutrients inhibit the 'normal' pathway for PLP synthesis, since *pdxB* over-expression promoted cell growth on all media.

I also determined that dropping out glycine from the M9/glucose +EZ (aa/10) re-enables the ability of over-production of each HisB, Php, and YjbQ to complement the Δ*pdxB* Δ*ltaE* strain (Figure 2.8). As positive and negative controls, I tested for complementation of the Δ*pdxB* Δ*ltaE* strain by transforming this strain with *pdxB* or the blank pTrcHisB vector, respectively. Over-

expression of *pdxB* complemented this strain within two days on each media, while cells transformed with the pTrcHisB vector failed to produce visible colonies within seven days.

Interestingly, glycine does not inhibit complementation of the Δ*pdxB* Δ*ltaE* strain by over-production of HisB when glycine is added to M9/glucose by itself (Table 2.3; Figure 2.9). Although complementation by *hisB* was once again shown to be ineffective on M9/glucose +EZ (aa/10) compared with M9/glucose or M9/glucose +EZ (aa/10) –Gly, the addition of glycine by itself to M9/glucose did not interfere with complementation.

Table 2.3. Growth on plates containing M9/glucose and indicated supplements at 37 ºC of *E. coli* Δ*pdxB* Δ*ltaE* when indicated genes are over-expressed from pTrcHisB

	Δ*pdxB* Δ*ltaE*		
Supplements	*pdxB*	*hisB*	no gene insert
--	+	+	-
+EZ (aa/10)	+	+/-	-
+EZ (aa/10) -Gly	+	+	-
+Gly	+	+	-

[a]+ indicates growth, - indicates no growth, and +/- indicates abnormal, restricted growth

Discussion

The double knockout complementation experiments indicate that HisB, Php, and YjbQ are not epistatic. Yet, the common effect that glycine has on complementation of the Δ*pdxB* Δ*ltaE* strain by over-production of each of these enzymes suggests that they take part in a single pathway that is down-regulated in the presence of this amino acid. Alternatives to this scenario require glycine to inhibit multiple serendipitous pathways, which is less likely. Glycine could easily down-regulate multiple serendipitous pathways by inhibiting enzymes in the 'normal' pathway for PLP synthesis that are downstream of the serendipitous branch-point; however, this possibility can be ruled out since the over-expression of *pdxB* complements the Δ*pdxB* Δ*ltaE* strain on the complex media, even in the presence of glycine.

Combined, these results suggest that HisB, Php, and YjbQ operate in the same serendipitous pathway, but that these enzymes are functionally redundant with other promiscuous enzymes, making HisB, Php, and YjbQ inessential to the pathway. Figure 2.2 diagrams three serendipitous

pathways that are structured in simple ways that are consistent with this interpretation. In this figure, the inessentiality of HisB, Php, and YjbQ arises from the presence of other promiscuous enzymes in *E. coli* that catalyze the same reaction. Possibly, the presence of a second promiscuous catalyst is unnecessary if the uncatalyzed reaction is itself efficient enough to allow an increase in pathway flux by over-production of another enzyme in the pathway.

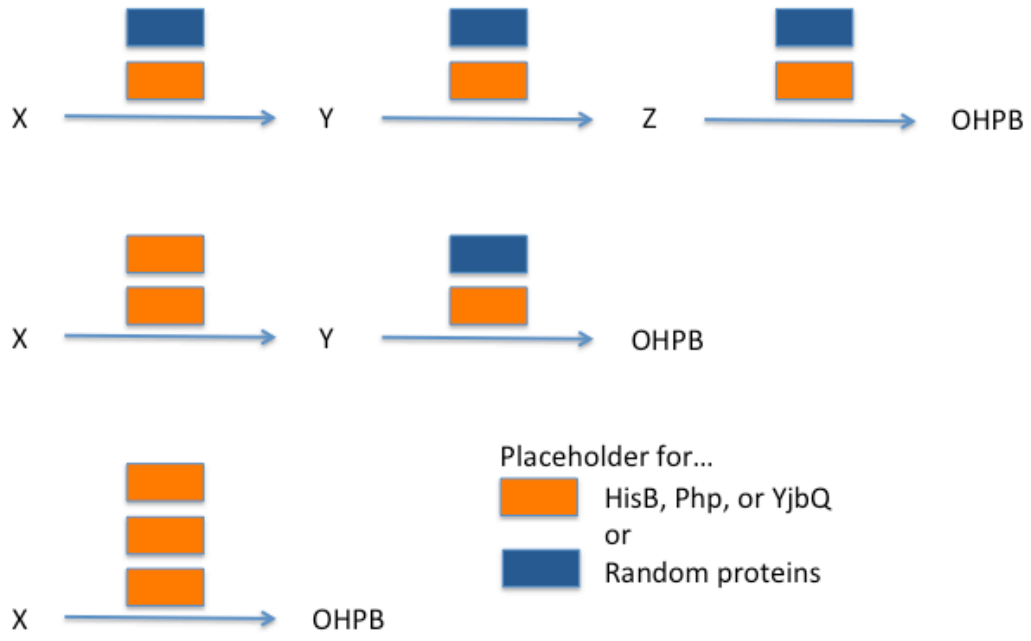


Figure 2.2. Three simple scenarios where HisB, Php, and YjbQ operate in the same serendipitous pathway to produce OHPB, but are not epistatic. Placeholders represent indicated proteins, which catalyze denoted steps in the serendipitous pathway.

Curiously, adding only glycine to M9/glucose did not inhibit complementation of the Δ*pdxB* Δ*ltaE* strain by HisB over-production. Possibly, this result indicates that glycine inhibits complementation in a way that involves other metabolites. In the future, it would be interesting to conduct additional dropout experiments to find which other nutrients glycine depends on to inhibit complementation. Alternatively, increasing the concentration of glycine as the sole supplement on M9/glucose would be sufficient to cause inhibition despite the absence of other nutrients.

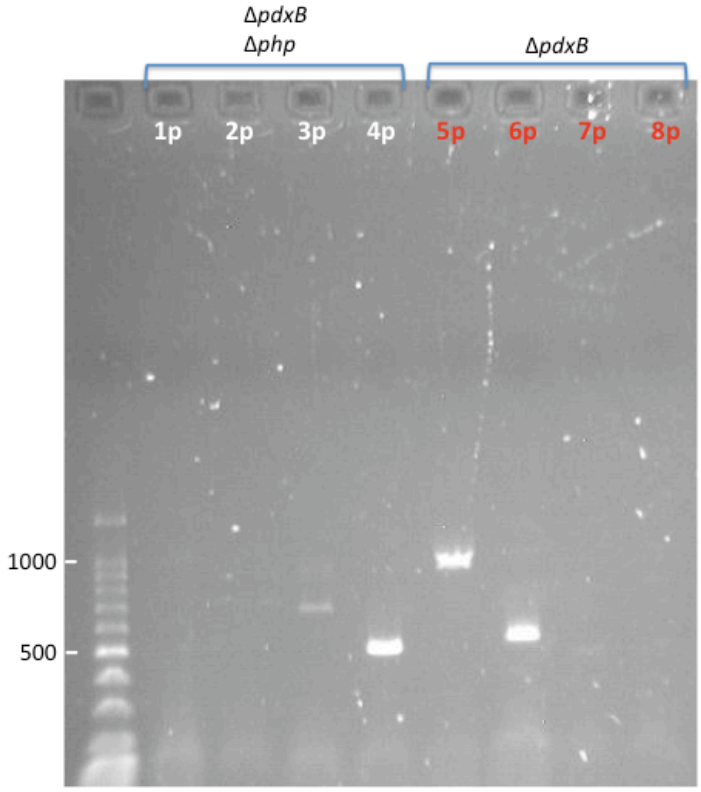


Figure 2.3. Visualization of the PCR reactions used to confirm the replacement of *php* by *cm*^{r}, along side a 100 bp ladder (Promega) after electrophoresis on a 1% agarose gel with ethidium bromide. The ladder's 500 bp and 1000 bp bands are denoted.

Table 2.4. Anticipated sizes of PCR products from indicated colony PCR reactions used to confirm the replacement of *php* by *cm*^{r} with indicated strains and primer pairs.

	Δ*pdxB* Δ*php*		Δ*pdxB*	
primer pairs	rxn	expected length (bp)	rxn	expected length (bp)
php-oF *php-iR*	1p	--	5p	966
php-iF *php-oR*	2p	--	6p	538
php-oF *cm-iR*	3p	667	7p	--
php-oR *cm-iF*	4p	497	8p	--

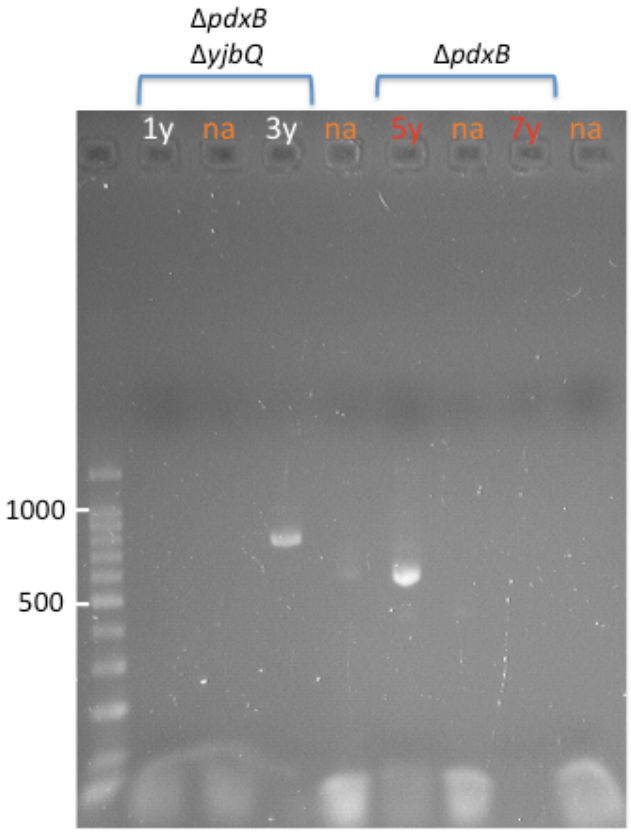


Figure 2.4. Visualization of the four of the PCR reactions used to confirm the replacement of *yjbQ* by *cm*^{r} (the other four reactions are shown in Figure 2.4), along side a 100 bp ladder (Promega) after electrophoresis on a 1% agarose gel with ethidium bromide. The ladder's 500 bp and 1000 bp bands are denoted.

Table 2.5. Anticipated sizes of PCR products from indicated colony PCR reactions used to confirm the replacement of *yjbQ* by *cm*^{r} with indicated strains and primer pairs.

	Δ*pdxB* Δ*yjbQ*		Δ*pdxB*	
primer pairs	rxn	expected length (bp)	rxn	expected length (bp)
yjbQ-oF *yjbQ-iR*	1y	--	5y	605
yjbQ-iF *ybjQ-oR*	2y	--	6y	501
yjbQ-oF *cm-iR*	3y	784	7y	--
yjbQ-oR *cm-iF*	4y	638	8y	--

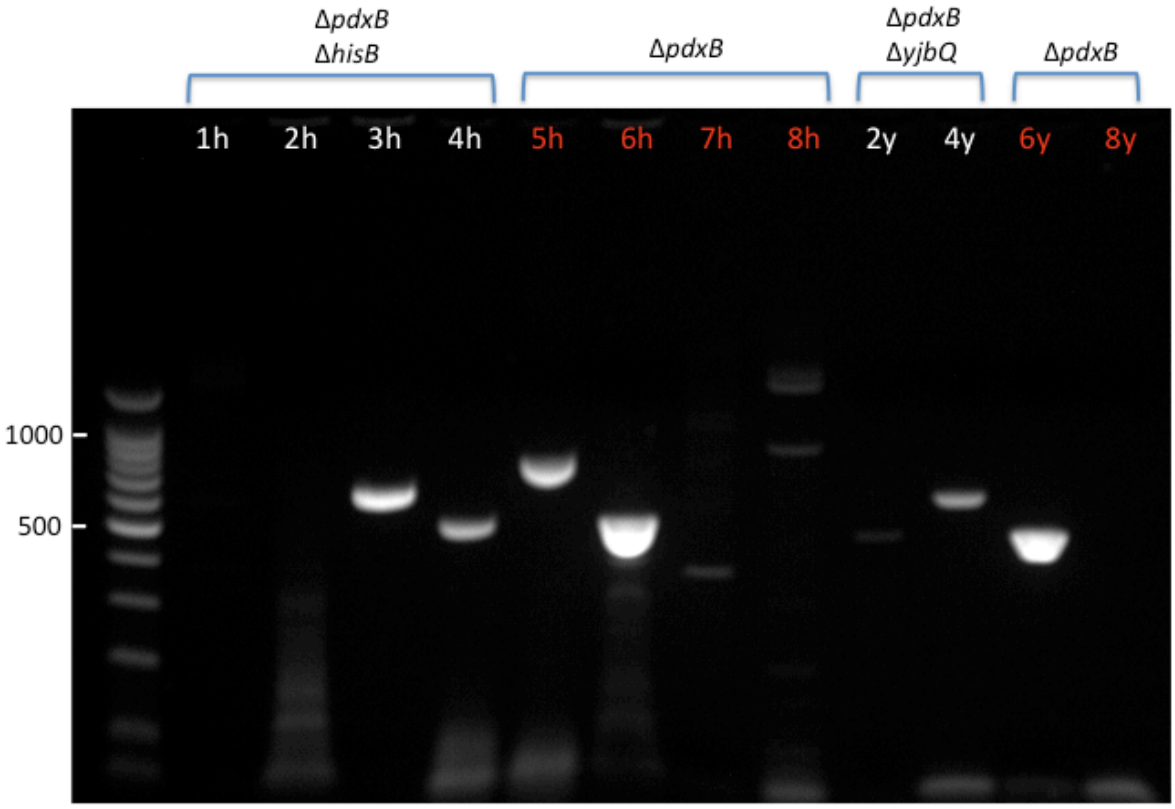


Figure 2.5. Visualization of the PCR reactions used to confirm the replacement of *hisB* or *yjbQ* by *cm*^{r}, along side a 100 bp ladder (Promega) after electrophoresis on a 1% agarose gel with ethidium bromide. The ladder's 500 bp and 1000 bp bands are denoted.

Table 2.6. Anticipated sizes of PCR products from indicated colony PCR reactions used to confirm the replacement of *hisB* by *cm*^{r} with indicated strains and primer pairs.

	Δ*pdxB* Δ*hisB*		Δ*pdxB*	
primer pairs	rxn	expected length (bp)	rxn	expected length (bp)
hisB-oF *hisB-iR*	1h	--	5h	787
hisB-iF *hisB-oR*	2h	--	6h	525
hisB-oF *cm-iR*	3h	613	7h	--
hisB-oR *cm-iF*	4h	507	8h	--

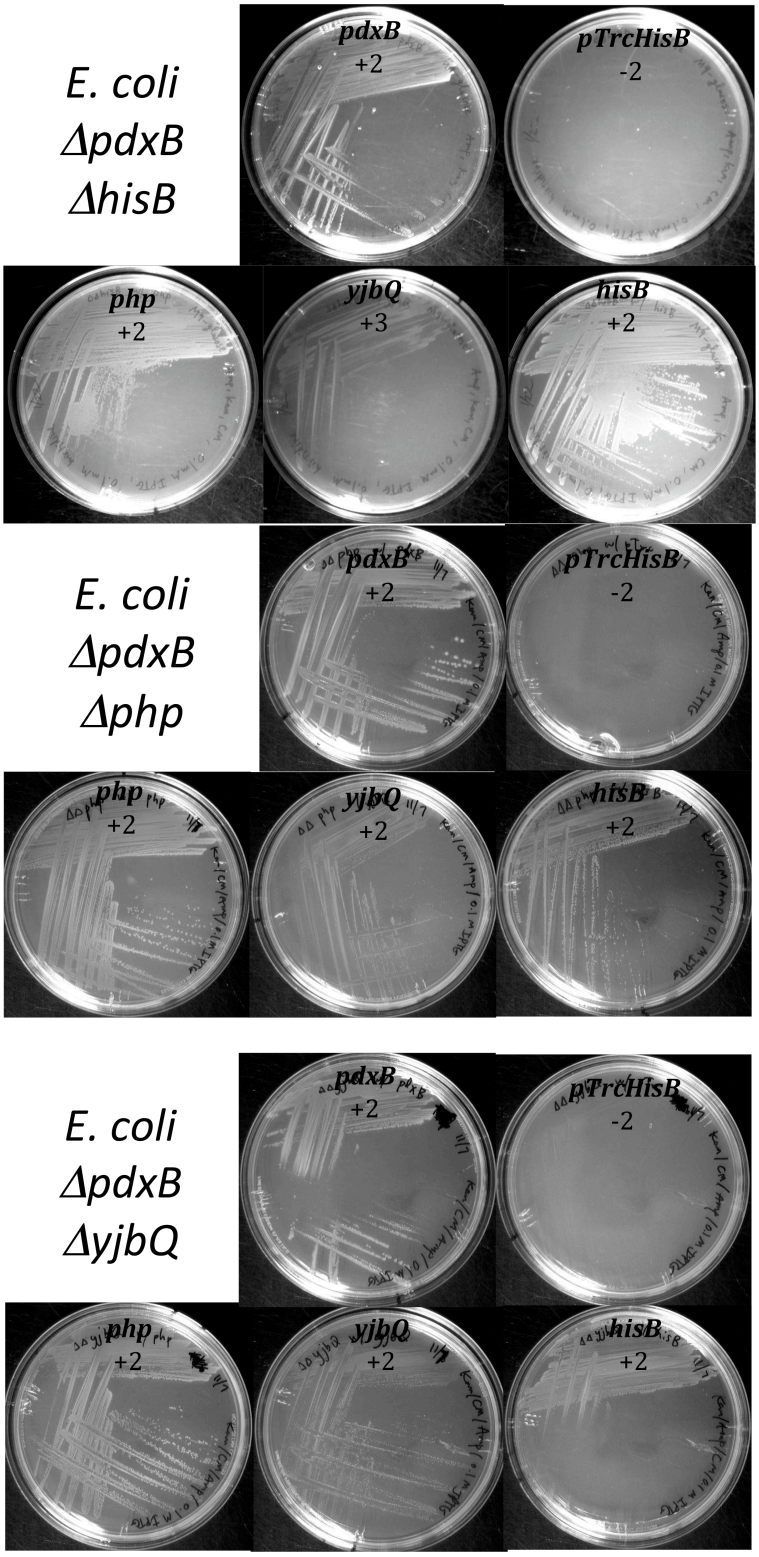


Figure 2.6. Growth of the Δ*pdxB* Δ*hisB*, Δ*pdxB* Δ*php*, and Δ*pdxB* Δ*yjbQ* strains when indicated genes are over-expressed from pTrcHisB on M9/glucose plates at 37 ºC. +N or -N indicates growth or no growth, respectively, of visible colonies after N days.

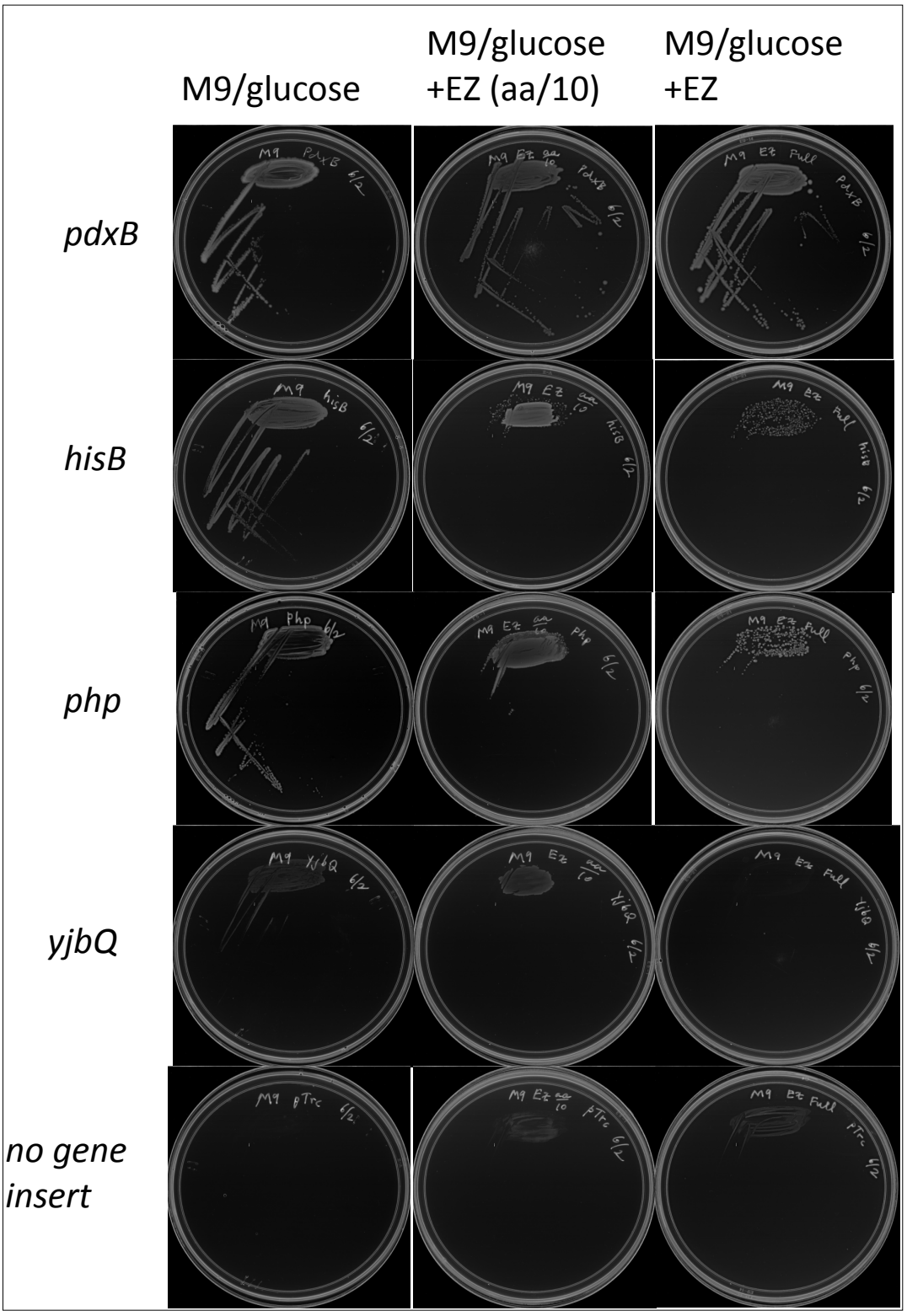


Figure 2.7. Growth of the Δ*pdxB* Δ*ltaE* strain when transformed with *pdxB*, *hisB*, *php*, *yjbQ*, or the blank pTrcHisB vector on plates of M9/glucose, M9/glucose +EZ (aa/10), and M9/glucose +EZ at 37 ºC within seven days.

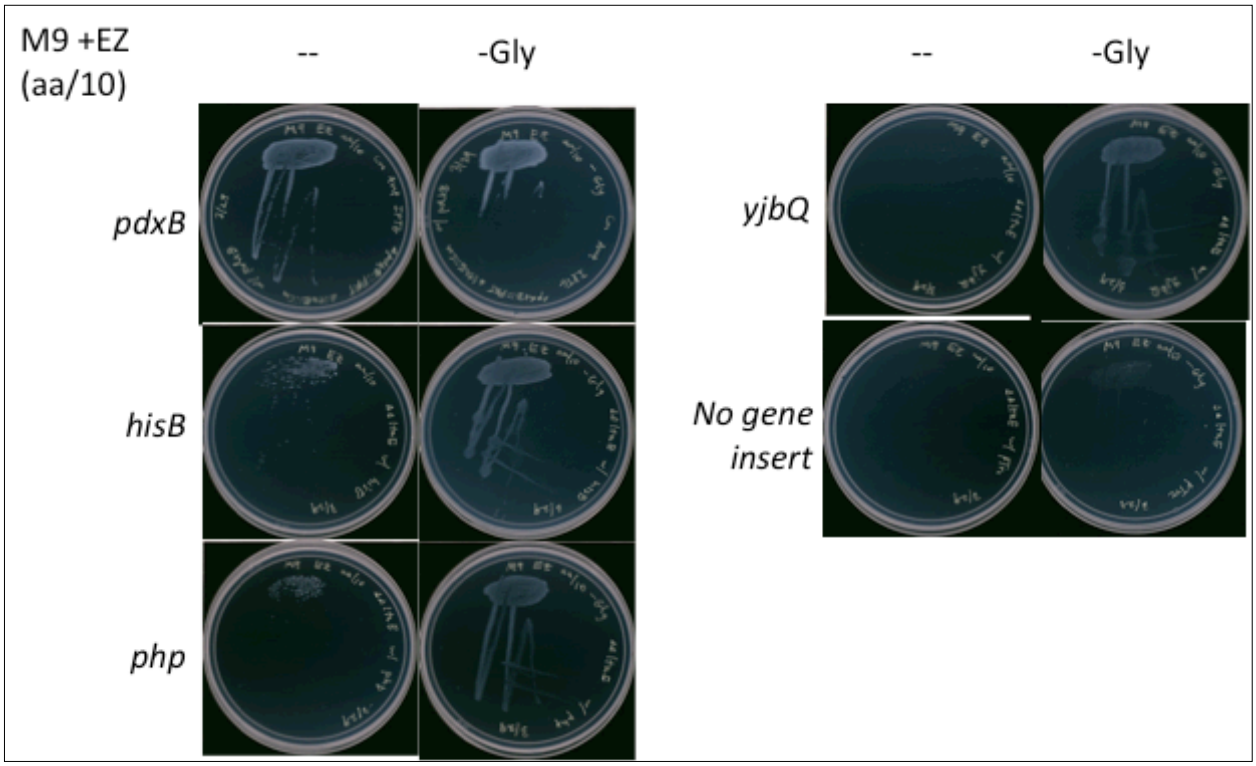


Figure 2.8. Growth at 37 ºC of the the Δ*pdxB* Δ*ltaE* strain on plates of M9/glucose +EZ (aa/10) and M9/glucose +EZ (aa/10) –Gly when this strain is transformed with the indicated gene inserts.

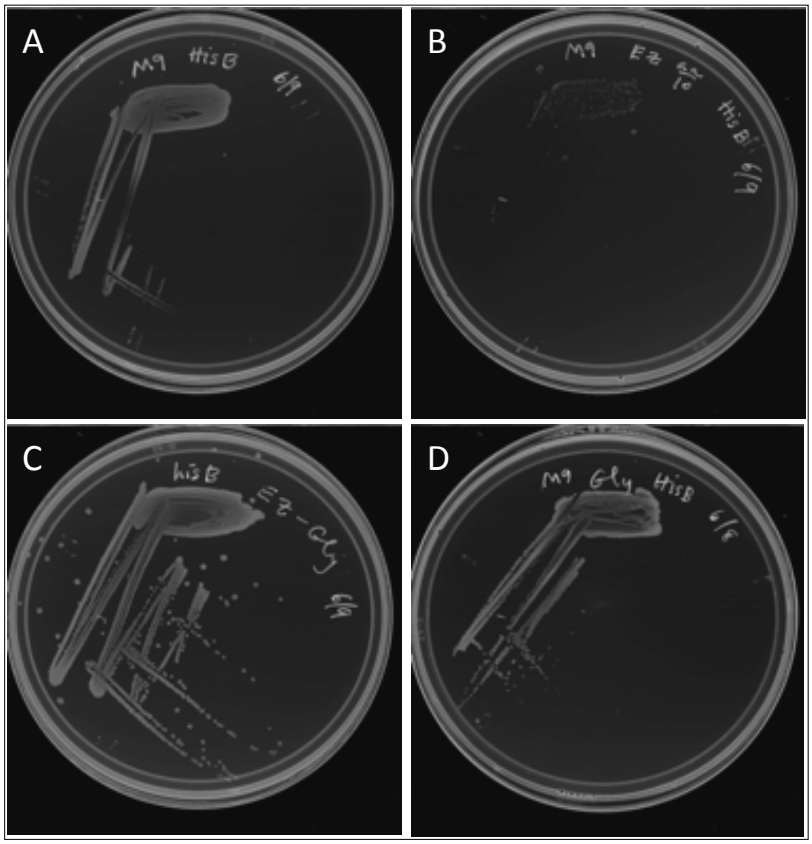


Figure 2.9. Growth of the Δ*pdxB* Δ*ltaE* strain when *hisB* is over-expressed from pTrcHisB on plates of: A) M9/glucose; B) M9/glucose +EZ (aa/10); C) M9/glucose + EZ (aa/10) - Gly; and D) M9/glucose + Gly, at 37 ºC for two days (A,B,C) or three days (D).

Methods

The ΔpdxB strain

The Δ*pdxB* strain was obtained from the Keio collection, where *pdxB* was replaced with a kanamycin resistance gene (*kan*^{r}) within *E. coli* K-12 strain BW25113 using the method diagramed in Figure 2.1 (Baba *et al*, 2006).

The ΔpdxB ΔltaE strain

Dr. Juhan Kim from the Copley Lab replaced *ltaE* with a chloramphenicol resistance gene in the Δ*pdxB* strain. The Δ*pdxB* Δ*ltaE* strain is a variant of this strain, wherein *kan*^{r} was removed from the *pdxB* locus so that this resistance gene, which is encoded in Tn5, could be used as a selectable marker for insertion mutants. I removed *kan*^{r} using the TSS method (Chung *et al*, 1989) to transform the above strain with the pCP20 vector, which encodes a recombinase that was used to

catalyze homologous recombination at the FRT sites that flank the *kan*^{r} cassette, eliminating *kan*^{r} from the genome as a small, circular fragment of DNA (Figure 2.1). Cells were grown on LB at 42 ºC to remove the temperature sensitive pCP20 vector, and the Δ*pdxB* Δ*ltaE* strain was selected on LB with chloramphenicol. This strain did not grow on LB in the presence of ampicillin, indicating that pCP20 had been successfully removed.

Knocking out hisB, php, *or* yjbQ *in the* ΔpdxB *strain*

The three knockout strains (Δ*pdxB* Δ*hisB*, Δ*pdxB* Δ*php*, and Δ*pdxB* Δ*yjbQ*) were made by replacing *php*, *hisB*, or *yjbQ* with a chloramphenicol resistance gene (*cm*^{r}) within the Δ*pdxB* strain according to the method used in Baba *et al*, 2006 and diagramed in Figure 2.1. I used a forward and reverse pair of PCR primers first to amplify *cm*^{r} from the pACYC184 vector and then to direct homologous recombination of *cm*^{r} with the target gene in the genome of the Δ*pdxB* strain (Table 2.7). The primers were each 70 bp long. In the 5' to 3' direction, the first 50 bp of the forward/reverse primers were complementary to the upstream/downstream regions of the target gene, respectively. The next 20 bp were complementary to the upstream/downstream regions of the *cm*^{r}, respectively.

Table 2.7. Sequence of PCR primers used to amplify the chloramphenicol resistance gene and direct homologous recombination with *hisB*, *php*, and *yjbQ*. Sequences in black are complementary to the target gene; sequences in red are complementary to the chloramphenicol resistance gene.

Gene		Primers (5' to 3')
hisB	Forward	GAA GAA AGC CAG CGC GTC ATT GAC GCC TTA CGT GCG GAG CAA GTT TGA TGC CTG GTG TCC CTG TTG ATA C
	Reverse	CGC AGC CGG TAT CAA GGA TCA CCA CGT TCA TTA CAG CAC TCC TTT CGA CGT TAC GCC CCG CCC TGC CAC T
php	Forward	AAT GGT CGC TCT TTT CTT TCC TCC CTC CGG TTT GCA GGA GAC ACC CTA TGC CTG GTG TCC CTG TTG ATA C
	Reverse	GCA ACG CCA ATC TTT TTC ATG AGT CTG TCC TTA TTG GAA AAA TTG AGA GGT TAC GCC CCG CCC TGC CAC T
yjbQ	Forward	CTG TCT TGC TGC AAA CTG ATT AAG AGA GTT TTA TCA AGG AGC ACG ACA TGC CTG GTG TCC CTG TTG ATA C
	Reverse	ATA TTG TAG CAA CTC CGA AAT GGT CAT TTT TTA CTC CCC TTG TAG TGT CGT TAC GCC CCG CCC TGC CAC T

In conducting PCR to amplify *cm*^{r} from the pACYC184 vector, I used approximately 200 ng template and 1 µM each primer in a final volume of 100 µL H_2O. PCR was conducted with 30 cycles of denaturation at 94 ºC for 30 seconds, primer annealing at 54 ºC for 30 seconds, and extension by the Ex Taq polymerase (TaKaRa) at 72 ºC for 1 minute. The 30 cycles were preceded by 2 minutes of denaturation at 94 ºC and followed by 2 minutes of extension at 72 ºC. The PCR product was then digested with DpnI to degrade the methylated pACYC184 template, and purified using the E.Z.N.A. Cycle Pure Kit (Omega Bio-Tek, Inc.).

Next, I used the TSS method (Chung *et al*, 1989) to transform the Δ*pdxB* strain with the pKD46 vector, which encodes lambda-red recombinase. Cells were grown at 30 ºC instead of 37 ºC since the vector is temperature sensitive. I selected for transformants on LB with ampicillin and kanamycin. Next, I transformed competent cells of this strain prepared using 10% glycerol (Sambrook *et al*, 1989) with either of the chloramphenicol resistance cassettes designed for the targeted deletion of *hisB*, *php*, or *yjbQ*. Cells were immediately recovered by adding 1 mL LB and incubated with shaking for two days at 30 ºC. Knockouts were selected on plates of LB, kanamycin, and chloramphenicol at 37 ºC. Removal of the temperature sensitive pKD46 vector from the cell was confirmed by growing cells on LB with ampicillin, kanamycin, and chloramphenicol at 30 ºC. Indeed, knockouts failed to grow in the presence of ampicillin, indicating that these cells have lost the pKD46 vector, which encodes *amp*^{r}.

Using PCR to confirm replacement of target genes with cm^r

Colony PCR with each double knockout strain was used to confirm the replacement of *hisB*, *php*, and *ybjQ* with *cm*^{r} using the strategy described in the results section, with the primers from Table 2.8. PCR was conducted with 0.2 µL of 3-4 single colonies of the double knockout strain suspended in 10 µL H_2O and 2 µM of each primer in a final volume of 10 µL, and carried out with 30 cycles of denaturation at 94 ºC for 30 seconds, primer annealing at 54 ºC for 30 seconds, and

extension by the Go Taq polymerase (Promega) at 72 ºC for 1 minute. The 30 cycles were preceded by 5 minutes of denaturation at 94 ºC and followed by 3 minutes of extension at 72 ºC.

Table 2.8. Sequence of PCR primers used to verify the replacement of *hisB*, *php*, or *yjbQ* with cm^r.

Gene		Primers (5' to 3')
hisB	i-F	CAT ATG CTG GAT CAG ATC GC
	i-R	CCG GTA TCT TCG ACG GTG TG
	o-F	AGC GGC TGC CTG CGA ATT AC
	o-R	GCG CAA TGG CAG ACT TCA CC
php	i-F	AGA TTA CGC CGC TGG AAG AG
	i-R	CGG CCT GAC TGA ATC CTG AC
	o-F	ATG ATC GCT ACG CAG CTC GG
	o-R	TAA TCT GCT CAC GCT GTA AG
yjbQ	i-F	ACT GCA TCT GTT GCT GCA AC
	i-R	CCA AAT GCC TTG CCA GGT GC
	o-F	GAT TCT GAT AAT CAT ATT AC
	o-R	TGC TGC TGA CGT AGC AGC TC
Cm^R	i-F	CGG TGA GCT GGT GAT ATG GG
	i-R	TCT TGC GAA TAT ATG TGT AG

pTrcHisB constructs

The ORFs of *pdxB*, *hisB*, *php*, and *yjbQ* from *E. coli* K-12 strain BW25113 were cloned into the pTrcHisB vector from Invitrogen. The genes were amplified using the PCR primers listed in Table 2.9 and unidirectional cloning was achieved in each case using the 5' BamHI and 3' EcoRI cut sites introduced by the PCR primers. These clones were made and given to me by Dr. Juhan Kim from the Copley Lab.

Transformation of E. coli with pTrcHisB constructs

Unless otherwise stated, pTrcHisB constructs were transformed into *E. coli* using the TSS method (Chung *et al*, 1989).

Table 2.9. Sequence of PCR primers used to amplify *pdxB*, *hisB*, *php*, and *yjbQ*, all of which introduce upstream and downstream BamHI and EcoRI cut sites, respectively.

Gene		Primers (5' to 3')
pdxB	Forward	TCA ATT GGA TCC GAA AAT CCT TGT TGA TG
	Reverse	TTA TCC GAA TTC TTA ACG TGC CGG ATG
hisB	Forward	TCA ATT GGA TCC AAG TCA GAA GTA TCT TTT TAT C
	Reverse	TTA TCC GAA TTC TTA CAG CAC TCC TTT CG
php	Forward	TCA ATT GGA TCC AAG TTT TGA TCC GAC GGG T
	Reverse	TTA TCC GAA TTC TTA TTG GAA AAA TTG AGA GGG AT
yjbQ	Forward	TCA ATT GGA TCC ATG GTA TCA AAA GAC GCT CAC
	Reverse	TTA TCC GAA TTC TTA CTC CCC TTG TAG TGT C

Complementation on M9/glucose or complex media

Complementation of strains lacking *pdxB* was conducted in the presence of 0.1 mM IPTG, which induces transcription of the recombinant gene by relieving *lacI* inhibition at the pTrcHisB promoter. Cells were washed before plating to remove extra-cellular pyridoxine, which could enable background growth. Single colonies of transformants of the Δ*pdxB* strain were selected on LB with kanamycin and ampicillin. Next, 1-3 single colonies were washed 5X in 50 µL filter-sterilized phosphate buffered serine (PBS), centrifuging between washes at 13,400-16,000 g for 1 minute. The final cell pellet was resuspended in 10 µL PBS and 1-3 µL of the resuspension was streaked on plates of M9/glucose or complex media with IPTG (0.1 mM) and grown at 37 ºC.

Complementation on M9/glucose in the presence of histidine

In complementation experiments on M9/glucose, when specified, histidine was included in the growth medium at a concentration of 0.1 mM.

Antibiotics

When appropriate, kanamycin was included at 50 µg/mL and was used to select cells with kan^r at the *pdxB* locus; ampicillin was included at 100 µg/mL and was used to select cells with amp^r from pTrcHisB vector; and chloramphenicol was included at 35 µg/mL and was used to select for cells with cm^r at the *hisB*, *php*, *yjbQ*, or *ltaE* loci.

Complex media of M9/glucose supplemented with nutrients from EZ Rich Defined Medium

Below are descriptions for the following complex media referred to in this thesis. The protocol for making these media is shown at the end of this chapter. The ingredients in the following media are listed below:

a) M9/glucose +EZ: Consisted of M9/glucose supplemented with the amino acids, nucleotide bases, and vitamins from EZ Rich Defined Medium (Neidhardt *et al*, 1974).

b) M9/glucose +EZ (aa/10): Is the same as 'a,' but with a 10-fold reduction in the concentration of amino acids.

c) M9/glucose +EZ (aa/10) -Gly: Is the same as 'b,' but minus glycine. This was the medium used to screen the insertion library.

Protocol for making plates of M9/glucose with amino acid, nucleotide base, and vitamin supplements from EZ Rich Defined Medium

M9/glucose EZ Rich Plates (amino acids diluted 10X) (1L – 40 plates)

In 0.36 L di-H_2O add and autoclave M9 salts (amnt. for 1 L M9):
- -6g Na_2HPO_4
- -3g KH_2PO_4
- -1g NH_4Cl

In 0.5 L di-H_2O add and autoclave bactoagar:
- -15 g bactoagar

Next, combine above batches and add (filtered in di-H_2O):
- -1 mL 1 M $MgSO_4$
- -10 mL 20% glucose
- -0.1 ml 1 M $CaCl_2$

Add or dropout following EZ Rich supplements:
- -100 mL 10X AGCU
- -20 mL 5X amino acids (dilutes a.a. 10X from EZ Rich protocol)
- -10 mL VA vitamins

Depending on which nutrients were used, add sterilized di-H_2O to 1 L

M9/glucose EZ Rich Plates (full amino acid supplement) (1L – 40 plates)

Follow the above protocol but autoclave M9 salts in 0.18 L di-H_2O and add 200 mL 5X amino acid supplement.

Chapter III: The imidazoleglycerol-phosphate dehydratase domain of HisB has a promiscuous activity that complements the Δ*pdxB* strain during the over-production of HisB

Introduction

HisB is a bi-functional enzyme that catalyzes the 6th and 8th steps of histidine biosynthesis in *E. coli* (Figure 3.1). These functions are catalyzed by two structurally distinct domains, where the N-terminal domain is a histidinol-phosphate phosphatase (HPP) and the C-terminal domain an imidazoleglycerol-phosphate dehydratase (IGPD). The bi-functionality of HisB raises the question of which domain is responsible for complementation of the Δ*pdxB* strain.

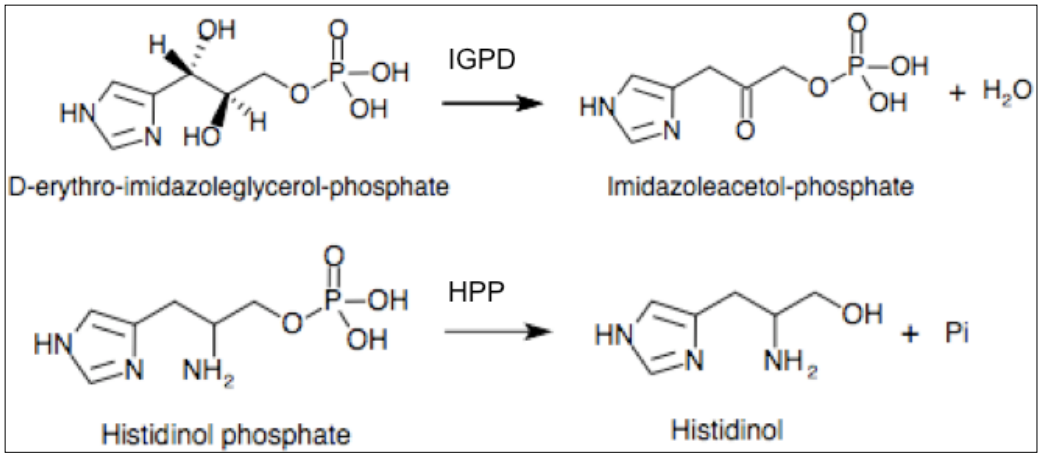


Figure 3.1. *E. coli* HisB has two structurally distinct catalytic domains. The C-terminal domain, imidazoleglycerol-phosphate dehydratase (IGPD), and the N-terminal domain, histidinol-phosphate phosphatase (HPP), catalyze the 6th and 8th steps in histidine biosynthesis, respectively.

To address the above question, I tested the ability of mono-functional, mutant versions of HisB to complement the Δ*pdxB* strain on M9/glucose. First, I sub-cloned the two parts of *hisB* that encode the enzyme's N- and C-terminal domains and tested if over-production of each domain rescued growth of the Δ*pdxB* strain on M9/glucose. I anticipated that over-production of one of the two domains would recapitulate the phenotype seen by over-producing wild-type HisB, indicating that this domain catalyzes a promiscuous activity involved in complementation. Indeed, only over-production of the IGPD domain complemented the Δ*pdxB* strain.

To confirm the importance of IGPD in complementing the Δ*pdxB* strain, I also tested whether mutating single IGPD active site residues interfered with complementation by over-

production of an otherwise wild-type HisB protein. This experiment offers greater precision than the first for two reasons. The first reason is that sub-cloning the HPP domain may have caused this domain to mis-fold and lose its ability to complement the Δ*pdxB* strain. In comparison, making single amino acid changes in the IGPD domain of HisB is more likely to leave the HPP domain functionally intact. Thus, if these mutations do disable complementation, it can be said with greater confidence that HPP does not have a promiscuous activity that is capable of complementing the Δ*pdxB* strain. Another source of precision comes from being able to see the effect of single amino acid changes on IPGD's promiscuous activity. The first experiment indicated that the IGPD domain catalyzes a promiscuous activity that is relevant in complementation. However, knowing the consequence of specific amino acid changes on the promiscuous activity can be used to more precisely pin the activity to or away from the IGPD active site and may even offer insight into the mechanism of the promiscuous reaction, based on the essentiality of different residues.

I constructed two mutant alleles of *hisB* that encode proteins with single amino acid changes of either E177A or E327A. I targeted these residues based on a study of HisB from *A. thaliana*, which predicts that residues that are homologous with *E. coli* residues E177 and E327 (i.e. *A. thaliana* residues E21 and E173, respectively) mediate catalysis in the *A. thaliana* ortholog (Figure 3.2) (Glynn *et al*, 2005). This prediction was made using molecular modeling in context of the crystal structure of *A. thaliana* HisB, where E21 and E173 are proposed to mediate acid-base catalysis, the former in eliminating the hydroxyl group and abstracting the proton from C2 to convert between the diazafulvene intermediate and the enolic intermediate, and the latter in the enol to keto tautomerization that produces imidazoleacetol-phosphate. E173 is also predicted to assist in the coordination one of two Mn^{2+} ions involved in catalysis. This study strengthens their prediction by showing the evolutionary conservation of E21 and E173 across various eukaryotic and bacterial species, including *E. coli*. I tested whether the over-expression of *E. coli hisB-E177A*

and *hisB-E327A* complemented not only the Δ*pdxB* strain, but also a strain of *E. coli* lacking *hisB*, to gauge the importance of these residues in IGPD's 'normal' activity as well.

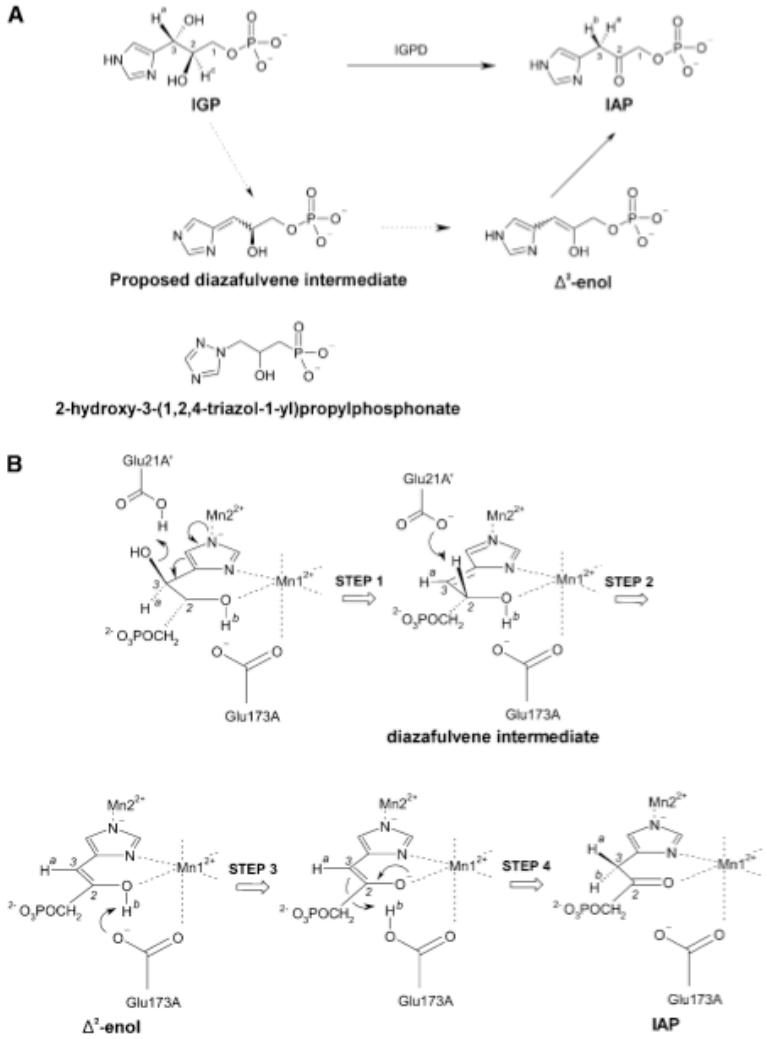


Figure 3.2. From Glynn *et al*, 2005, showing the proposed mechanism of IGPD from *A. thaliana*, where it is put forth that: A) the mechanism proceeds through a diazafulvene intermediate and B) that two glutamate residues mediate acid-base catalysis and two Mn^{2+} ions help stabilize intermediates.

Results

Over-production of the IGPD domain, but not the HPP domain, complements the ΔpdxB strain on M9/glucose

The N- and C- terminal domains of HisB encode HPP and IGPD, respectively. Dr. Juhan Kim sub-cloned the sequences of *hisB* that encode HPP or IGPD using the pTrcHisB vector and I checked

that these constructs contained the expected gene inserts by fully sequencing the gene inserts, including each constructs' multiple cloning site, immediately upstream and downstream of the restriction sites that were used to ligate the gene inserts with the pTrcHisB vector. The sequencing data matched the expected sequences.

Next, I transformed these constructs into the Δ*pdxB* strain, and tested the ability of each strain to grow on M9/glucose during the over-production of the respective recombinant domain (Table 3.1; Figure 3.3). Over-production of the HPP domain did not complement the Δ*pdxB* strain, as visible colonies failed to appear within seven days. However, complementation was seen in cells that over-produced the IGPD domain, which formed visible colonies after two days of growth. As positive and negative controls, I grew the Δ*pdxB* strain on M9/glucose while over-producing PdxB, or when transformed with the blank pTrcHisB vector, respectively. Over-production of PdxB produced visible cell growth after two days, while cells that were transformed with pTrcHisB failed to produce visible colonies after seven days.

Table 3.1. Growth of the Δ*pdxB* strain on plates containing M9/glucose at 37 ºC when the indicated genes were over-expressed from pTrcHisB.

Gene	Function of HisB Domain	Δ*pdxB*
pdxB	n.a.	+2[a]
hisB N-terminal domain	HPP	-7
hisB C-terminal domain	IGPD	+2
no gene insert	n.a.	-7

[a]+N and -N indicates growth or no growth, respectively, of visible colonies after N days.

Side directed mutagenesis of the hisB construct to generate the mutant constructs hisB-A530C and *hisB-A980C*

To confirm the importance of the IGPD domain in complementing the Δ*pdxB* strain during HisB over-production, I tested how mutating single residues from the IGPD active site effected this process. First, I used site-directed mutagenesis to create two mutant alleles of *hisB*, each with a single point mutation: *hisB-A530C* and *hisB-A980C*. These mutations translate to changes in HisB's primary sequence of E177A and E327A, respectively. pTrcHisB constructs with the mutant *hisB* alleles were generated by amplifying the *hisB* construct with primers that contain the desired SNP,

as diagramed in Figure 3.4. PCR products generated in this fashion consisted of a single band of approximately 5 kb when visualized after electrophoresis on a 1% gel (Figure 3.5). The amplification product was then digested with DpnI to degrade the wild-type template, purified, transformed into *E. coli* DH5α, and isolated in greater quantity from ampicillin-resistant single-colonies. Constructs digested with EcoRI and BamHI yielded two bands corresponding in expected sizes to the *hisB* gene inserts and the pTrcHisB backbone.

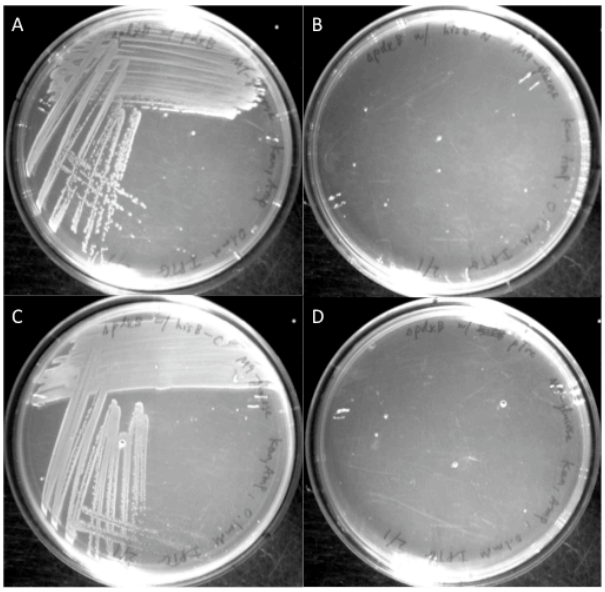


Figure 3.3. Growth of the Δ*pdxB* strain on plates containing M9/glucose at 37 ºC when *pdxB*, *hisB-N*, or *hisB-C* were over-expressed from pTrcHisB (A-C, respectively) or when the Δ*pdxB* strain was transformed with the pTrcHisB vector (D). Plates incubated for the number of days indicated in Table 3.1.

Next, I took steps to protect against the possibility that the wild-type *hisB* construct escaped degradation by DpnI and co-transformed with the mutant constructs into *E. coli* DH5α. Along these lines, I digested the *hisB-A530C* and *hisB-A980C* constructs isolated in the above step, separated the products on a 1% agarose gel, and excised and purified the DNA band corresponding to the expected size of the mutant *hisB* inserts. I then ligated this DNA a new pTrcHisB backbone and transformed the ligation product into *E. coli* DH5α, which produced no more than 30 single colonies with ampicillin resistance, indicating a low transformation efficiency and a low chance of co-transformation.

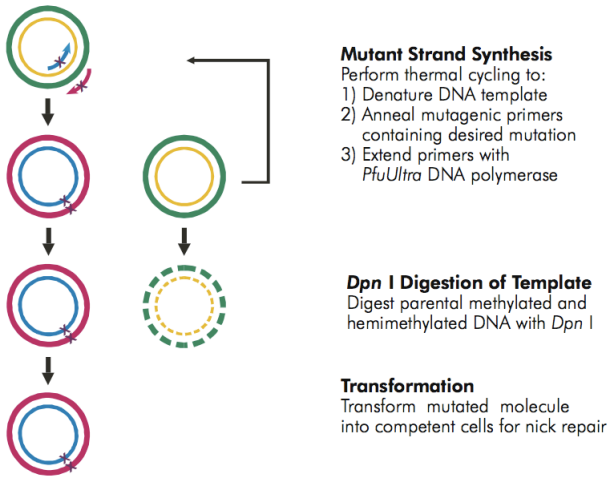


Figure 3.4. From Stratagene, describing the use of PCR to introduce site-directed point mutations into constructs of circular DNA.

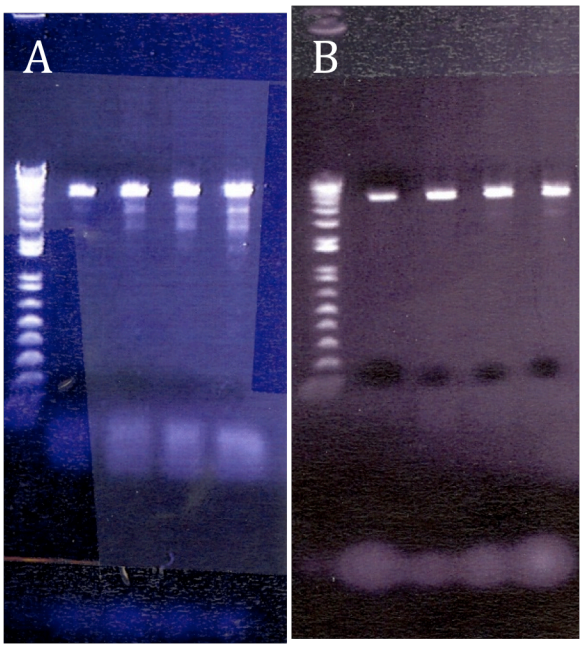


Figure 3.5. DNA generated by PCR with the *hisB* construct amplified using primers to create the mutations A) A530C or B) A980C, with increasing amounts of template, visualized after electrophoresis on a 1% agarose gel alongside a 1 kb plus ladder (Invitrogen).

I isolated the *hisB-A530C* and *hisB-A980C* constructs from *E. coli* DH5α transformants and confirmed that these constructs contained the anticipated inserts by fully sequencing these inserts, along with the regions of the multiple cloning site immediately upstream and downstream of the restriction cut sites I used to ligate the inserts with pTrcHisB. In each case, the sequence data matched the anticipated sequence. I also checked the sizes of the *hisB-A530C* and *hisB-A980C* gene

inserts by digesting these constructs with BamHI and EcoRI and visualizing the product after electrophoresis on a 1% agarose gel. Each construct generated two bands of DNA that corresponded to the expected lengths of the pTrcHisB backbone (ca. 4.4 kb) and the mutant *hisB* inserts (ca. 1 kb) (Figure 3.6).

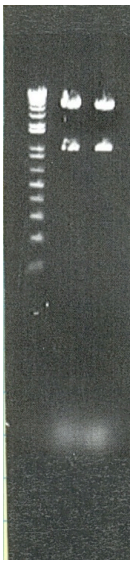


Figure 3.6. The re-purified *hisB-A530C* and *hisB-A980C* constructs digested with BamHI and EcoRI and visualized after electrophoresis on a 1% agarose gel alongside a 1 kb plus ladder (Invitrogen).

The mutations E177A and E327A either partially or completely disabled the over-production of HisB from complementing the ΔpdxB strain on M9/glucose

The *hisB-A530C* and *hisB-A980C* constructs, which have changes in HisB's primary sequence of E177A and E327A, respectively, were transformed into the Δ*pdxB* strain, and it was seen if over-production of HisB E177A and HisB E327A complemented the Δ*pdxB* strain on M9/glucose (Table 3.2; Figure 3.7). Compared with HisB over-production, which results in growth of visible colonies of the Δ*pdxB* strain after two days (determined from previous experiment; not in parallel), the over-production of HisB-E327A produced visible colonies after five days, while cells over-producing HisB-E177A failed to form visible colonies after seven days. As positive and negative controls, I tested for growth of the Δ*pdxB* strain on M9/glucose when transformed with *pdxB* or the blank pTrcHisB vector, respectively. Over-expressing *pdxB* resulted in visible cell growth after two days,

while the strain transformed with the blank pTrcHisB vector failed to produce visible colonies after seven days.

The mutations E177A and E327A either completely disabled the over-production of HisB from complementing the ΔhisB strain on M9/glucose

To gauge the importance of E177 and E327 to IGPD's 'normal' activity, I also tested whether over-production of either mutant protein could complement a strain of *E. coli* lacking *hisB* (Table 3.2). Neither complementation experiment was successful, suggesting that IGPD activity was significantly diminished in these alleles in each case. As a positive control, I tested if over-expression of *hisB* from the pTrcHisB vector complements the *hisB* deletion. Also, as a negative control, I tested for growth of the Δ*hisB* strain on M9/glucose when transformed with the blank pTrcHisB vector. Cells from the positive control produced visible colonies after two days, while those from the negative control failed to grow within seven days.

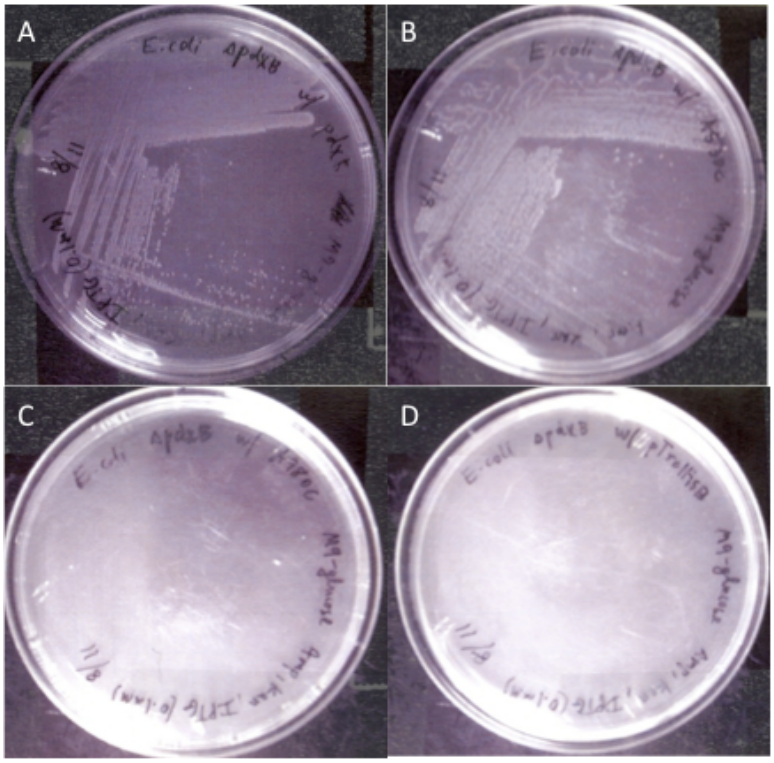


Figure 3.7. Growth of the Δ*pdxB* strain on plates containing M9/glucose at 37 ºC when *pdxB*, *hisB-E177A*, or *hisB-E327A* were over-expressed from pTrcHisB (A-C, respectively) or when the Δ*pdxB* strain was transformed with the pTrcHisB vector (D), after 2 days (A) or 7 days (B-D).

Table 3.2. Growth of the Δ*pdxB* strain or the Δ*hisB* strain on plates containing M9/glucose at 37 ºC when the indicated genes were over-expressed from pTrcHisB.

Gene	Δ*pdxB*	Δ*hisB*
pdxB	+2[a]	n.a.
hisB	+2[b]	+2
hisB E177A	+5	-7
hisB E327A	-7	-7
no gene insert	-7	-7

[a]+N and -N indicates growth or no growth, respectively, of visible colonies after N days.
[b]Determined previously.

Discussion

Both experiments in this chapter support the conclusion that IGPD catalyzes a promiscuous activity that contributes to a serendipitous pathway that complements the Δ*pdxB* strain upon the over-production of HisB. When the two domains of HisB were sub-cloned and individually over-produced in the Δ*pdxB* strain, only the over-production of IGPD resulted in complementation, indicating that the elevation of a promiscuous activity catalyzed by the IGPD domain drives complementation of the Δ*pdxB* strain during the over-production of HisB. This indication was supported by the ability of mutations at the IPGD active site to completely or partially disrupt complementation by HisB over-production.

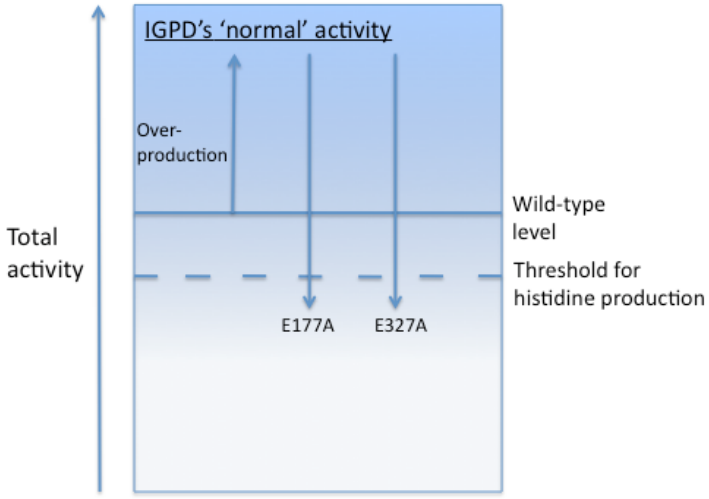
Conversely, these results indicate that the HPP domain does not catalyze a promiscuous activity that can be elevated to complement the Δ*pdxB* strain. Although over-production of only the HPP domain fails to complement the Δ*pdxB* strain, the most compelling result that supports this claim is the reduced ability or complete inability of over-production of HisB-E177A and HisB-E327A to do so. Whereas the HPP domain could mis-fold when it is expressed without the IGPD domain and lose a relevant promiscuous function, there is a greater probability of the HPP domain remaining structurally and functionally intact upon the introduction of single amino acid changes in the active site of a different domain.

It is interesting that on M9/glucose, over-production of neither HisB-E177A nor HisB-E327A complemented a deletion in *hisB*, but that over-expression of *hisB-E177A* was able to complement a deletion in *pdxB*. While mutating both E177 and E327 lowered IGPD's 'normal' activity below the physiologically relevant threshold need for histidine biosynthesis, only the act of mutating E327 lowered IGPD's promiscuous activity below the threshold needed for PLP synthesis.

These results also suggest that, qualitatively, deleting E177 must have caused a lesser relative decrease in IPGD's promiscuous activity than IPGD's 'normal' activity (Figure 3.8). These activities operate at some level under normal conditions. When HisB is over-expressed, both the 'normal' and promiscuous activities are expected to increase by the same factor. However, when HisB-E177A is over-produced, since only its promiscuous activity remains physiologically relevant, this mutation must decrease the 'normal' activity by a greater factor than it does the promiscuous activity. Considering, also, that when IGPD is over-produced its promiscuous activity is closer to a physiologically irrelevant level to begin with, this difference in effects is probably significant.

Theses results also suggest that E327 is more important than E177 in catalyzing the promiscuous reaction involved in complementing the Δ*pdxB* strain, bringing up the point that a promiscuous reaction need not follow the same mechanism, nor use the same catalytic residues, as the native reaction.

i)



ii)

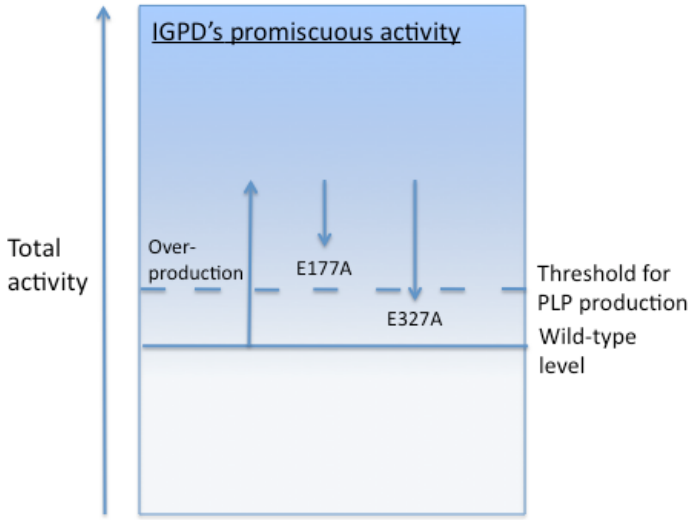


Figure 3.8. The qualitative effects of point mutations E177A and E327A on IGPD's i) 'normal' and ii) promiscuous activities, where arrows that pass below thresholds for physiological relevance show the minimum decrease in activity in each case.

Methods

Strains

The Δ*pdxB* strain and the Δ*hisB* strain were obtained from the Keio collection, where either gene was replaced with a kanamycin resistance gene (*kan*^{r}) within *E. coli* K-12 strain BW25113 using the method diagramed in Figure 2.1 (Baba *et al*, 2006).

Cloning of n- and c-terminal HisB domains

The sequences enconding the N- and C- terminal domains of *hisB* from *E. coli* K-12 strain BW25113 were amplified using the the PCR primers in Table 4.3 cloned into the pTrcHisB vector using 5' BamHI and 3' EcoRI cut sites introduced by the PCR primers. These constructs were given to me by Dr. Juhan Kim from the Copley Lab. I confirmed that these constructs contained the correct gene inserts by sequencing the entire inserts using the primers from Table 4.4.

Table 4.3. PCR primers used to amplify *pdxB*, *hisB*, *php*, *yjbQ* and *aroB* from Eco genomic DNA. In each case, BamHI and EcoRI were used as 5' and 3' cutters for uni-directional cloning.

Gene		Primers (5' to 3')
hisB-n	Forward	TCG ACC GGA TCC GAG TCA GAA GTA TCT TTT TAT CG
	Reverse	TTA CAA GAA TTC TTA ACG TCT GGT GAG TTG CTC GC
hisB-c	Forward	TCG ACC GGA TCC GGA CCG TTA CGC TCA CGT AGT G
	Reverse	TTA CAA GAA TTC TTA CAG CAC TCC TTT CGA CG

Table 4.4. Sequencing primers used in the sequencing of *hisB* constructs.

Gene	Primers (5' to 3')
pTrcHisB-F	GAG GTA TAT ATT AAT GTA TCG
hisB seq 2	TGC CGT AAG CCG AAA GTA AAA CTG G
hisB seq 3	TCA AAG GCG ACC TCT ATA TCG ACG ATC

Site-directed mutagenesis of the hisB construct to generate the hisB-A530C and *hisB-A980C constructs*

Site-directed mutagenesis of the pTrcHisB vector cloned with the *hisB* gene insert was used to generate constructs with the point mutations: *A530C* and *A980C*, according to the QuickChange II Site-Directed Mutagenesis Kit (Stratagene). PCR of the *hisB* construct was conducted in a total volume of 50 µL, including 5 µL of the 10X reaction buffer from the QuickChange II Site-Directed Mutagenesis Kit; 5, 10, 20, or 50 ng of template DNA; and 10 uM of the forward and reverse PCR primers from Table 4.5 that have the desired changes in sequence. The PCR product was visualized after electrophoresis on a 1% agarose gel containing ethidium bromide.

Table 4.5. Primers used in site-directed mutagenesis of *hisB* where changes from the complementary sequence of wild-type *hisB* are denoted in red

Gene		Primers (5' to 3')
hisB-E177A	Forward	GCG TAA TAC CAA AGC GAC GCA GAT TGA CG
	Reverse	CGT CAA TCT GCG TCG CTT TGG TAT TAC GC
hisB-E327A	Forward	CAT CAC CGT GTA GCG AGT CTG TTC AAA GC
	Reverse	GCT TTG AAC AGA CTC GCT ACA CGG TGA TG

The PCR product was treated with DpnI to degrade the wild-type *hisB* template and then electroporated into competent cells of *E. coli* DH5α using prepared using calcium chloride (Sambrook *et al*, 1989). The *hisB-A530C* and *hisB-A980C* constructs were isolated from single colonies selected on LB with ampicillin using the E.Z.N.A Mini-Prep Kit (Omega Bio-Tek).

To further protect against the possibility that the wild-type *hisB* construct co-transformed with the *hisB* mutant constructs, *hisB* mutant gene inserts were separated from the pTrcHisB backbone by restriction digestion with BamHI and EcoRI, gel purified using the E.Z.N.A. Cycle Pure Kit (Omega Bio-Tek), and re-ligated with a 'fresh' pTrcHisB vector using T4 DNA Ligase (New England BioLabs). Then, the ligation product was electroporated into competent cells of *E. coli* DH5α prepared using calcium chloride (Sambrook *et al*, 1989). The *hisB-A530C* and *hisB-A980C* constructs were then again isolated from single colonies selected on LB with ampicillin using the E.Z.N.A. Mini-Prep Kit (Omega Bio-Tek). I confirmed that these constructs had the correct inserts by fully sequencing the gene inserts using the primers in Table 4.4, as well as digesting these constructs with BamHI and EcoRI and visualizing the product after electrophoresis on a 1% agarose gel with ethidium bromide.

pTrcHisB constructs and transformation into E. coli

The pTrcHisB constructs containing *pdxB*, *hisB*, or no gene insert, and the method used to transform these constructs into *E. coli*, is the same as described in the methods section of *chapter II*.

Complementation of the double knockout strains on M9/glucose

Complementation of *E. coli* knockouts on M9/glucose follows the same protocol as in the methods section of *chapter II*.

Antibiotics

When appropriate, kanamycin was included at 50 μg/mL and to select for cells with kan^r at the *pdxB* locus; ampicillin was included at 100 μg/mL and was used to select for cells with amp^r from pTrcHisB vector.

Chapter IV: Using transposon mutagenesis to uncover additional enzymes in the serendipitous pathway used to complement a deletion in *pdxB* by HisB over-production

Introduction

In chapter II, double knockout complementation experiments were used to test for epistasis between *hisB*, *php*, and *yjbQ*. In comparison, genome-wide transposon mutagenesis offers a high-throughput way to test for epistasis between either of these enzymes and the entire *E. coli* proteome. In this chapter, I used this method to search for additional enzymes involved in the serendipitous pathway that complements the Δ*pdxB* strain upon *hisB* over-expression.

I created an insertion library generated from a strain of *E. coli* with deletions in both *pdxB* and *ltaE* (hereafter refered to as the Δ*pdxB* Δ*ltaE* strain), and which over-expressed *hisB*. I used a strain lacking *ltaE* to prevent pseudo-reversion via pathway #1, which occasionally occurs on M9/glucose (Kim *et al*, 2010). In the screen of the insertion library for PLP auxotrophs in this chapter, growth was sometimes severely reduced. In these instances it was useful to know that no amount of growth was due to pseudo-reversion via pathway #1.

The addition of metabolic supplements to M9/glucose was used to more effectively screen the insertion library for PLP auxotrophs

A simple way to screen the insertion library for cells that are unable to synthesize PLP would be to patch cells onto M9/glucose +/- pyridoxine and search for cells that only grow in the presence of pyridoxine. However, the screen is improved if M9/glucose is supplemented with metabolites that *E. coli* must synthesize on minimal medium. If insertion inactivates an enzyme in the serendipitous pathway that is also required for growth on M9/glucose in its 'normal' role, the addition of pyridoxine alone will be unable rescue growth, and the strain's auxotrophy for PLP would be undetectable. In comparison, a larger number of PLP auxotrophs will theoretically be detectable on a complex medium since fewer enzymes would be needed for growth in their 'normal' roles.

It was important to find a growth medium that neither directly complemented the Δ*pdxB* Δ*ltaE* strain nor inhibited *hisB* complementation. In chapter II, I found that although the total amino acid, nucleotide base, and vitamin supplements from EZ Rich Defined Medium (Neidhardt, 1974) inhibited *hisB* complementation, inhibition was relieved with the omission of glycine. Thus, I chose to use the M9 +EZ (aa/10) –Gly medium for the screening in this chapter. The amino acids were diluted ten-fold because the amounts in the full-strength medium seemed excessive. Significantly, I also showed that this medium does not promote growth of the Δ*pdxB* Δ*ltaE* strain in the absence of an exogenous source of PLP.

This complex medium includes various essential metabolites that *E. coli* must synthesize on minimal medium: the 20 L-amino acids minus glycine; the nucleotide bases: adenine, guanine, cytosine, and uracil; and the vitamins: thiamine, pantothenate, *p*-aminobenzoic acid, *p*-hydroxybenzoic acid, and 2,3-dihydroxy benzoic acid. Thus, a broad range of enzymes that are essential on M9/glucose can be deleted on this complex medium without disrupting cell growth from a lack of their 'normal' functions, resulting in tremendous improvement in the capacity of this screen.

The potential for transposon mutagenesis to identify additional enzymes involved in complementation of the pdxB deletion by hisB over-expression

Indeed, it is possible that other enzymes work alongside HisB, Php, and YjbQ in complementing the Δ*pdxB* strain. Even though no others were revealed in the initial multi-copy suppression experiment, this experiment did not include every *E. coli* ORF. In addition, the multi-copy suppression experiment is not suited to detect promiscuous enzymes that participate in a relevant serendipitous pathway, but do not elevate the flux of this pathway when over-produced, disallowing these enzymes to complement the Δ*pdxB* strain. This scenario is possible if the enzyme does not catalyze a rate limiting step and so has little control over pathway flux. In comparison, any enzyme that is essential to a serendipitous pathway would be detectable in a screen for deletions

that disrupt the function of the pathway. In this way, transposon mutagenesis has the potential to identify additional enzymes in pathway #2.

There are multiple ways that transposition could disrupt *hisB* complementation. One way is if the transposon inserts in a gene, or the regulatory region of a gene, that encodes an enzyme in the serendipitous or native part of PLP synthesis. The number of enzymes in the serendipitous leg is unknown; however, there are five enzymes downstream of OHPB in native PLP synthesis that are expected to be epistatic with *hisB*: SerC, PdxA, Dxs, PdxJ, and PdxH. In addition, since serendipitous pathways divert material at some branch-point in core metabolism, transposition of a gene involved in the production of the branch-point intermediate could disrupt *hisB* complementation. In this way, transposon mutagenesis could be useful in determining the origin of the serendipitous pathway. For example, inactivation of *serA* in serine biosynthesis would inactivate pathway #1. In fact, another reason why adding metabolic supplements to the growth medium enhances the potential for this experiment is that less enzymes in metabolic pathways upstream of potential branch-points are essential for growth and would be identifiable in this screen.

There are also numerous ways transposon mutagenesis could indirectly disrupt *hisB* complementation—any alteration in cell physiology that results in a decreased flux through the pathway. To give one example, the inactivation of a general regulatory protein could down-regulate an entire suite of genes, including one of the promiscuous enzymes involved in the serendipitous pathway. If regulatory proteins are identified in this experiment, a systematic targeting of genes that this protein is known to regulate could subsequently be used to identify enzymes that participate in the serendipitous synthesis of PLP.

Results

*Patching the Δ*pdxB* Δ*ltaE* strain from LB to a complex medium that lacks PLP with out washing the cells in PBS does not allow growth from the presence of residual sources of PLP*

In growth experiments that test for complementation of the Δ*pdxB* strain or the Δ*pdxB* Δ*ltaE* strain, cells are normally washed in multiple times in PBS to remove residual sources of PLP

that could promote background growth; however, this washing step is impractical in the screening of hundreds of cells. I intended to screen the insertion library for PLP auxotrophs on the M9 +EZ (aa/10) –Gly complex medium. So, I tested the effect of directly patching the Δ*pdxB* Δ*ltaE* strain, transformed with either *pdxB*, *hisB*, or the blank pTrcHisB vector, from LB onto either: M9/glucose, M9/glucose +EZ (aa/10), and M9/glucose +EZ (aa/10) -Gly (Figure 5.1). After two days of growth on M9/glucose and M9/glucose +EZ (aa/10) -Gly, cells over-expressing either *pdxB* or *hisB* grew to a noticeably larger size than cells transformed with pTrcHisB, indicating that direct patching does not blur the phenotypes of cells that are able to synthesize PLP by *pdxB* or *hisB* over-expression vs. cells that cannot synthesize PLP and rely on residual PLP for growth. In addition, cells over-expressing *pdxB* grew to a noticeably larger size than the other two strains on M9/glucose +EZ (aa/10), which is consistent with the fact that the presence of glycine in the medium inhibits PLP synthesis by *hisB* over-expression.

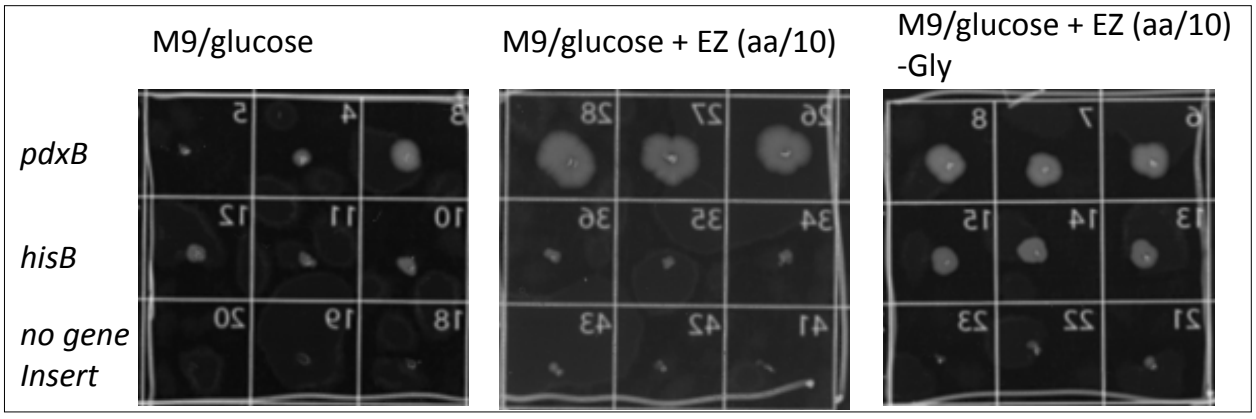


Figure 5.1. Growth after two days at 37 ºC of the Δ*pdxB* Δ*ltaE* strain when transformed with *pdxB*, *hisB*, or the blank pTrcHisB vector and patched directly from LB onto the indicated media.

A screen of 750 insertion mutants of the ΔpdxB ΔltaE strain over-expressing hisB uncovered several strains that were either partially or completely auxotrophic for PLP

I conducted genome-wide transposon mutagenesis with the Δ*pdxB* Δ*ltaE* strain transformed with *hisB* on the pTrcHisB vector, *in vivo* using the EZ-Tn5 Transposome (Epicenter) (Figure 5.1). The transposome is a stable complex of the Tn-5 transposon bonded with a transposase, which is

activated by intracellular Mg^{2+} concentrations, resulting in random, catalyzed transposition. An insertion library of 2,500 single colonies was isolated on plates of LB that included kanamycin to select for cells that had undergone transposition, which can express *kan*^{r} from the Tn5 transposon.

Next, I screened 750 insertion mutants for PLP auxotrophy by patching single colonies onto gridded plates with the M9/glucose +EZ (aa/10) -Gly complex medium +/- pyridoxine (Figure 5.2). Cells were also patched onto a third plate of LB as a positive control. This screen searched for mutants where pyridoxine rescued growth on complex medium, indicating that transposon mutagenesis had inactivated the pathway for PLP synthesis.

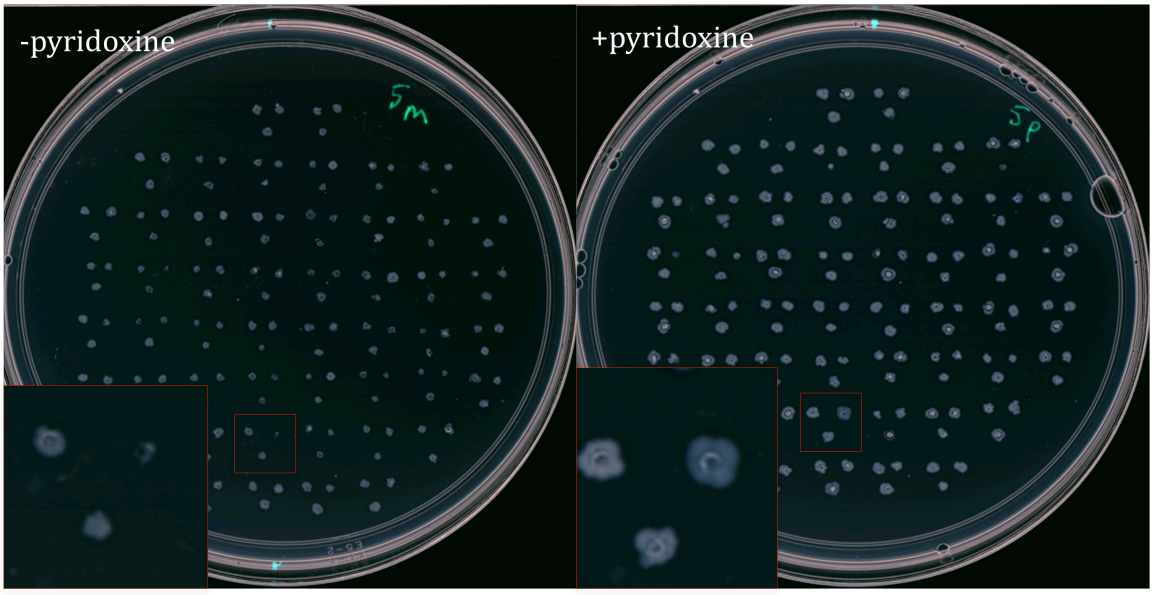


Figure 5.2. Single colonies of the insertion library patched onto plates with M9/glucose +EZ (aa/10) –Gly complex medium +/- pyridoxine on plates A/B, respectively. Cells grown from three insertion mutants are enlarged, where the upper right insertion mutant grows only in the presence of pyridoxine.

I chose to examine the phenotypes of 24 strains that formed larger sized single colonies in the presence than in the absence of pyridoxine. These strains were re-purified on LB, whereafter single colonies were washed in PBS and re-streaked on M9/glucose EZ (aa/10) -Gly +/- pyridoxine (Figure 5.3). Two strains were completely unable to grow on this medium in the absence of pyridoxine, but were rescued in its presence, as was apparent with strain #5 from Figure 5.3. Several other strains that grew in the presence of pyridoxine showed unusual, restricted growth in

its absence, as demonstrated by strain #4 from Figure 5.3. In the next step, I will locate Tn5 in one of the strains with the former phenotype (the other was mishandled) and four of the strains with the latter phenotype.

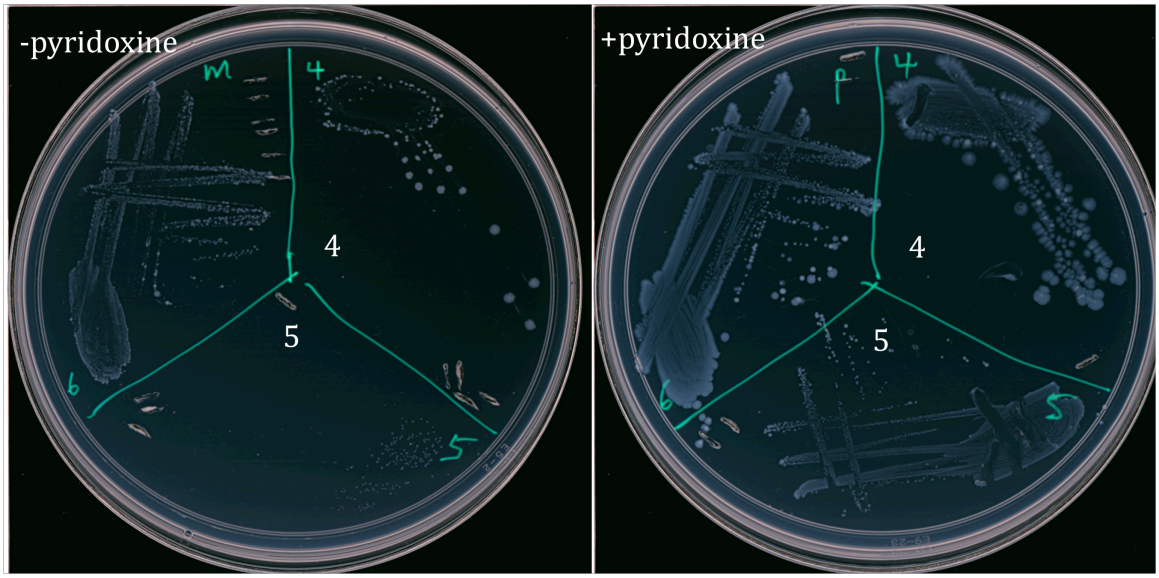


Figure 5.3. Insertion mutants #4 and 5 of 24 that were re-purified on LB, washed, and re-streaked on the M9/glucose +EZ (aa/10) –Gly complex medium +/- pyridoxine.

Using inverse PCR to locate Tn5 in the genomic DNA of select insertion mutants

Transposons of known sequence can be located in genomic DNA by digesting the DNA, circularizing the fragments, and using PCR with primers that anneal to either end of the transposon, and which are directed outward from the transposon, to amplify the flanking genomic DNA (Figure 5.4) (Ochman *et al*, 1988). These portions of genomic DNA can then be sequenced using the primers that were used to generate the PCR product.

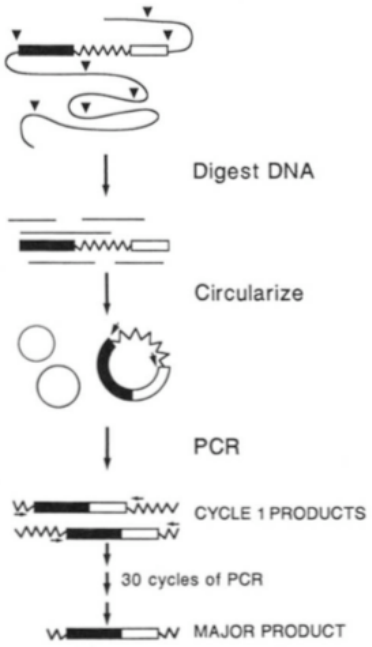


Figure 5.4. From Ochman *et al*, 1988, diagramming the use inverse PCR to locate the insertion site of transposon of known sequence in genomic DNA.

I chose to locate Tn5 in five insertion strains. After isolating the genomic DNA from these strains, I digested the DNA in multiple ways using either EcoRI, KpnI, or NcoI. I purified the digested DNA generated from each of the 15 insertion mutant/restriction enzyme combinations and circularized the DNA using T4 DNA Ligase. The DNA was once again purified and then used as a PCR template with primers that anneal to the ends of the transposon, positioned to amplify the flanking genomic DNA. The PCR product was visualized after electrophoresis on a 1% agarose gel and of the 15 gDNA/restriction enzyme combinations, inverse PCR was successful in at least one instance with each of the five insertion strains. In the future, I will purify and sequence the these PCR products, which will allow me to identify the position of Tn5 in each insertion mutant.

Discussion

In this chapter, I used the Tn5 transposon in genome-wide tranposon mutagenesis to generate an insertion library of 2,500 cells from the Δ*pdxB* Δ*ltaE* strain, transformed with *hisB* on the pTrcHisB vector. I screened 750 insertion mutants by patching cells onto a complex medium +/- pyridoxine, selecting for cells that were rescued by pyridoxine. Preceding this screen, as a control

experiment, I ensured that direct patching of the Δ*pdxB* Δ*ltaE* strain onto the complex medium without washing cells to remove residual sources of pyridoxine did not result in background growth. After two days of growing cells transformed with either *hisB* or the blank pTrcHisB vector, there was a clear distinction between the two, indicating that the method of direct patching is suitable for differentiating between cells that can and cannot synthesize PLP.

After re-screening 24 insertion mutants by re-purification on LB, washing cells in PBS to remove residual sources of PLP, and plating on the complex medium +/- pyridoxine, I found two strains that grew only in the presence of pyridoxine and several other strains that grew abnormally in the absence of pyridoxine, but not in the presence of pyridoxine. I am currently in the process of locating the Tn5 insertion site in one strain with the former phenotype and four strains with the latter phenotype.

Once I complete this process, the entire protocol used in this chapter can be optimized and used for the screening of thousands of insertion mutants. Considering that a screening of 750 insertion mutants identified two strains with clear PLP auxotrohpy, as well as several more that have an abnormal phenotype in the absence of PLP, a screening with thousands of cells could generate tens of interesting insertion mutants, each of which could be used to provide information about the serendipitous pathway used to complement the Δ*pdxB* Δ*ltaE* strain during *hisB* over-expression.

Methods

Genome-wide transposon mutagenesis

Genome wide transposon mutagenesis of the Δ*pdxB* Δ*ltaE* strain transformed with *hisB* on the pTrcHisB expression vector was conduced *in vivo* using the EZ-Tn5<Kan-2>Tnp Transposome (Epicentre). The transposome was electroporated into competent cells of the Δ*pdxB* Δ*ltaE* strain, previously transformed with the *hisB* construct, prepared using 10% glycerol (Sambrook *et al*, 1989). Cells were immediately recovered in 1 mL SOC medium and then incubated with shaking at

37 ºC for 1 hr. Next, this culture was spread in 50 µL aliquots on plates of LB with kanamycin, ampicillin, and chloramphenicol. These plates were grown at 37 ºC, yielding a library of approximately 2,500 cells.

Screening the insertion library

Single colonies of the insertion library isolated on LB were screened for PLP auxotrphy by using autoclaved toothpicks to directly patch the same colony onto gridded plates in the following order: i) M9/glycose EZ-gly (aa/10) minus pyridoxine; ii) M9/glucose EZ-gly (aa/10) plus pyridoxine; and iii) LB—which served as a positive control. Cells were then grown for two days at 37 ºC. The plates were inspected for cells whose growth was rescued by pyridoxine. 24 strains with this phenotype were re-streaked by growing the strains on LB with kanamycin, ampicillin, and chloramphenicol and then washing single colonies 5X in 50 µL PBS and plating on EZ-gly (aa/10) +/- pyridoxine. These plates also contained 0.1 mM IPTG and the appropriate antibiotics.

Using inverse PCR to locate Tn5 in select insertion mutants

I isolated genomic DNA from five of the strains from the second phenotypic screen using the Wizard Genomic DNA Purification Kit (Promega). Next, I digested the genomic DNA in three ways with either EcoRI, KpnI, or NcoI, at 37 ºC in a final volume of 20 µL with 4, 2, or 1 units/µg gDNA of EcoRI, KnpI, or NcoI (Invitrogen), respectively, and 5 µg of gDNA. As digestion was not complete after two hours, as determined by visualizing the digestion product after electrophoresis on a 1% gel, reactions were carried out overnight.

After purifying DNA from the digestion product using the E.Z.N.A. Cycle Pure Kit (Omega Bio-Tek), the DNA was ligated overnight at 16 ºC in a final volume of 500 µL with 0.5 µg DNA and 2 U/µL T4 DNA Ligase (Invitrogen). Next, the DNA was precipitated with ethanol and collected by centriguation.

Finally, I used 2.5, 0.25, or 0.025 µL of the purified ligation product as a DNA template in PCR with 0.5 uM of the primers that anneal to either end of Tn5 (Table 5.1) in a final volume of 25

μL. PCR was conducted using 30 cycles of denaturation at 94 ºC for 30 seconds, primer annealing at 54 ºC for 30 seconds, and extension by the Taq polymerase (New England BioLabs) at 68 ºC for 1 minute. The 30 cycles were preceded by 2 minutes of denaturation at 94 ºC and followed by 5 minutes of extension at 72 ºC. PCR products were visualized after electrophoresis on a 1% agarose gel with ethidium bromide.

Table 5.1. Primers that anneal to either ends of the Tn5 transposon.

	Primers (5' to 3')
Forward	ACCTACAACAAAGCTCTCATCAACC
Reverse	GCAATGTAACATCAGAGATTTTGAG

Antibiotics

When appropriate, kanamycin was included at 50 μg/mL and was used to select cells with *kan*^{r} at the Tn5 locus; ampicillin was included at 100 μg/mL and was used to select cells with *amp*^{r} from pTrcHisB vector; and chloramphenicol was included at 35 μg/mL and was used to select for cells with *cm*^{r} at the *ltaE* locus.

Works cited

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of Escherichia coli K12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006.0008

Chung CT, Niemela SL, Miller RH (1989) One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci USA* **86**: 2172-2175

Copley SD (2011) Toward a systems biology perspective on enzyme evolution. *Journal of Biological Chemistry* **287**: 3-10

Glynn SE, Baker PJ, Sedelnikova SE, Davies CL, Eadsforth TC, Levy CW, Rodgers HF, Blackburn GM, Hawkes TR, Viner R, Rice DW (2005) Structure and echanism of imidazoleglycerol-phospahte dehydratase. *Structure* **13**: 1809-1817

Kim J, Kershner JP, Novikov Y, Shoemaker RK, Copley SD (2010) Three serendipitous pathways in *E. coli* can bypass a block in pyridoxal-5'-phospahte synthesis. *Mol Syst Biol* **6**: 2010.88

Mukherjee T, Hanes J, Tews I, Ealick SE, Begley TP (2011) Pyridoxal phosphate: biosynthesis and catabolism. *Biochimica et Biophysica Acta* **1814**: 1585-1596

Neidhardt FC, Bloch PL, Smith DF (1974) Culture medium for enterobacteria. *J Bacteriol* **119**: 736-747

Ochman H, Gerber AS, Hartl DL (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**: 621-623

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning; a laboratory manual, 2nd ed.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.